Picking Up Speed in Understanding: Speech Processing Efficiency and Vocabulary Growth Across the 2nd Year

Anne Fernald, Amy Perfors, and Virginia A. Marchman
Stanford University

To explore how online speech processing efficiency relates to vocabulary growth in the 2nd year, the authors longitudinally observed 59 English-learning children at 15, 18, 21, and 25 months as they looked at pictures while listening to speech naming one of the pictures. The time course of eye movements in response to speech revealed significant increases in the efficiency of comprehension over this period. Further, speed and accuracy in spoken word recognition at 25 months were correlated with measures of lexical and grammatical development from 12 to 25 months. Analyses of growth curves showed that children who were faster and more accurate in online comprehension at 25 months were those who showed faster and more accelerated growth in expressive vocabulary across the 2nd year.

Keywords: infant language comprehension, infant speech processing, lexical development, eye-tracking, online measures

Children in the early stages of learning a language are often credited with "acquiring" new vocabulary, as if words come one by one into the child's possession. When we speak of acquiring something like a piano or a piece of property, the emphasis is on ownership, an odd way to characterize the complex and incremental processes involved in word learning. However, we also speak of acquiring skills, such as playing the piano, in which the emphasis is on gradual mastery rather than possession. It is increasingly evident that learning to recognize, understand, and speak a new word appropriately is a gradual process. Not only do infants respond meaningfully to more and more words over the 2nd year, they also respond with increasing speed and efficiency to each of the words they are learning. That is, rather than "acquiring" a new word in an all-or-none fashion, they get better at recognizing and interpreting the same word in more diverse and challenging contexts.

Because comprehension is a mental activity not easily observable in infants' spontaneous behavior, the gradual emergence of understanding has been difficult to study with precision. However, with the refinement of procedures that track listeners' eye movements as they scan a visual array in response to speech, a technique used widely in research with adults (Tanenhaus, Magnusen, Dahan, & Chambers, 2000), it is now possible to monitor the time course of spoken language understanding in very young children. Using a looking-while-listening procedure with infants from 15 to 24 months of age, Fernald, Pinto, Swingley, Weinberg, and McRoberts (1998) found that speed and accuracy in spoken word understanding increase dramatically over the 2nd year. In that study, infants looked at pictures of objects while listening to speech naming one of the objects with a familiar word. Whereas 15-month-olds responded inconsistently and shifted their gaze to the correct picture only after the end of the target word, 24-month-olds were faster and more reliable in their responses, initiating a shift in gaze midway through the target word based on partial phonetic information. These results showed that over the same period when most children experience rapid growth in lexical production, that is, the "vocabulary spurt" (Bloom, 1973; Ganger & Brent, 2004; Goldfield & Reznick, 1990), they also become much more efficient in interpreting familiar words in fluent speech.

How are these changes in receptive language abilities over the 2nd year related to the rapid development in expressive skill that occurs in the same time period? We explored this question in a longitudinal study of receptive and expressive language abilities in English-learning infants from 12 to 25 months of age. The first goal of this research was to extend Fernald et al.'s (1998) cross-sectional findings by tracking developmental changes in the speed and accuracy of spoken word recognition in the same infants across the 2nd year. The second goal was to determine whether recently developed online measures could be used to study individual differences in the development of speech-processing abilities by asking whether infants who are faster and more accurate in word recognition at one age are also faster and more accurate at other ages as well. The third and most important goal in this

research was to determine how the development of competence in spoken language understanding relates to development in other domains of linguistic competence, such as growth in expressive vocabulary and the emergence of grammatical abilities. We motivate these goals in the introduction by comparing speech-processing skills in adults and children and exploring what is known about how these abilities relate to early lexical development.

Efficiency in Spoken Word Recognition by Adults

The effortlessness with which adults make sense of speech belies the complexity of the task that faces the infant. To follow a conversation, skilled listeners must integrate acoustic information with linguistic and contextual knowledge, processing strings of speech sounds at rates of 10 to 15 phonemes per second (Cole & Jakimik, 1980). The ability to process speech continuously is central to this remarkable efficiency. By making use of phonetic information as it becomes available, adults can identify spoken words very rapidly. Online measures of spoken word recognition reveal that listeners evaluate hypotheses about word identity incrementally, on the basis of what they have heard up to that moment (Marslen-Wilson & Zwitserlood, 1989). For example, the word onset /ar/ activates numerous English words including *art, arbor, ardent, aardvark, arduous,* and others consistent with the initial phonetic information. When the listener hears /ard/, most of these candidates can be eliminated; then, when the /v/ is heard, the word *aardvark* is uniquely specified even before the final syllable. Efficient processing of word-initial information can facilitate rapid decisions about the identity of many spoken words (Grosjean, 1985). Moreover, activation of the possible meanings of a spoken word typically begins within 150 ms of word onset (Zwitserlood, 1989). If the listener could only process one phoneme at a time and each sound in the sequence was unexpected, recognizing words in fluent speech would be impossible. Similarly, the listener who interprets words one at a time is not able to follow the meaning of fluently spoken sentences, a discouraging experience familiar to anyone who studies a second language by memorizing lists of words but with little experience hearing them strung together meaningfully in speech.

Efficiency in Speech Processing by Infants

Research on early speech processing skills has shown that infants in the 1st year attend to sound patterns relevant to language structure. By 6 months, they are attuned to the phonological system of the ambient language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994) and use language-specific parsing strategies to identify word-size units in speech (Jusczyk, 1997). When briefly familiarized with strings of nonsense syllables, 8-month-olds notice which syllables co-occur (Saffran, Newport, & Aslin, 1996). Such accomplishments have been cited as evidence for early "word recognition," although this selective response to familiar syllable sequences can occur without any association between sound patterns and meanings. By the end of the 1st year, children begin to associate sound patterns with meanings, speaking a few words and appearing to understand many more. However, the processes involved in comprehension are only partially and inconsistently apparent in children's behav-

ioral responses to speech in everyday situations and thus are less accessible to observation than developments in speech production.

Studies of adult spoken language understanding rely on online measures that monitor the time course of listeners' responses to speech (e.g., Marslen-Wilson, 1987; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1996). Until recently, most developmental studies of infants' ability to identify referents of spoken words have had to rely on offline measures, assessing responses made after the offset of the speech stimulus that do not tap into the real-time properties of understanding. Because comprehension occurs automatically without time for reflection, it is important to examine listeners' interpretation *during* speech processing and not just afterward. When the preferential-looking methods used widely in infancy research were extended to studies of language comprehension, researchers began to use summary measures of looking time to assess infants' lexical knowledge (e.g., Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Schafer & Plunkett, 1998). However, because processing efficiency was not their focus, studies using these techniques did not capture time-course information about speed of response as was standard in the adult literature. With further refinement, measures of children's eye movements as they look at pictures while listening to speech now make it possible to assess the speed and efficiency of interpretation as the sentence unfolds (e.g., Fernald, McRoberts, & Swingley, 2001; Swingley & Aslin, 2002; Trueswell, Sekerina, Hill, & Logrip, 1999). Fernald et al. (1998) provided the first evidence for age-related changes in the speed and accuracy of spoken word recognition by English-learning infants, research that has now been extended to Latino children learning Spanish (Hurtado, Marchman, & Fernald, 2005). Other studies using this looking-while-listening procedure have shown that 2-year-olds interpret spoken language incrementally, associating a familiar word with the appropriate picture after hearing only the first 300 ms of the speech signal (Fernald, Swingley, & Pinto, 2001; Swingley, Pinto, & Fernald, 1999). Moreover, young English learners make use of prosodic and distributional information to anticipate the upcoming target word at the end of the sentence (Fernald & Hurtado, in press; Thorpe & Fernald, in press). Researchers are also using online measures of eye movements to explore syntactic processing by preschool children (e.g., Snedeker & Trueswell, 2004; Song & Fisher, 2005). These developmental studies of real-time language comprehension reveal that as children are building a functioning linguistic system, they become increasingly efficient in interpreting the speech they hear.

Relation of Receptive and Expressive Competence in Early Language Development

In the diary studies that provided the first systematic observations of early language development, Tiedemann (1787/1927) and Darwin (1877) both reported that their infants first showed signs of comprehension several months before beginning to produce recognizable words. Subsequent studies based on larger samples and more reliable measures have confirmed these early estimates, showing that comprehension is first evident around 8 months with word production beginning around the first birthday (Bloom, 1973; Snyder, Bates, & Bretherton, 1981). Research methods have changed considerably, but children evidently have not. The most extensive study of the rate of lexical development in English-

learning children provided norms for the MacArthur–Bates Communicative Development Inventory (CDI), a set of parental report questionnaires used to assess vocabulary size and grammatical competence from 8 to 30 months of age (Fenson et al., 1993). Parents were also asked to report the words understood by the child, but only through the age of 16 months, when it becomes increasingly more difficult for parents to provide valid estimates of receptive vocabulary. Fenson et al. (1993) found that by the time children can produce 50 words, their comprehension vocabulary is typically reported to be four times as large.

Links between early comprehension and production have been investigated primarily through observational studies showing modest correlations between words understood and spoken by children in the 2nd year (Bates, Bretherton, & Snyder, 1988). In a longitudinal study of six children from 6 to 18 months, Harris, Yeeles, Chasin, and Oakley (1995) found substantial individual differences in the rate at which comprehension developed and in the lag between comprehension and production. The limited evidence for associations between early comprehension and expressive abilities only hints at a potential connection between fluency in understanding and speaking. Because estimates of receptive language have relied largely on observational methods and parental report, they reveal little about the cognitive processes that might underlie such a connection.

Several researchers have speculated as to how speech-processing skills might relate to early vocabulary growth, although there is disagreement about the likely direction of influence. Some claim that infants' early representations of word forms are imprecise and that it is vocabulary growth that forces a shift to more efficient segmentally based processing (Charles-Luce & Luce, 1990; Walley, 1993). Stager and Werker (1997) have argued that younger infants have difficulty in distinguishing phonetic detail when trying to map new word forms onto meanings but that this limitation is short-lived; thus, by the time of the vocabulary spurt infants develop additional cognitive capacities that enable them to discriminate and learn words more quickly. From another perspective, Bloom (1993) suggested that children's growing ability to use diverse cues to retrieve words in memory is a critical aspect of cognitive development leading to more rapid lexical growth. One implication of her argument is that enhanced efficiency in the information-processing skills underlying comprehension could accelerate growth in expressive language.

One way to investigate links between early speech processing skill and lexical development is to ask whether differences among children in abilities essential for language understanding are associated with differences in the size of their productive lexicons, concurrently and at earlier or later times. Studies of speech perception skills in preverbal infants have focused on group data rather than on individual differences and thus have not attempted to relate measures of early perceptual competence to other measures of linguistic development (see Aslin, Jusczyk, & Pisoni, 1998). One recent exception is a study showing that the performance of 6-month-olds in a phonetic discrimination task was correlated with size of production vocabulary in the 2nd year (Tsao, Liu, & Kuhl, 2004). In another relevant study, Werker, Fennell, Corcoran, and Stager (2002) asked whether infants' ability to learn phonetically similar novel words was correlated with reported vocabulary size. They found that those 14-month-olds who were successful in word learning had relatively larger vocab-

ularies; however, at 18 months infants were more proficient overall in learning novel words, and there was no relation between performance and vocabulary size. Werker et al. concluded that vocabulary size may predict infants' ability to use phonetic detail in word learning before the onset of the vocabulary spurt, but that this relation holds in only the earliest stages of building a vocabulary.

Three recent studies have provided preliminary evidence for a relation between speech processing skills and vocabulary size in infants in the 2nd year. Fernald et al. (2001) used the looking-while-listening paradigm to show that 18- and 21-month-olds could recognize words on the basis of partial phonetic information. Grouped by response speed, infants with faster mean reaction times (RTs) had larger production vocabularies than infants with slower response latencies. Fernald (2002) also found convergent results in a study of online understanding of verbs by 26-month-olds; those children with larger vocabularies were more efficient in using information from the verb to predict the upcoming target noun in the sentence. In a third recent study, Zangl, Klarman, Thal, Fernald, and Bates (2005) presented infants between 12 and 24 months with words that were either naturally spoken or perceptually degraded. Using a variant of the looking-while-listening procedure, Zangl et al. also found accuracy and speed in word recognition to be correlated with vocabulary size.

Although these studies provide preliminary evidence that efficiency in comprehension may be associated with early lexical growth, two other recent studies found no association between vocabulary size and success in online speech processing. When Swingley and Aslin (2000, 2002) investigated the ability of 14- and 18-month-olds to identify words pronounced correctly and incorrectly, infants' performance was uncorrelated with level of vocabulary development. Inconsistencies among these findings may be explained by differences in study design and age of participants. By following a large sample of children and examining their emerging speech processing skills in relation to measures of lexical and grammatical development at multiple time points, we aimed to provide some clarification in the present research.

Overview of Goals and Design of the Research

The longitudinal design of this research enabled us to investigate the development of children's receptive language abilities from three different perspectives. One goal was to replicate the finding that efficiency in online speech processing increases across the 2nd year (Fernald et al., 1998). Developmental changes in the speed and accuracy of spoken word recognition were assessed by testing 59 infants four times each in the looking-while-listening procedure, at 15, 18, 21, and 25 months of age. Because speed of processing is operationalized in terms of the latency with which infants shift their gaze to the named target picture, it was not clear from the original cross-sectional findings whether the age-related decrease in RT was in fact related to linguistic processing or whether infants simply made faster eye movements as they grew older. This latter explanation seemed unlikely, given that by 12 months of age infants' mean visual reaction time (VRT) approaches adult values (Canfield, Smith, Brezsnyak, & Snow, 1997; Rose, Feldman, Jankowski, & Caro, 2002). However, to explore this possibility, we added a control measure to assess VRT in a nonlinguistic task. We expected to find an increase in accuracy and

a decrease in RT in the word recognition task across the four time points, with no appreciable change in VRT in the nonlinguistic task.

A second goal of this research was to investigate the stability of our measures of speed and accuracy in online speech processing. As all previous researchers using comparable measures have either focused on a single age group or used a cross-sectional design, nothing is known about the test–retest reliability of these measures. Do those infants who respond most reliably at any given age also respond most rapidly? Are those infants with the fastest mean RTs at one age among the fastest at later ages? Exploring such questions helps to determine whether eye-movement measures could potentially be valuable in research on individual differences in the early development of speech processing abilities.

The third goal was to establish the validity of the online measures used here and in other recent studies by asking how infants' development in processing efficiency between 15 and 25 months relates to other aspects of language competence. To what extent are speed and accuracy in speech processing associated with lexical growth and grammatical development? We addressed this question by examining relations between measures of infants' performance in online word recognition at each age with measures of vocabulary size and grammatical competence, as reported by parents on the CDI at 12, 15, 18, 21, and 25 months. As a convergent measure of lexical knowledge, the Peabody Picture Vocabulary Test—Revised (PPVT–R; Dunn & Dunn, 1981) was administered at 25 months. Growth trajectories were then analyzed to examine how the development of competence in spoken language understanding relates to patterns of vocabulary growth and the emergence of grammatical abilities across the 2nd year.

These analyses could yield three possible patterns of results with quite different developmental implications: First, it could be that processing efficiency and vocabulary growth draw on fundamentally different cognitive capacities and thus are not consistently related. Second, if individual differences in speech processing skill are to some extent stable from infancy through later childhood, consistent with numerous studies of early information processing in other domains (e.g., Colombo & Fagan, 1990), processing efficiency could drive vocabulary growth from the very beginning. In this case significant relations between speed and accuracy in spoken word recognition and later lexical development should be evident from 15 months onward. A third scenario is that relations between processing efficiency and vocabulary size are not evident at the outset but emerge over the 2nd year. This pattern of results is consistent with the view that multiple factors influence early vocabulary learning and that children with more extensive lexical knowledge benefit by developing more efficient processing strategies by the end of the 2nd year.

Method

Participants

The initial sample consisted of 63 full-term infants (37 boys, 26 girls) recruited through a university hospital. Participants were from middle-class families in which English was the primary spoken language, with 85.7% of families Caucasian, 12.7% Asian, and 1.6% Hispanic–Caucasian. The highest educational level attained by at least one parent in the family was some graduate-level education (62%), undergraduate degree (35%), and some college (3%).

Infants were tested in the laboratory at the ages of 15, 18, 21, and 25 months. Data for some subjects at each age were not included in the analyses for the following reasons: (a) fussiness (15 months: $n = 2$, 18 months: $n = 4$, 21 months: $n = 3$, 25 months: $n = 3$), (b) failure to fixate one of the stimulus pictures on at least 70% of trials (15 months: $n = 12$, 18 months: $n = 6$, 21 months: $n = 6$, 25 months: $n = 6$), (c) experimenter error (18 months: $n = 1$), and (d) missed session (18 months: $n = 1$). At 18 months, 4 participants dropped out of the study permanently because of illness ($n = 1$) or moving from the area ($n = 3$); hence, most analyses were conducted on the final sample of 59 children. Growth curve analyses included only those children ($n = 50$) for whom we had online measures of speech-processing efficiency at 25 months as well as parent-report measures of vocabulary for at least four of the five sample points between 12 and 25 months. At the 25-month test session, 10 participants who completed the word-recognition procedure did not complete the PPVT–R because of fussiness or fatigue.

Parental Report Measures of Lexical and Grammatical Development

When their infant reached the ages of 12, 15, 18, 21, and 25 months, parents were mailed the appropriate version of the CDI. At 12 and 15 months, they filled out the CDI: Words and Gestures. At 18, 21 and 25 months, they filled out the CDI: Words and Sentences. At the 12-month time point, parents returned the completed CDI by mail; at older ages, they brought the CDI when they visited the laboratory. The CDI yields several different measures of vocabulary and grammar:

Receptive vocabulary. The measure of receptive vocabulary was the number of words understood by the child as reported on the CDI: Words and Gestures at 12 and 15 months. Because the validity of parental report assessments of comprehension decreases beyond 16 months, information regarding receptive vocabulary was limited to the 12- and 15-month samples.

Expressive vocabulary. The measure of vocabulary production was the number of words that the child "understands and says," as reported by parents at 12 and 15 months (CDI: Words and Gestures) and at 18, 21, and 25 months (CDI: Words and Sentences).

Mean length of the three longest utterances (M3L). M3L is defined as the average number of morphemes in the three longest utterances spoken by the child, as reported by parents on the CDI: Words and Sentences at 18, 21, and 25 months.

Grammatical complexity. The grammatical complexity measure is designed to assess the use of word combinations and closed-class morphemes. Parents are asked to choose one of two utterances typical of children's early word combinations, for example, *Kitty go away* versus *Kitty went away* or *Doggie table* versus *Doggie on table,* that sounds most like the way their "child talks right now." The grammatical complexity score is the number of times the parents chose the second (i.e., more complex) example, assessed at 18, 21, and 25 months.

PPVT–R

At 25 months, we administered the PPVT–R after the looking-while-listening procedure, following standard procedures.[1] The child sat on the parent's lap, and the experimenter explained the task to the child by saying,

---

"Now I'm going to show you some pictures. The way this game works is, if you hear me say the name of a picture, you can show me the picture by touching it with your finger." Once the child learned the task, the procedure began. The child was shown a series of line drawings arranged four per page. One picture on each page was the designated target, with target location counterbalanced across pages. When the child had looked at the page for about 10 s, the experimenter called attention to one of the pictures by saying "Point to the [target]." If the child failed to respond within 10 s, the instruction was repeated. Testing continued until the child chose the incorrect picture on 6 of 8 consecutive pages or refused to participate further. The child's score was his or her ceiling score minus errors.

The Looking-While-Listening Procedure for Monitoring Word Recognition Online

Speech stimuli. In designing stimuli to be used at different ages, we had competing goals. On the one hand, it was important to include comparable tokens at each age to enable controlled comparison of age-related differences in the efficiency of speech processing. For this reason, infants in the Fernald et al. (1998) study were tested at all three ages on the same stimuli, consisting of four words already familiar to the youngest infants. Because this earlier study was cross-sectional, there was no concern about effects of repeated testing. However, in the present longitudinal study it was also important to make the stimulus sets increasingly challenging to keep older children engaged in the task. Because we were exploring individual differences in processing abilities, it was critical to include some stimuli known to be difficult for children at each age to avoid ceiling effects.

The speech stimuli at all ages consisted of prerecorded sentences containing target words in a familiar carrier phrase such as "Where's the [target]? Do you see it?" At 15 months, the four target words (*doggie, baby, ball, shoe*) used by Fernald et al. (1998) were each presented six times. In addition to these 24 trials, 6 filler trials were included to increase the variability of the stimulus set. At 18 months, the target words were *doggie, baby, ball,* and *car*, each presented twice as whole words and twice as truncated words (*daw, bei, baw, ka*). This manipulation was motivated by the finding that the ability to recognize familiar words using partial phonetic information develops between 18 and 21 months of age and is related to level of lexical development (Fernald et al., 2001). In addition to these 16 trials, 13 filler trials were included.[2] At 21 months, *doggie, baby, birdie*, and *kitty,* were presented both as whole words and as truncated words (*daw, bei, ber, ki*) for a total of 16 trials. On 8 additional trials, the 21-month-olds also heard *juice* and *cookie* as target words, preceded by frames in which the verb was either semantically related (*drink, eat*) or semantically neutral (*take, look at*), with four filler trials. At 25 months, the standard set of target words (*doggie, baby, ball, car*) was augmented by target words typically learned later in the 2nd year (*monkey, cow, flower, tree, animal*), presented as whole words. To diversify the sentence frames, we included adjectives (*nice, pretty*) before target words on 8 trials. Four filler trials were included.

To prepare these stimuli, a female native speaker of English produced several tokens of each sentence, matching them closely in intonation contour. These candidate stimuli were recorded on a Revox B77 tape recorder, then digitized, analyzed, and edited using the SoundEdit waveform editor on a Macintosh computer. The tokens for the final stimulus set were chosen on the basis of comparability in duration of the carrier phrase and target word in each vocalization. Preparation of the truncated words used as stimuli in the 18- and 21-month test sessions is described in Fernald et al. (2001).

Visual stimuli. Visual stimuli consisted of digitized photographs of colorful objects corresponding to the target words. Two object tokens were used for each target word at every age. All images were matched in size and brightness and were presented on a gray background on 15-in.-diagonal (38.1 cm) computer monitors. Each object served equally often as target and distracter.

Apparatus. The looking-while-listening procedure was conducted in a sound-treated room. The testing booth had three cloth-covered panels; the side panels measured 1×2 m, and the front panel measured 1×1.2 m. Mounted in front at the infant's eye level were two computer monitors, separated horizontally by 60 cm on center. The image on each monitor subtended a visual angle of approximately $12°$. The infant sat on the parent's lap 60 cm from the monitors, such that an eye/head movement of approximately $72°$ was necessary to shift gaze from one picture to the other. A curtain behind the infant's head obstructed the parent's view of the monitors while allowing the infant access to the parent during testing. A loudspeaker for presenting the auditory stimuli was on the floor between the monitors. A video camera mounted behind the front panel was focused on the infant's face. The camera was connected to a VCR in the adjacent control room where the computer controlling the experiment was also located.

Procedure. Each session began with a 20-min familiarization period during which the experimenter talked with the parents, obtained informed consent, and interacted with the child. When parent and child appeared comfortable, they were seated in the booth. Room lights were dimmed as identical pictures appeared on the monitors to attract the infant's attention. A second experimenter in the control room spoke briefly over the loudspeaker to acquaint the child with the sound source. When the child was attentive, the experimental session began. Trial types were presented in a quasirandom order, with side of presentation of target and distracter objects counterbalanced across trials. On each test trial the two pictures were shown in silence for 3 s prior to the speech stimulus, continuing for 1 s after offset of the sound stimulus. During the 1-s intertrial interval the screens were black. The entire test session lasted about 4 min.

Coding eye movements. A digital time code accurate to 33 ms was recorded onto the videotape, with a visual marker indicating the onset of the speech stimulus on each trial. The tape was digitized and coded offline by highly trained observers using custom software on a Macintosh computer. Coding was done without sound by observers blind to trial type and side of target picture. For each trial, coders analyzed the time course of the infant's gaze patterns frame by frame, noting on each frame whether the infant's eyes were oriented to the left picture, to the right picture, between the two pictures, or away from both pictures. The software aligned these data with the onset of the target word for each trial. The computer also calculated the duration of each look and indicated the time at which the infant initiated each shift in gaze.

Interobserver reliability analyses focused on those trials on which shifts in gaze from one picture to the other occurred most frequently for two reasons: First, the potential for disagreement on these trials was greater than on trials in which infants continued to fixate one picture only and, second, trials involving shifts were critical in the analysis of response speed. To prepare for the reliability analysis, an observer otherwise uninvolved in the analysis prescreened tapes for 20% of the infants at each age. This observer viewed the tape in real time to identify trials on which shifts occurred most frequently, assigning these trials to be independently coded. For the 15-month sample, 7 trials were coded for each of 13 infants; on 90.4% of the trials, shifts were judged to be within one frame of each other. At 18 months, 9 trials were coded for each of 13 infants; 88.7% of shifts were within one frame of each other. At 21 months, 8 trials for each of 14 children were coded twice; 91.4% of shifts were within one frame of each other. Finally, at 25 months, 8 trials for each of 13 children were coded twice; 90.8% of shifts were within one frame of each other.

---

[2] Given the use of partial-word trials in the 18-month test session, a greater number of filler trials with engaging novel pictures and animated speech (e.g., "Hey, look at that!") was used to reduce the overall proportion of anomalous sentences that might be confusing to the infants.

Measures of Accuracy and Speed in Online Speech Processing

Because children cannot know in advance which picture will be labeled, about half the time they will by chance be looking at the distracter picture at target-word onset (distracter-initial trials), and half the time they will already be looking at the target picture (target-initial trials). On distracter-initial trials, the correct response is to shift to the target picture, whereas on target-initial trials, the correct response is to continue looking at the target picture without shifting away. Thus, a child with perfect accuracy would shift to the target picture on 100% of the distracter-initial trials and would never shift away on target-initial trials. Shifts on all distracter- and target-initial trials were assessed within the time window from 300 to 1,800 ms following target word onset. Shifts prior to 300 ms were excluded because they presumably occurred before the child had had time to process sufficient acoustic input and mobilize an eye movement (Haith, Wentworth, & Canfield, 1993), and shifts occurring between 1,800 and 3,000 ms after target word onset were considered to be outliers that were less clearly in response to the target word.[3]

Proportion of correct shifts to target picture. This measure of correct shifts represents the number of first shifts to the target picture that occurred on distracter-initial trials within the 300–1,800 ms window following target word onset, as a proportion of all distracter-initial trials.

Proportion of incorrect shifts to distracter picture. This measure of incorrect shifts represents the number of first shifts to the distracter picture that occurred on target-initial trials within the 300–1,800-ms window following target word onset, as a proportion of all target-initial trials.

Proportion of looking time to target picture. Combining the data from both target- and distracter-initial trials, this measure of accuracy represents the total amount of time the infant spent fixating the target picture as a percentage of total time fixating either picture during the 300–1,800-ms window from target word onset.

RT. Speed of response to the spoken word was calculated as the mean latency (in milliseconds) to shift from the distracter to the target picture on all distracter-initial trials on which a correct shift occurred within the 300–1,800-ms window from target word onset. It is important to note that the mean RT for different children and for any given child at different ages could be based on different numbers of trials. One reason is that children differed in how frequently they by chance were looking at the distracter at target word onset and thus differed in the numbers of trials on which RT could be calculated. Another reason is that there were fewer trials to begin with at younger ages, and younger infants shifted less reliably than older infants. Thus, the mean RT for an infant at 15 or 18 months might reflect performance on only 2–4 trials, because the infant started on the distracter on only a few trials or failed to shift reliably from distracter to target picture. In contrast, the mean RT for the same infant at 21 or 25 months might reflect performance on 6–10 or more trials. At all ages, the mean RT scores for each child was based on at least two shifts, as data from those with only one codeable shift were excluded from the relevant analyses.

VRT Procedure

To control for individual differences and age-related changes in infants' speed in shifting from one picture to another when no language processing was involved, we longitudinally assessed RT to a peripheral visual stimulus in a nonlinguistic context. We used a modified version of the Visual Expectation Paradigm developed by Haith, Hazan, and Goodman (1988), a procedure designed to measure infants' raw RT to peripheral visual stimuli presented in silence.

Apparatus and procedure. Upon completion of the word-recognition test at 15, 18, and 21 months, infants participated in the VRT procedure. The infant sat on the caregiver's lap at a distance of 72 cm from a single 17-in.-diagonal (42.5 cm) computer monitor positioned at the infant's eye level. The caregiver was instructed to close her or his eyes during presentation of the stimuli. A video camera recorded the child's eye movements.

Visual stimuli. Infants viewed computer-generated pictures that alternated in a random sequence on the left and right sides of the computer monitor (see Haith et al., 1988). Each picture subtended a visual angle of approximately $2.3°$; the angle of the saccade required to shift fixation from one side to the other was $15.8°$. The stimuli consisted of 10 pictures (e.g., diamond, bulls eye, triangle), each 4-cm square with a different dynamic pattern (e.g., spinning, blinking, expanding). On each trial, one picture was displayed for 700 ms on the right or left side of the screen. Infants saw a total of 20 pictures displayed in a random sequence with no interstimulus interval, with the constraint that the stimulus appeared on the same side no more than three trials in a row. This procedure lasted approximately 2.5 min.

Coding eye movements. A digital time-code accurate to 33 ms was recorded onto the videotape, with a visual marker indicating the stimulus onset on each trial. The tape was digitized and coded off-line by trained observers using custom software. For each trial, coders analyzed the time course of the infant's gaze patterns frame by frame, noting on each frame whether the infant's eyes were oriented to the left of the screen, to the right of the screen, or away from the screen. Eye movements were coded relative to the onset of the visual stimulus.

VRT. Mean VRT was defined as the average latency to begin an eye movement toward the stimulus after it appeared, based on those trials on which the stimulus switched from one side to the other. The calculation of VRT was based on those eye-movement latencies ≥ 133 ms (Canfield et al., 1997), with a cutoff of 700 ms corresponding to the stimulus offset (Dougherty & Haith, 1997). Thus, the mean VRT measure for each infant represents the average of all response latencies between 133 and 700 ms.[4]

Results

Following an initial section outlining our data analysis strategy, the results are presented in five sections related to the major research questions. In the second section, we describe changes in speech-processing efficiency between 15 and 25 months in our longitudinal sample of English-learning children. The third section focuses on interrelations between measures of speed and accuracy in online comprehension and their stability over time. The fourth section explores the relation of speed of spoken word recognition to VRT in a nonlinguistic task. The fifth section examines infants' lexical and grammatical development from 12 to 25 months in relation to the emergence of speed and accuracy in spoken word recognition. The final section investigates how the development of efficiency in online speech processing by 25 months relates to patterns of growth in children's production vocabulary across the 2nd year of life.

Rationale for Data Analysis Strategy

Although the verbal stimuli at each age consisted of words the children were highly likely to know, there was variability among individuals in level of vocabulary development at every time point.

---

[3] The cutoff used most frequently in previous studies is 367 ms (e.g., Fernald et al., 2001; Swingley & Aslin, 2000), an "educated guess" as to the minimum time required to identify a word and mobilize an eye movement. Since none of these studies included children as old as 25 months, we chose the more conservative cutoff of 300 ms for use with our wider age range so as not to disadvantage older infants by eliminating very fast responses from the analysis.

[4] For further details on the coding criteria in this procedure, see Dougherty and Haith (1997).

As children may not have been equally familiar with all the words in the stimulus sets at each age, an important consideration was whether to focus analyses of changes in processing efficiency on children's responses to all of the target words or to limit them to those words reported to be in the child's vocabulary. A possible objection to including all available data in the analyses is that correlations between speed and accuracy in word recognition and either age or vocabulary size could be an artifact of individual differences in children's knowledge of the words we happened to test them on. For example, if older infants perform better in this task than younger infants, this could reflect an age-related increase in processing efficiency or it could simply be that older children "know" a greater proportion of the target words. To the extent that children respond randomly to unfamiliar target words, younger children and children with smaller vocabularies will miss more trials and thus perform less well overall than older children or children with larger vocabularies. In this case, any significant correlations between efficiency in word recognition and both age and vocabulary size could be attributable to task demands rather than reflecting an interesting relation between online measures of processing efficiency and offline measures of lexical development.

One way to address this concern is to limit the analyses of processing efficiency to words reportedly "known" by the child at each age, thus eliminating trials on which responses may have been random because the target word was unfamiliar. However, this approach has disadvantages as well. First, as in all experiments with infants, the number of trials is low, and dropping data can increase variability and reduce statistical power. Focusing on subsets of target words necessitates dropping trials for individual children at each age and disproportionately dropping participants with very low vocabulary scores. Second, it is difficult to know which criterion of vocabulary knowledge to apply when excluding trials. Whereas the CDI: Words and Gestures used at 15 months solicited parents' estimates of words that the child "understands" as well as "understands and says," the CDI: Words and Sentences used at 18, 21, and 25 months only included the production measure. Although words produced can be used as an index of "words known" by the child from 18 months on, this is a very conservative measure, as infants typically demonstrate understanding of words before they can speak them. Given arguments for and against each of these approaches, we report both types of analysis in the first section to demonstrate the comparability of the results.

Changes in the Speed and Accuracy of Spoken Word Recognition From 15 to 25 Months

One goal of this research was to extend the cross-sectional findings of Fernald et al. (1998) with a longitudinal design. To compare speech-processing efficiency in the same children at different ages, we analyzed infants' speed and accuracy in word recognition at 15, 18, 21, and 25 months using repeated measures analyses of variance. For both measures, the first analysis included all target words for which data were available (all words), whereas the second analysis included just those target words that the child was reported to understand at 15 months or to understand and produce at 18, 21, and 25 months (words understood and produced).[5]

RT. The mean RT was calculated for each child at each age by averaging response latencies on those distracter-initial trials on

which a correct shift occurred in the 300–1,800 ms interval following word onset. In the all-words analysis, mean RT decreased significantly with age, as shown in Figure 1, with an overall decrease of 210 ms across the 10-month period (15 months: $M = 981$ ms, $SD = 223$ ms; 18 months: $M = 962$ ms, $SD = 215$ ms; 21 months: $M = 802$ ms, $SD = 186$ ms; 25 months: $M = 771$ ms, $SD = 128$ ms), $F(3, 75) = 14.4$, $p < .001$. Note that the variance in response speed also decreased substantially by 25 months, suggesting that indices of processing speed may be less stable at the younger than older ages. The second analysis focused only on words understood and produced and yielded parallel results, also shown in Figure 1. Mean RT decreased significantly with age, with an overall decrease of 215 ms across the 10-month period (15 months: $M = 984$ ms, $SD = 232$ ms; 18 months: $M = 943$ ms, $SD = 227$ ms; 21 months: $M = 792$ ms, $SD = 188$ ms; 25 months: $M = 769$ ms, $SD = 138$ ms), $F(3, 60) = 9.7$, $p < .0001$. A decrease in variability was also observed, as in the all-words analysis.

Accuracy. Because the behavior indicating a correct response differs as a function of where the child happens to be looking at the onset of the target word, we first evaluated responses on distracter-initial and target-initial trials separately. For each individual at each age, two measures were calculated: the mean proportion of correct shifts from the distracter to the target picture and the mean proportion of incorrect shifts away from the target to the distractor. These correct and incorrect shifts were then compared in a 4 (age) × 2 (trial type: target-initial vs. distracter-initial) repeated measures analysis of variance. As shown in Figure 2, the data from the all-words analysis revealed significant main effects of age, $F(3, 108) = 10.1$, $p < .001$, and trial type, $F(1, 36) = 217.8$, $p < .001$, as well as an Age × Trial Type interaction, $F(3, 108) = 8.3$, $p < .001$. Follow-up tests indicated that the tendency to shift correctly to the target picture on distracter-initial trials increased significantly with age (15 months: $M = .52$, $SD = .31$; 18 months: $M = .56$, $SD = .32$; 21 months: $M = .73$, $SD = .25$; 25 months: $M = .86$, $SD = .16$), $F(3, 108) = 16.3$, $p < .001$. In contrast, the tendency to shift incorrectly from the target to the distracter picture did not change significantly with age (15 months: $M = .26$, $SD = .23$; 18 months: $M = .32$, $SD = .24$; 21 months: $M = .31$, $SD = .20$; 25 months: $M = .36$, $SD = .19$), $F(3, 108) = 1.4$, *ns*.

The second analysis of correct and incorrect shifts, which focused on only those trials on which the target words were understood and produced by the child, yielded comparable results to the all-words analysis. There were significant main effects of age, $F(3, 93) = 7.2$, $p < .0001$, and trial type, $F(1, 31) = 136.8$, $p < .0001$, as well as a significant Age × Trial Type interaction, $F(3, 93) = 6.3$, $p < .001$. Follow-up repeated measures contrasts indicated again that the tendency to shift correctly to the target picture on

---

[5] The analyses that included only trials with "known" target words required dropping from zero to four trials from the data of individual participants, that is, those trials with target words *not* reported as either "understood" (at 15 months) or "understood and produced" (at 18, 21, and 25 months) on the CDI were dropped. This also required eliminating some subjects altogether who for various reasons did not have enough codeable data after these trials were eliminated. The numbers of children dropped from the analyses at each age, for RT and accuracy, respectively, were 15 months: *n*s = 2, 0; 18 months: *n*s = 8, 6; 21 months: *n*s = 4, 3; 25 months: *n*s = 1, 0.
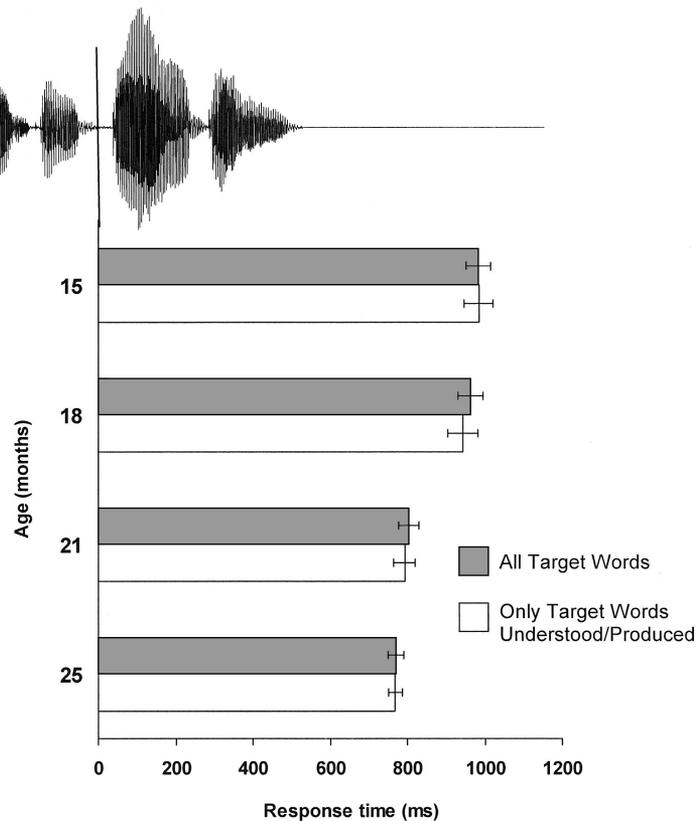
Figure 1. Mean latencies (with *SE*) to initiate a shift in gaze from the distracter to the target picture, measured from the beginning of the target word, shown for children at 15, 18, 21, and 25 months. Gray bars show means for analyses based on all available trials; white bars show means for analyses based on the subset of trials for which the target words were understood and produced by the child at each age according to parental report. Note that the means for these subsets are depicted side by side for visual comparison only; they were not compared statistically. The graph is aligned with the amplitude envelope of one of the stimulus sentences.

distracter-initial trials increased significantly with age (15 months: $M = .53$, $SD = .30$; 18 months: $M = .54$, $SD = .36$; 21 months: $M = .71$, $SD = .29$; 25 months: $M = .86$, $SD = .15$), $F(3, 93) = 12.5$, $p < .001$, whereas the tendency to shift incorrectly from the target to the distracter picture did not (15 months: $M = .28$, $SD = .25$; 18 months: $M = .35$, $SD = .27$; 21 months: $M = .35$, $SD = .24$; 25 months: $M = .35$, $SD = .20$), $F(3, 93) = 0.7$, *ns*.

An alternative approach to operationalizing accuracy was to measure the total time the child fixated the target picture during the relevant time window following target word onset. The proportion of looking time to the target picture was calculated on each trial and then averaged for each infant at each age. In the all-words analysis, mean looking time increased significantly with age (15 months: $M = .62$, $SD = .12$; 18 months: $M = .61$, $SD = .17$; 21 months: $M = .72$, $SD = .13$; 25 months: $M = .75$, $SD = .12$), $F(3, 108) = 9.2$, $p < .001$. A similar age-related increase in accuracy was observed when we analyzed only words understood and produced trials (15 months: $M = .63$, $SD = .13$; 18 months: $M = .59$, $SD = .20$; 21 months: $M = .70$, $SD = .19$; 25 months: $M = .76$, $SD = .12$), $F(3, 93) = 6.1$, $p < .001$. Thus, older children

spent relatively more time fixating the target picture, regardless of where they were looking at target word onset.

Both of these measures of accuracy are valid—the first based on correct shifts, the second based on time spent fixating the target picture. However, because the looking time measure combines data from both target- and distracter-initial trials, it is partially redundant with the proportion of correct shifts measure. Moreover, the differential patterns of response observed on distracter- and target-initial trials suggest that the ability to shift correctly from distracter to target picture may undergo greater developmental change during this period than correct responses that require the child merely to keep looking at the target, as shown in Figure 2. For this reason, proportion of correct shifts was used as the measure of accuracy in the remaining analyses.

The first three analyses of developmental changes in RT and accuracy gave very similar results, regardless of whether we included all available trials for each child or only those trials with target words understood and produced by the child. All of the remaining analyses were also conducted using both trial sets; however, because similar patterns of results were obtained in all
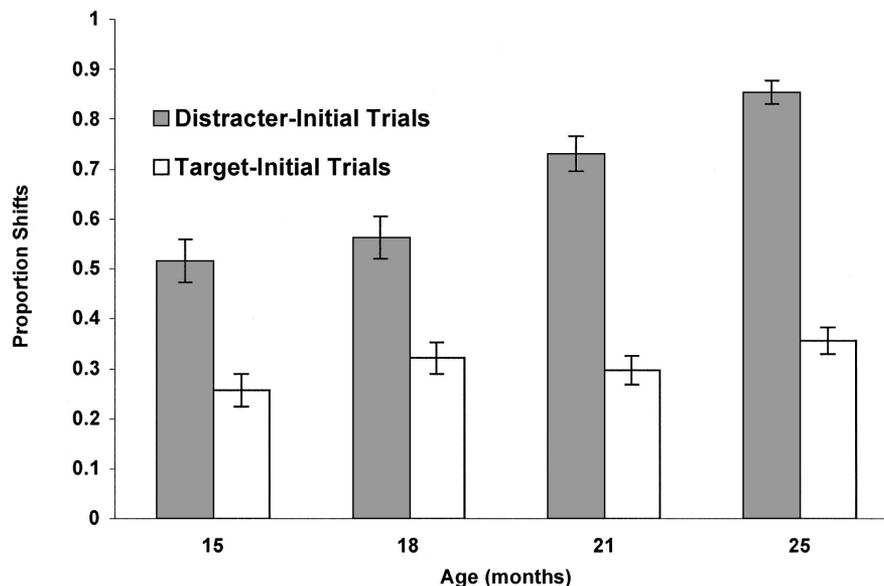
Figure 2. Mean proportion (with *SE*) of shifts away from picture initially fixated on distracter-initial and target-initial trials shown at each age. Gray bars represent correct shifts (the measure of accuracy); white bars represent incorrect shifts.

cases, we only report analyses that were based on responses to all target words.

Stability and Intercorrelation of Online Speech-Processing Measures

Month-to-month correlations for RT and accuracy from 15 to 25 months are shown in Table 1. Because multiple tests of correlation were conducted, all probability values reported here and in subsequent tables were adjusted using a Bonferroni correction based on the number of comparisons within each variable. The correlation between mean RTs at 18 and 21 months was reliable, although it was only marginally significant between 21 and 25 months with the adjusted criterion value ($p < .07$) and was not reliably stable between 15 and 18 months. Month-to-month measures of accuracy also showed moderate stability across the period, with significant associations between mean proportion of correct shifts at 18 and 21 months and at 21 and 25 months, although not at 15 and 18 months.[6]

Table 1
Month-to-Month Correlations for Three Online Measures

Age	RT (ms)[a]	Accuracy[b]	VRT (ms)[c]
15 to 18 months	.21	.27	.56**
18 to 21 months	.39*	.49**	.71**
21 to 25 months	.31	.34*	—

Note. RT = reaction time; VRT = visual reaction time. Dash indicates data are not available.
[a] Mean response latency to shift to the target picture.   [b] Mean proportion of correct shifts to the target picture.   [c] Mean response latency to orient to a peripheral visual stimulus with no speech stimulus.
* $p < .05$.   ** $p < .01$.

To explore the relation of speed and accuracy, we asked whether the proportion of correct shifts was negatively correlated with mean RT at each age. Correlations between speed and accuracy in word recognition were significant at 18 months ($r = -.44$, $p < .01$), 21 months ($r = -.43$, $p < .01$), and 25 months ($r = -.53$, $p < .01$). The correlation at 15 months was in the same direction, but did not reach significance with the adjusted criterion value ($r = -.31$, $p < .09$).[7]

Relation of RT in Online Speech Processing to VRT in a Nonlinguistic Task

The VRT measure was designed to assess speed of orienting in a nonlinguistic task. As in the looking-while-listening procedure, the VRT task required infants to shift rapidly from one picture to another. The question motivating this control measure was whether the increase in speed of orienting to a named picture over the 2nd year could be accounted for by a more general increase in speed of oculomotor response during visual orienting. A comparison of mean VRT (in milliseconds) at 15 months ($M = 342$, $SD = 45$), 18 months ($M = 335$, $SD = 62$), and 21 months ($M = 332$, $SD = 61$) revealed no significant changes with age. Mean VRT was highly stable within individuals over time: Correlations between VRT scores at 15 and 18 months and at 18 and 21 months were both reliable (see Table 1). However, speed of orienting in the VRT procedure was not related to speed of orienting to a named picture in the looking-while-listening procedure. Correlations between children's mean VRT in the nonlinguistic task and their

---

[6] The analyses in this section based on only those trials with "known" target words yielded comparable results to the all-words analyses reported in the text.

[7] See footnote 6.

mean RT in the word-recognition task at 15 months ($r = .01$), 18 months ($r = -.06$), and 21 months ($r = -.17$) were haphazard and nonsignificant. This dissociation between nonverbal and verbal measures suggests that the developmental decreases in RT in the looking-while-listening procedure reflect improved efficiency in interpreting visual stimuli in the context of linguistic reference rather than speed in visual orienting more generally, at least as reflected in this low-level perceptual control task.

Relations Between Speed and Accuracy in Online Word Recognition and Off-Line Measures of Lexical and Grammatical Development

The most important goal of this study was to determine how the development of spoken word recognition over the 2nd year relates to other domains of linguistic competence. Table 2 presents descriptive statistics for measures of vocabulary and grammar from the CDI and PPVT–R. The number of words comprehended increased between 12 and 15 months, $t(55) = 11.5$, $p < .001$, and at both time points children were reported to understand more words than they produced ($p < .001$). Vocabulary production increased substantially over the 2nd year, with children moving from < 10 words on average at 12 months to almost 400 words at 25 months, $F(4, 196) = 159.7$, $p < .001$. These scores fell near the median percentile levels for vocabulary production ($M = 48.8\%$) based on normative data reported in Fenson et al. (1993). Improvements were also observed in both grammar measures from 18 to 25 months: complexity, $F(2, 106) = 53.7$, $p < .001$; M3L, $F(2, 88) = 84.4$, $p < .001$. Again, these scores were consistent with expected levels for these measures, indicating that children in this sample were making progress in lexical and grammatical development comparable to that seen in larger datasets. The range of individual differences was also consistent with the variation typical of children in this age range (Fenson et al., 1993).

To assess relations between online measures of speech-processing efficiency and offline measures of language growth, we first examined concurrent correlations at each age. As shown in Table 3, mean RT in the looking-while-listening procedure was significantly correlated with all vocabulary and grammar measures as well as the PPVT–R at 25 months, although concurrent relations at younger ages were much weaker. As shown in Table 4, online accuracy was significantly correlated with vocabulary production at both 21 and 25 months. Relations between online accuracy and grammar were more modest at 21 months, with only the correlation with M3L reaching significance at Bonferroni-corrected alpha levels. At 25 months, however, both grammar measures were significantly correlated with online accuracy. Thus, greater accuracy in spoken word recognition at the older ages was related to higher production vocabulary and grammar scores. As with the RT measure, correlations were weak at the younger ages, providing further evidence that links between online measures of language processing and offline measures of language abilities were most evident in children likely to be experiencing rapid growth in linguistic abilities.[8]

Given that both speed and accuracy in word recognition at 25 months were concurrently correlated with CDI measures and the PPVT–R, we next examined relations between speech-processing skills at 25 months and measures of vocabulary and grammar at the four previous time points.[9] As shown in Table 5, faster mean RT

at 25 months was significantly related to larger comprehension vocabularies at 12 and 15 months, as well as to production vocabularies at 12, 15, 18, and 21 months. Online RT at 25 months was also significantly related to grammar measures at 21 months. Relations between 25-month RT and grammar measures at 18 months were weaker, not surprising given that children at this age are just beginning to produce complex grammatical forms. Table 6 shows that accuracy at 25 months was also related to earlier lexical and grammatical measures. Higher accuracy at 25 months was significantly related to larger comprehension vocabularies at 12 and 15 months as well as larger production vocabularies at 12, 18, and 21 months. At 15 months, the correlation between 25-month accuracy and production vocabulary just missed significance ($p = .06$) with the conservative Bonferroni-corrected value. As with RT, associations between 25-month accuracy and earlier grammar were most consistent at the 21-month time point; at 18 months, the correlation was significant for only the M3L grammar measure.[10] Taken together, these findings show that efficiency in spoken language understanding at 25 months was consistently related to prior as well as concurrent linguistic accomplishments.

Relations Between Efficiency in Online Spoken Language Comprehension and Growth in Productive Vocabulary

The analyses presented so far have examined online measures of speech processing in relation to age and to concurrent and prior language skills. However, these analyses provided no information about how individual children changed over time. Because there was no concept of trend in these analyses, it was not possible to tell whether individual differences in the same children accounted for the effects at each time point. An important advantage of a longitudinal repeated measures design is that it enables one to assess not only group-level effects but also trajectories of growth in individual children from 12 to 25 months of age.

To examine these patterns of change, we used statistical techniques that evaluate the data in terms of a hierarchical or multilevel structure (Raudenbush & Bryk, 2002). At Level 1, repeated observations of individuals can be assessed with respect to individualized growth functions described by a unique set of parameters (e.g., starting point, or *intercept,* and linear rate of change, or *slope*). Because vocabulary was assessed at five time points, it was also possible to explore the degree to which individual trajectories of vocabulary growth were characterized by gradual increases in rate of change across the period (i.e., *acceleration*). Using these three parameters, we obtained both an average starting point in vocabulary (i.e., at 12 months) and the average linear and quadratic

---

[8] See footnote 6.

[9] Here and in the growth curve analyses, we focus on processing measures at 25 months because correlations with concurrent and prior language measures were strongest and most consistent at this age. Correlations between RT and accuracy at 15 and 18 months and expressive vocabulary size at later ages were mostly low, with one notable exception: Mean RT at 15 months was significantly correlated with vocabulary size at 25 months ($r = -.42$, $p < .01$). At 21 months, accuracy was significantly correlated with expressive vocabulary size at each age from 15 to 25 months (rs $= .38$ to $.50$, $p < .05$), although correlations between accuracy at 15 months and later vocabulary size were not significant.

[10] See footnote 6.

Table 2
Descriptive Statistics for the Offline Vocabulary and Grammar Measures From the CDI and PPVT–R

Age	Words understood[a]		Words produced[b]		Grammatical complexity[c]		M3L[d]		PPVT–R[e]	
	M	SD	M	SD	M	SD	M	SD	M	SD
12 months	82.5	62.3	9.6	10.8	—		—		—	
15 months	166.2	88.7	28.9	30.7	—		—		—	
18 months	—		95.6	101.4	1.1	4.4	1.8	1.0	—	
21 months	—		208.5	158.9	5.2	7.4	3.1	1.7	—	
25 months	—		391.7	176.8	13.5	11.5	5.5	3.1	13.7	5.4

Note. CDI = MacArthur–Bates Communicative Development Inventory; PPVT–R = Peabody Picture Vocabulary Test—Revised; M3L = mean of the three longest utterances; dash = data are not available.
[a] Number of words reported as "understands" on the CDI: Words and Gestures (reported at 12 and 15 months). [b] Number of words reported as "understands and says" on the CDI: Words and Gestures (reported at 12 and 15 months) or CDI: Words and Sentences (reported at 18, 21, and 25 months). [c] Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words and Sentences (reported at 18, 21, and 25 months). [d] M3L reported on the CDI: Words and Sentences (reported at 18, 21, and 25 months). [e] Ceiling score minus errors from the PPVT–R, administered at 25 months.

terms that characterized individual vocabulary growth from 12 to 25 months. Preliminary analyses of individual trajectories that were run using regression-based curve estimation techniques reflected both linear and quadratic growth. Statistical estimates of variance accounted for (or goodness of fit as indicated by r^2) indicated that linear functions (in the general form, $y = ax + b$) provided a strong fit ($M = .84$), averaging across individuals. However, linear functions accounted for significantly less variance than quadratic functions, $M = .97$, $t(49) = 8.5$, $p < .001$, expressed in the general form, $y = ax^2 + bx + c$, where a represents acceleration in learning rate.[11]

Once these three Level 1 parameters were defined for each individual, Level 2 analyses determined whether variation in the linear and quadratic parameters could be accounted for by other person-level (between-subjects) characteristics. RT and accuracy in online processing at 25 months were the two Level 2 factors

Table 3
Concurrent Correlations Between Reaction Time and Offline Vocabulary and Grammar Measures

Age	Words understood[a]	Words produced[b]	Grammatical complexity[c]	M3L[d]	PPVT–R[e]
15 months	−.30*	−.25	—	—	—
18 months	—	−.21	−.29	−.22	—
21 months	—	−.20	−.19	−.12	—
25 months	—	−.38*	−.40**	−.41**	−.60**

Note. Mean response latency (ms) to shift to the target picture. M3L = mean of the three longest utterances; PPVT–R = Peabody Picture Vocabulary Test—Revised. Dash indicates data are not available.
[a] Number of words reported as "understands" on the MacArthur–Bates Communicative Development Inventory (CDI): Words and Gestures (reported at 15 months). [b] Number of words reported as "understands and says" on the CDI: Words and Gestures (reported at 15 months) or CDI: Words and Sentences (reported at 18, 21, and 25 months). [c] Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words and Sentences (reported at 18, 21, and 25 months). [d] M3L reported on the CDI: Words and Sentences (reported at 15, 18, and 25 months). [e] Ceiling score minus errors from the PPVT–R, administered at 25 months.
* $p < .05$. ** $p < .01$.

evaluated, for two reasons. First, our previous analyses indicated that online measures at 25 months were correlated with both concurrent and prior linguistic accomplishments. Second, online performance at 25 months provided more stable estimates of spoken word recognition than earlier assessments, as indicated by reduced variability in performance across the sample.

A series of Level 1 and Level 2 hierarchical growth curve models was constructed using the linear mixed models procedure in SPSS (Version 12.0), with maximum likelihood estimation used for all models. Table 7 presents a summary of the results, including overall model fits (goodness of fit expressed as −2 restricted log likelihood [−2RLL] in "smaller-is-better" form), number of parameters in each model, evaluation of estimates of coefficients for fixed effects, and evaluation of covariance parameter estimates. The first stage of the analysis examined Level 1 models with no Level 2 variables (i.e., "unconditional" models). A comparison of linear (Model 1) versus quadratic (Model 2) unconditional models indicated a significantly better fit for the quadratic model (−2RLL = 2,728.3) than for a model including only a linear term (−2RLL = 2,859.5), $p < .01$. Examination of estimates of the covariance parameters (random effects) indicated significant individual variation in both linear and quadratic parameters (all $ps < .001$). Model 2 further reflected significant covariance (i.e., col-

---

[11] It could be argued that other nonlinear growth functions (e.g., logistic) are better candidates for modeling growth in vocabulary. For example, in contrast to the quadratic, the logistic function constrains the starting point to be zero and limits the upper bound, forcing growth to slow down and trajectories to asymptote. Further, the logistic function incorporates a parameter that reflects an "inflection point," that is, a sudden transition from slower to more rapid learning that could characterize a "spurt" or "burst" in learning rate. We chose to use a quadratic function here for several reasons. First, quadratic functions are mathematically simpler than logistic functions and can more easily be incorporated into statistical analyses. In addition, few ceiling effects were observed; that is, few children showed a leveling off or slowing down of growth. Finally, Ganger and Brent (2004) have recently shown that, for most children, increases in the rate of vocabulary growth are gradual across the period, best captured by a quadratic model, rather than characterized by an abrupt change, or spurt in growth, as modeled by a logistic function.

Table 4
Concurrent Correlations Between Accuracy and Offline Vocabulary and Grammar Measures

Age	Words understood[a]	Words produced[b]	Grammatical complexity[c]	M3L[d]	PPVT–R[e]
15 months	.20	.10	—	—	—
18 months	—	.05	−.16	−.04	—
21 months	—	.47**	.25	.34**	—
25 months	—	.49**	.44**	.50**	.62**

Note. Mean proportion of correct shifts to the target picture. M3L = mean of the three longest utterances; PPVT–R = Peabody Picture Vocabulary Test—Revised. Dash indicates data are not available.
[a] Number of words reported as "understands" on the MacArthur–Bates Communicative Development Inventory (CDI) Words & Gestures (reported at 15 months). [b] Number of words reported as "understands and says" on the CDI: Words and Gestures (reported at 15 months) or CDI: Words and Sentences (reported at 18, 21, and 25 months). [c] Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words & Sentences (reported at 18, 21, and 25 months). [d] M3L reported on the CDI: Words & Sentences (reported at 18, 21, and 25 months). [e] Ceiling score minus errors from the PPVT–R, administered at 25 months.
* $p < .05$. ** $p < .01$.

Table 6
Correlations Between Accuracy at 25 Months and Offline Vocabulary and Grammar Measures at 12, 15, 18, and 21 Months

Age	Words understood[a]	Words produced[b]	Grammatical complexity[c]	M3L[d]
12 months	.29*	.32*	—	—
15 months	.28*	.31	—	—
18 months	—	.39*	.18	.32*
21 months	—	.42*	.36*	.41*

Note. Mean proportion of correct shifts to the target picture. M3L = mean of the three longest utterances. Dashes indicate data are not available.
[a] Number of words reported as "understands" on the MacArthur–Bates Communicative Development Inventory (CDI): Words and Gestures (reported at 12 and 15 months). [b] Number of words reported as "understands and says" on the CDI: Words and Gestures (reported at 12 and 15 months) or CDI: Words and Sentences (reported at 18 and 21 months). [c] Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words and Sentences (reported at 18 and 21 months). [d] M3L reported on the CDI: Words and Sentences (reported at 18 and 21 months).
* $p < .05$.

linearity) between slope and acceleration (Linear × Quadratic), suggesting that trajectories of growth differed in acceleration depending on the average rate of change across the period. Examination of the fixed effects in Model 2 indicated that neither vocabulary size at 12 months (intercept) nor linear growth were significantly different from zero, with estimated coefficients for these parameters similar to or smaller than their respective standard errors. However, the parameter for acceleration (quadratic) differed significantly from zero, $t(49) = 9.2$, $p < .001$, indicating that children on average increased in their rate of vocabulary growth over time.

Table 5
Correlations Between Reaction Time at 25 Months and Offline Vocabulary and Grammar Measures at 12, 15, 18, and 21 Months

Age	Words understood[a]	Words produced[b]	Grammatical complexity[c]	M3L[d]
12 months	−.45**	−.39*	—	—
15 months	−.36*	−.35*	—	—
18 months	—	−.36*	−.13	−.28
21 months	—	−.45**	−.35*	−.48**

Note. Mean response latency (ms) to shift to the target picture. M3L = mean of the three longest utterances. Dash indicates data are not available.
[a] Number of words reported as "understands" on the MacArthur–Bates Communicative Development Inventory (CDI): Words and Gestures (reported at 12 and 15 months). [b] Number of words reported as "understands and says" on the CDI: Words and Gestures (reported at 12 and 15 months) or CDI: Words and Sentences (reported at 18 and 21 months). [c] Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words and Sentences (reported at 18 and 21 months). [d] M3L reported on the CDI: Words and Sentences (reported at 18 and 21 months).
* $p < .05$. ** $p < .01$.

In the second stage of this analysis, we examined whether individual differences in these parameters were related to individual differences in speed and accuracy of online language comprehension, covariates in the Level 2 models. Because these two Level 2 factors were significantly intercorrelated at 25 months ($r = -.53$), their effects were modeled individually. Further, because the linear and quadratic terms were strongly collinear (as determined by the covariance estimates of Model 2), individual models were constructed to evaluate the relationship of each Level 2 covariate to the linear (Models 3 and 5) and quadratic (Models 4 and 6) terms separately. As shown in Table 7, all of the resulting Level 2 models had similar random effects, with significant variance in both growth terms and significant Linear × Quadratic covariance.

An examination of the fixed effects for RT (Models 3 and 4) indicated significant relationships between RT and both the linear, Model 3 RT × Linear: $t(55) = -3.8$, $p < .0001$, and quadratic parameters, Model 4 RT × Quadratic, $t(49) = -3.2$, $p < .002$. That is, children with steeper rates of change and more accelerated growth tended to have faster RTs during the looking-while-listening task at 25 months. To illustrate this effect, we classified children as either fast (≤ 750 ms) or slow (> 750 ms) in response speed on the basis of a median split of the mean RT scores at 25 months. Figure 3 plots growth in vocabulary from 12 to 25 months in these two RT groups. Children in the faster RT group had steeper and more accelerated vocabulary growth across the 2nd year compared with children with slower RTs.

A similar pattern was observed for accuracy (Models 5 and 6). These Level 2 models indicated that online accuracy was significantly related to both linear rates of change, Model 5 Accuracy × Linear, $t(56) = 3.8$, $p < .0001$, and acceleration, Model 6 Accuracy × Quadratic, $t(50) = 49.7$, $p < .002$, in vocabulary growth. To illustrate this effect, we classified children as either high ($\geq 92\%$) or low ($< 92\%$) in accuracy based on a median split of the accuracy scores at 25 months. Figure 4 shows that those children who were more accurate in word recognition at 25 months had

Table 7
Summary of Level 1 (Unconditional) and Level 2 (RT and Accuracy) Growth Models

	Level 1: Unconditional		Level 2: RT		Level 2: Accuracy	
	Model 1: Linear	Model 2: Quad	Model 3: RT × Linear	Model 4: RT × Quad	Model 5: Acc × Linear	Model 6: Acc × Quad
Model fit (−2RLL)	2,859.5	2,728.3	2,715.1	2,719.2	2,715.4	2,719.4
No. of parameters	4	7	9	9	9	9
Fixed effects' coefficient (SE)						
Intercept	**−39.9 (7.0)**[a]	4.6 (5.3)	−0.6 (31.2)	25.9 (30.5)	13.3 (27.3)	−9.6 (26.7)
Linear	**30.6 (2.0)**[a]	3.9 (3.4)	**43.6 (11.0)**[a]	3.9 (3.4)	**−31.1 (9.8)**[a]	3.9 (3.4)
Quad		**2.1 (0.2)**[a]	**2.1 (0.2)**[a]	**4.3 (0.8)**[a]	**2.1 (0.2)**[a]	0.1 (0.7)
RT × Intercept			6.8 (40.0)	−27.6 (39.0)		
RT × Linear			**−51.9 (13.7)**[a]			
RT × Quad				**−3.0 (0.9)**[a]		
Acc × Intercept					−10.4 (31.7)	16.7 (30.9)
Acc × Linear					**41.0 (10.8)**[a]	
Acc × Quad						**2.4 (0.7)**[a]
Random effects, covariance estimate (SE)						
Linear	**159.8 (34.8)**[b]	**397.8 (96.6)**[b]	**−356.0 (88.5)**[b]	**387.9 (96.0)**[b]	**360.2 (89.5)**[b]	**391.5 (96.8)**[b]
Linear × Quad		**−18.7 (5.9)**[b]	**−18.6 (5.8)**[b]	**−20.5 (6.3)**[b]	**−18.8 (5.8)**[b]	**−20.6 (6.3)**[b]
Quad		**1.5 (0.4)**[b]	**1.5 (0.4)**[b]	**1.6 (0.5)**[b]	**1.5 (0.4)**[b]	**1.6 (0.5)**[b]

Note. Significant effects are indicated in bold. Acc = accuracy; quad = quadratic.
[a] t value, $p < .001$. [b] Wald z, $p < .001$.

steeper and more accelerated trajectories of lexical growth across the 2nd year compared with children with lower accuracy scores at 25 months. Although visual inspection of Figures 3 and 4 might suggest that rate and acceleration of vocabulary growth was related less strongly to accuracy than to response speed, fits for the two sets of Level 2 models were virtually identical (see Table 7).

These hierarchical growth models showed that the course of vocabulary growth from 12 to 25 months was related to speech-processing efficiency at the end of the 2nd year. The Level 1 models indicated that individual growth trajectories were best characterized by functions that included a nonlinear (quadratic) component, capturing individual differences in the degree to which vocabulary growth accelerated over this period. However, linear slope and acceleration were highly intercorrelated, suggesting that these two factors go hand-in-hand in describing trajectories of growth in this period; that is, children with steeper slopes of change also accelerated more rapidly. Level 2 models indicated that individual differences in trajectories of growth were significantly related to individual differences in both the speed and accuracy of spoken language comprehension. A comparison of the model fits indicated significantly better fits for those models that incorporated the Level 2 factors of RT (−2RLL = 2,719.2) and accuracy (−2RLL = 2,719.4) compared with the best-fitting unconditional Level 1 model (−2RLL = 2,728.3) that did not include these additional factors ($p < .05$). Because speed and accuracy of spoken language comprehension were intercorrelated
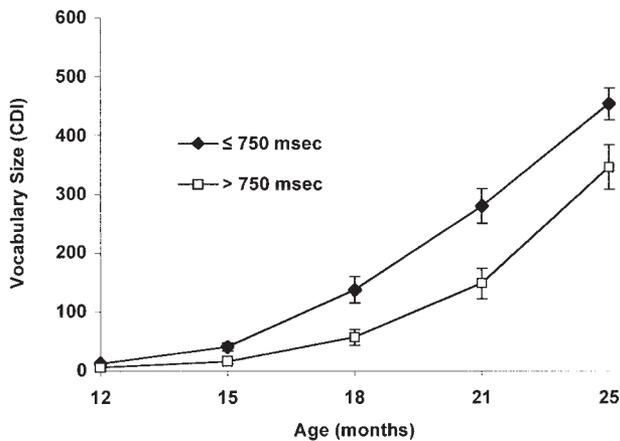

Figure 3. Mean trajectories of growth in vocabulary production as a function of reaction time in spoken word recognition shown at 25 months. Groupings are based on a median split of mean reaction time to shift to target picture at 25 months. Error bars represent standard errors.
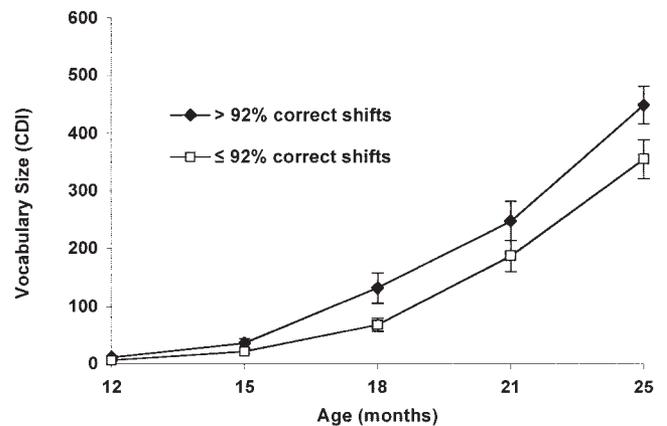

Figure 4. Mean trajectories of growth in vocabulary production as a function of accuracy in spoken word recognition shown at 25 months. Groupings are based on a median split of mean proportion of correct shifts to target picture at 25 months. Error bars represent standard errors.

in this sample, our analyses could not assess their unique contribution in accounting for trajectories of vocabulary growth. Nevertheless, the results showed that descriptions of trajectories of vocabulary growth were improved when these indices of online performance were incorporated into the models.

Discussion

This study provides the first longitudinal data on the emergence of efficiency in spoken language understanding across the 2nd year, relating developmental changes in speech-processing abilities to growth in vocabulary and grammatical competence. Children's speed and accuracy in spoken word recognition increased significantly over this period, consistent with earlier cross-sectional research. To explore the relation of online measures of speech-processing skill to more traditional measures of linguistic development, we gathered parental reports of vocabulary and grammatical usage at 12, 15, 18, 21, and 25 months, and a standardized test of lexical knowledge was administered at 25 months. Although efficiency of word recognition at 15 and 18 months was not strongly correlated with concurrent and later measures of language development, speed and accuracy in speech processing at 25 months were found to be robustly related to lexical and grammatical development across a range of measures from 12 to 25 months. Analyses of growth curves revealed that children who were relatively faster and more accurate in spoken word recognition at 25 months were those who had experienced faster and more accelerated vocabulary growth across the 2nd year.

Developmental Changes in the Efficiency of Spoken Language Processing

One goal of this research was to replicate earlier cross-sectional results showing that efficiency in spoken language processing increases substantially over the 2nd year. Observing separate groups of 15-, 18-, and 24-month-olds, Fernald et al. (1998) found age-related changes in speed and accuracy of word recognition over the 9-month period. The present longitudinal findings confirmed this overall pattern across the four time points sampled. The mean RT for 15-month-olds in the present study ($M = 981$ ms) was comparable to the mean RT for 15-month-olds in the Fernald et al. study ($M = 995$ ms), a replication that was not surprising given that the same speech stimuli were used. For 18-month-olds, however, the results were somewhat different. In the earlier cross-sectional study, the mean RT at 18 months ($M = 827$ ms) was significantly lower than it was at 15 months, whereas in the present longitudinal sample, the mean RT at 18 months ($M = 962$ ms) did not reflect a substantial decrease over this 3-month period. Although in both studies the change in response speed across the 2nd year was highly significant, RT decreased somewhat less overall in this longitudinal sample ($M = 210$ ms) than in the previous cross-sectional study ($M = 314$ ms).

One likely explanation for this discrepancy is that the stimuli used at the 18-month time point in the present study differed in important ways from those used by Fernald et al. (1998). To make the task more challenging at 18 months than at 15 months, we included a number of partial word trials consisting of truncated versions of familiar object names in which only the first 300 ms of the initial consonant and vowel was presented. Fernald et al.

(2001) had shown that 18-month-olds are able to recognize such partial words; however, this stimulus set may have been especially difficult for many of the infants at this age.[12] Another challenging feature of this stimulus set is that more filler trials were included than in the earlier study of partial-word recognition in an effort to reduce the redundancy of the frequently repeated stimuli. This attempt to maintain infants' interest by increasing stimulus variability unfortunately had the opposite effect, with the result that infants completed on average only 55% of the critical trials at the 18-month test session, far fewer target word trials than they completed on average at 15 months (80%), 21 months (87%), or 25 months (94%).[13] At the 21-month time point, partial word stimuli were also included, but by this age the children were unperturbed by the variability in the stimulus set, perhaps because they had heard partial words before. Despite the somewhat anomalous results at the 18-month time point, the finding that speed and accuracy in spoken word recognition increased significantly between the ages of 15 and 25 months clearly replicated the overall results of Fernald et al.'s (1998) study.

Our analyses also enabled us to explore two possible alternative explanations for the developmental changes in speech-processing efficiency observed in this and other studies using online measures. One concern is that accuracy in the looking-while-listening procedure is based on shifts in fixation from the distracter to the target picture. If 2-year-olds tend to shift back and forth between pictures more frequently than do infants at younger ages, could this confound between age and spontaneous shifting account for the increases in accuracy scores observed in the older children? As shown in Figure 2, children did indeed increase their tendency to shift from one picture to the other with increasing age; however, this increase was restricted to distracter-initial trials when the child shifted to the correct picture in response to the target word. When the child was already looking at the target picture by chance, there was no significant increase with age in the tendency to shift away. This asymmetry reveals that the increase of more than 30% in mean accuracy between the ages of 15 and 25 months cannot be accounted for by a general tendency of older infants to shift more frequently between pictures.

Another possible alternative explanation is that older children may have performed better overall than younger children not

---

[12] Although a variety of different trial types were included in the stimulus sets at different time points, the numbers of each trial type were small, and their main purpose was to make the task challenging to children at each age. Thus, we did not conduct separate analyses of each subtype for purposes of statistical comparison. Descriptive statistics showed that children's performance on the partial-word trials at 18 and 21 months was very similar to that found by Fernald et al. (2001).

[13] Children were excluded from the data set at any given age if they were inattentive on > 70% of trials overall. However, the measure of inattentiveness reported here refers to the percentage of target word trials (not including fillers) completed by those children included in the sample. A methodological insight to be gained from the pattern of findings at 18 months is that the overall complexity of the stimulus set can have a strong influence on children's performance. Thus, infants may be fast and accurate in recognizing a familiar target word presented in a predictable sequence of trial types, yet slower and less accurate in recognizing the same target word in the same frame when the other stimuli in the sequence are highly variable, and susceptibility to such interference varies with age.

because of greater efficiency in identifying familiar words, but simply because they knew more of the words they were being tested on. To address this concern, we conducted two series of analyses on every measure, one based on all the trials available for each child and the other based only on those trials with target words that were reportedly understood and produced by the child. The fact that these two different approaches yielded such similar patterns of results indicates that increases in speed and accuracy observed in the older and more advanced children were not an artifact of the fact that they "knew" more of the words included in the stimulus sets. Instead, we can argue that the observed developmental changes reflected children's growing efficiency in identifying highly familiar as well as less familiar words and not just their success in recognizing words well known to them.

We also examined this question more directly by comparing RT and accuracy to "known" versus "unknown" target words within the same child, focusing on just the small subset of children at each age for whom at least one of the target words was reported on the CDI as either "not understood" (at 15 months) or as "not understood and produced" (at 18, 21, and 25 months). Mean values indicated comparable RTs for "known" versus "unknown" words at all ages (15 months: 975 ms vs. 995 ms; 18 months: 979 ms vs. 963 ms; 21 months: 835 ms vs. 819 ms; and 25 months: 835 ms vs. 858 ms). Accuracy scores were also comparable in children's responses to "known" versus "unknown" words (15 months: 58% vs. 46%; 18 months: 51% vs. 54%; 21 months: 64% vs. 67%; 25 months: 78% vs. 80%).[14] Thus, for children for whom "known" versus "unknown" words could be directly compared, we found a pattern of quick and accurate responses even to those target words they reportedly did not yet understand and produce. These analyses make it clear that assumptions about what words are "unknown" by a child based on parental report should be considered with caution. By using online measurement techniques, we were able to tap into children's emerging understanding of words presumed to be unfamiliar, revealing receptive knowledge that builds gradually and may not yet be evident in spontaneous behavior.

Stability of Online Measures of Speech Processing From 15 to 25 Months

Although the first analyses focused on infants grouped by age, a second goal of this research was to investigate individual differences in the speed and accuracy of online speech processing. To examine the stability of these measures over time, we calculated correlations within each measure at adjacent sampling points. Children's mean accuracy scores were significantly correlated from the 18- to 21-month and the 21- to 25-month time points, whereas mean RT scores were stable between the 18- and 21-month time points, but only marginally so between 21 and 25 months. One explanation for the low correlations between 15 and 18 months is that measurement error may have obscured any underlying stability in these measures. Because unanticipated problems with the stimulus set resulted in lack of engagement during the 18-month test session, the data from that session did not represent children's best performance, as reflected in the finding that both speed and accuracy were lower than expected on the basis of previous research. However, it is also very likely that the lack of stability in skills related to spoken word recognition observed between 15 and 18 months is not simply an artifact of task

demands and that components of speech-processing skill are *not* particularly stable when children are first learning to speak. In the real world as well as in the looking-while-listening task used here, success in word recognition depends not only on lexical knowledge but also on numerous factors not specific to language such as memory, attention, and skill in categorizing objects. If variability in such nonlinguistic abilities is greater in younger infants than in 2-year-olds, this could contribute to lack of stability in speech-processing efficiency in the early stages of language learning. Further longitudinal research with carefully designed stimuli and multiple test sessions at each age is necessary to resolve these measurement issues.

Relations Between Efficiency in Spoken Language Comprehension and Growth in Vocabulary Across the 2nd Year

The third and most important goal of this research was to investigate relations between online measures of speech-processing efficiency and traditional offline measures of language development. Previous studies using online measures with infants older than 18 months have found such relations between processing variables and vocabulary size (Fernald et al., 2001; Zangl et al., 2005), although studies with younger infants have not (Swingley & Aslin, 2000, 2002). Results in the present longitudinal study are generally consistent with these previous findings, indicating that significant relations between speech-processing efficiency and vocabulary production are evident by 21 months of age. However, the longitudinal design of the current study allowed a broader examination of these links as children progressed from first words to the period of rapid growth in expressive vocabulary and early use of grammar. We found that children's response speed and accuracy at 25 months were related not only to concurrent vocabulary size but also to almost all prior measures of language from the age of 12 months. Results from growth curve analyses underscored the continuity in language abilities apparent across the 2nd year (see Bates & Goodman, 1997). Consistent with previous research, the children in this study varied significantly in the degree to which vocabulary growth could be characterized by steeper and more accelerated rates of change. Both of these individual difference variables (i.e., slope and acceleration) were associated with RT and accuracy in word recognition by the end of the 2nd year. That is, those children who demonstrated faster and more accelerated rates of productive language growth across the 2nd year were those who responded more quickly and reliably in the looking-while-listening task at 25 months. Thus, it was not just that those children with more developed speech-processing abilities at age 2 had also had more fully developed productive language skills at previous time points. Rather, greater efficiency in language comprehension by 25 months was related more broadly to children's overall *trajectories* of vocabulary learning.

---

[14] These analyses were limited to children for whom at least one of the target words at a given age was reportedly not understood and produced. As the analyses required a sufficient number of codeable responses in both of the two subsets of trials, they were based on variable numbers of trials in a small subset of the children at each age.

How Is Lexical Learning Related to Speech-Processing Efficiency?

The most important finding in this research is that individual differences in children's efficiency in interpreting spoken language were related to individual differences in their lexical and grammatical development. However, this association did not emerge strongly until the end of the 2nd year, and the nature and direction of this relation are far from clear. One possibility is that we underestimated the proficiency with which children could identify familiar words at the younger ages. If so, it could still be argued that greater processing efficiency may facilitate word learning at the earliest stages of language development, although our findings failed to capture this relation. An alternative possibility is that infants differ in their rate of early vocabulary learning because of factors unrelated to preexisting differences in information processing abilities. According to this explanation, it is the quality of early language experience that accounts primarily for initial differences in vocabulary size. If children with richer language input begin to talk sooner, these lexically more advanced children could develop faster processing speed through increased experience both in hearing and using speech. This could then give them an advantage in recognizing familiar words and in learning new ones, so that by the end of the 2nd year greater speech-processing efficiency is strongly associated with more rapid vocabulary growth. Although this chicken-or-egg version of the alternatives is oversimplified, we consider each in turn and then describe how processing efficiency could interact with vocabulary size in ways that influence the rate of lexical and grammatical learning.

The first of these two explanations draws on research showing that processing speed increases with age and correlates with competence across a range of cognitive tasks. In a meta-analysis, Kail (1991) examined mean RT on both nonlinguistic and linguistic tasks including mental rotation, Stroop, reading, and visual search, concluding that age differences in processing speed reflect a general, non-task-specific component that matures rapidly during childhood. At every age, there is also substantial variability in mean RT for children as well as adults, individual differences that predict performance on numerous cognitive tasks (Kail & Salthouse, 1994). RT in adults correlates so robustly with measures of fluid intelligence, memory, analogical reasoning, and language performance (Kail, 1992; Kwong See & Ryan, 1995; Salthouse, 1991) that many researchers believe developmental increases in processing speed can account fundamentally for age-related growth in cognitive functioning (e.g., Salthouse, 1996).

This literature suggests that individual differences in RT in the age range we studied could be associated with cognitive skills that are not necessarily specific to language learning. In the looking-while-listening procedure, a pattern of correct responses presupposes that the child is also proficient in other cognitive processes not exclusively linked to language. The child must first encode the visual image, parse the sentence, and determine whether the target word matches the fixated picture. If the spoken word matches what the child is looking at, the correct response is to continue fixating that picture; if there is a mismatch, the child must reject the picture and mobilize an eye movement to search for a more appropriate referent. Accuracy in either case presupposes a range of capabilities, including focused attention, rapid encoding of visual images, integration of visual and auditory input, association of the target word with an appropriate picture, and the ability to disengage quickly from one picture to attend to another. These are only some of the perceptual, motor, and cognitive processes that could influence children's performance in this word-recognition task, and all entail abilities not specific to linguistic processing. Thus, one explanation for the relation we found between efficiency in online word recognition and vocabulary growth is that faster processing speed at 25 months reflects greater competence in a range of abilities that facilitate but are not limited to language learning. Although we found no relation between speed of word recognition and speed of orienting in a nonlinguistic task, individual differences in cognitive processing skills at higher levels are undoubtedly relevant to spoken language processing.

As such component skills may develop at different rates in different children in the very early stages of word learning, this argument could in part explain why vocabulary growth was more robustly related to speech-processing efficiency at 25 months than at earlier time points. Another plausible possibility is that measurement error due to fluctuations in infant attention and the use of some stimuli that were too challenging added noise to our measurements at the younger ages. Taking both factors into consideration, it is possible that relations between processing efficiency at 15 and 18 months and later vocabulary growth would emerge more strongly if children were tested on more homogeneous verbal stimuli in multiple sessions at each age, to offset the effects of transient inattentiveness in any particular session.

An alternative explanation for our main findings is that early differences in the size of children's productive vocabularies at first have nothing to do with processing efficiency but depend on environmental factors such as the amount and quality of the language experienced by the child. If children with richer language input advance more rapidly in word learning from the beginning, they would gain more experience in interpreting spoken language over the 2nd year compared with children exposed to less rich input who understand fewer words. Thus, greater facility in speech processing could emerge as a consequence of these early experiential differences, accounting for the relation we found between lexical development across the 2nd year and speech-processing measures at 25 months. Consistent with this argument, numerous studies have shown that the amount and complexity of talk directed to young children vary substantially across families, and these differences in early language input are associated with differences in various measures of language development (Hart & Risley, 1995; Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). Children with richer input may develop larger vocabularies because they hear familiar words more frequently, have more experience in making sense of speech, and have more opportunities to learn and use new words. In this case greater efficiency in spoken word recognition could be seen as a kind of practice effect. Another argument suggesting that vocabulary size may drive the development of speech-processing skills is that young children who have larger vocabularies require more refined and efficient word-recognition skills in order to distinguish among greater numbers of potentially confusable representations in the mental lexicon. Thus, Charles-Luce and Luce (1990) and Walley (1993) have suggested that vocabulary growth itself leads to the development of more efficient processing strategies and new forms of phonological organization.

Whether individual differences in infants' processing abilities influenced lexical learning from the outset, or whether speech-processing skills emerged more gradually over the 2nd year as a consequence of early learning, the links we found between acceleration in vocabulary growth and efficiency of spoken language interpretation were robust by 2 years of age. These factors must also interact synergistically in ways that continue to affect the rate of learning. A potentially important factor in this synergy is that faster processing speed can free additional cognitive resources (Salthouse, 1996), a benefit of particular value in the early stages of language learning. The child who can identify familiar words more rapidly and reliably will have more resources available for attending to subsequent words in the sequence. Although the richness of the linguistic environment continues to play an important role throughout childhood (e.g., Weizman & Snow, 2001), individual differences in processing efficiency would interact with experiential factors in later lexical and grammatical development. For example, child-directed speech functions as "language input" only to the extent that the child can actually process what is heard. The word *dog* may be spoken equally often in two environments, but the child who is a little more efficient in accessing the meaning of *dog* from the acoustic signal can identify that word more reliably whenever it is spoken compared with a child who fails to identify the word on some occasions. That is, faster recognition of familiar words may enable a child to build up lexical representations in less time because these words are processed as meaningful lexical items more reliably and thus, in effect, more frequently. Regardless of whether particular words are actually spoken more frequently to the child, the ability to recognize them more reliably might convey some of the processing advantages of the frequency effects observed so widely in research on adult comprehension (Monsell, 1991). This could lead to greater success in learning new words encountered later in the sentence as well as in tracking distributional information about relations among the words that is essential for learning grammar.

As a result of such positive-feedback processes, there could be cascading advantages for the child who has a larger lexicon and more efficient word-recognition skills by the age of 2. Rapid access to the meanings of familiar words through incremental processing of continuous speech would likely continue to enhance both lexical and grammatical learning. Even for fluent adults, efficient comprehension requires processing speech continuously rather than waiting to the end of a sentence to interpret what has been said (Marslen-Wilson, 1987; Tanenhaus et al., 1996). For young language learners, the challenge of interpreting continuous speech is immeasurably greater. They not only need to attend to each word as it comes along but also to remember and relate nonadjacent words in the sequence in order to appreciate long-range dependencies crucial for mastering syntax. Moreover, by the end of the 2nd year children increasingly rely on known words to infer the meanings of unknown words (Fernald, 2002; Goodman, McDonough, & Brown, 1998). Thus, a slight initial advantage in the efficiency of spoken word recognition could be strengthened through positive-feedback processes, leading to faster growth in vocabulary and grammar that in turn lead to further increases in receptive language competence. This accords well with the suggestion by Elman et al. (1996) that the nonlinear patterns of early vocabulary growth show "the more words you know, the easier it is to accumulate more" (p. 185). A positive-feedback loop between efficiency in processing familiar words and success in learning new words is one mechanism that can explain this insight.

Conclusions

Two major findings emerged from this research. First, it is clear that children become more competent in interpreting spoken language between 15 and 25 months because they can not only respond reliably to an increasing number of words but also because they can respond more quickly and accurately to the same words they learned months earlier. This kind of gradual increase in the ability to identify a word in continuous speech is inconsistent with the all-or-none view of what it means to "know" a word that is implicit in much of the developmental literature and thus has implications for understanding the nature of early word learning. For example, the phenomenon of "fast-mapping" demonstrated in so many word-learning experiments is often described as if it were an endpoint, rather than just the beginning of the process of learning to comprehend a new word. Our findings are more consistent with the view that lexical representations are built up gradually and through experience become steadily more robust. It is also revealing that responses to those "known" target words parents reported as understood and produced by the child did not differ appreciably from responses to those target words reportedly not known. Children at all ages were quite successful in identifying words that in their everyday behavior they did not yet show evidence of understanding and using productively, consistent with the view that learning starts with partial knowledge involving graded representations (Munakata, 2001). Because eye-tracking methods yield more sensitive measures of listeners' responses to spoken language, the looking-while-listening procedure offers a valuable tool for exploring how graded lexical representations grow stronger over time in the early stages of vocabulary learning.

The second major finding of this research is that those children who were faster and more accurate in online comprehension at 25 months were those who had shown greater acceleration in vocabulary growth across the 2nd year. This provides the first experimental evidence for a link between efficiency in spoken language understanding and rate of language learning in the same group of infants followed longitudinally. Strong relations between processing efficiency and vocabulary size were not apparent at the outset of language learning but began to emerge around 21 months when children could typically produce 200 words or more. Thus, this pattern of findings still leaves open the question of what role individual differences in receptive language abilities play in very early word learning. One explanation considered is that those children who developed language at a faster rate were slightly more proficient all along in at least some of the skills involved in interpreting speech in real time, although our measures were not sensitive enough to capture these differences at the younger ages. Perhaps when first tested at 15 and 18 months, children were inattentive and immature in various cognitive abilities involved in identifying spoken words in this procedure, resulting in greater response variability than at later ages. If such component skills mature and become better integrated with age, this could account for the more stable assessment of individual differences in children observed at 21 and 25 months. An alternative explanation is that stable differences among children in the efficiency in spoken language understanding develop only gradually, and that language

experience and learning shape this development. Some children may have learned to speak more words at younger ages for reasons unrelated to their processing abilities at the time, becoming faster and more accurate as a result of their more extensive experience in hearing, producing, and interpreting speech. Both explanations are consistent with our finding of strong relations between speed and accuracy in language understanding and lexical development by the end of the 2nd year, although the causal role of processing skills in early vocabulary learning is unclear. Whether infants' initial processing abilities influence vocabulary growth from the very beginning of building a lexicon or individual differences in speech-processing skills emerge more gradually as a result of early learning are questions for future research. However, these data are consistent with a dynamic view of acquisition in which processing efficiency and linguistic knowledge operate in a synergistic fashion over the course of language development.

References

Aslin, R., Jusczyk, P., & Pisoni, D. (1998). Speech and auditory processing during infancy: Constraints on and precursors to language. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology* (5th ed., Vol. 2, pp. 147–198). New York: Wiley.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge, England: Cambridge University Press.

Bates, E., & Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes, 12,* 507–584.

Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression.* Cambridge, England: Cambridge University Press.

Canfield, R., Smith, E., Brezsnyak, M., & Snow, K. (1997). Infant 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, 62*(2, Serial No. 250).

Charles-Luce, J., & Luce, P. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language, 17,* 205–215.

Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.

Colombo, J., & Fagan, J. (1990). *Individual differences in infancy: Reliability, stability, prediction.* Hillsdale, NJ: Erlbaum.

Darwin, C. (1877). A biographical sketch of an infant. *Mind, 7,* 285–294.

Dougherty, T., & Haith, M. (1997). Infant expectations and reaction time as predictors of childhood speed of processing and IQ. *Developmental Psychology, 33,* 146–155.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test—Revised.* Circle Pines, MN: American Guidance Service.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge MA: MIT Press.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., et al. (1993). *User's guide and technical manual for the MacArthur Communicative Development Inventories.* San Diego, CA: Singular Press.

Fernald, A. (2002, March). *How infants develop expectations about what's coming next in speech.* Paper presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York.

Fernald, A. & Hurtado, N. (in press). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science.*

Fernald, A., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in understanding and recognizing words in fluent speech. In J. Weissenborn & B. Hoehle (Eds.), *Approaches to boot-strapping in early language acquisition* (pp. 97–123). Amsterdam: Benjamins.

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9,* 72–75.

Fernald, A., Swingley, D., & Pinto, J. (2001). When half a word is enough: Infants can recognize spoken words using partial acoustic-phonetic information. *Child Development, 72,* 1003–1015.

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology, 40,* 621–632.

Goldfield, B., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language, 17,* 171–183.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14,* 23–45.

Goodman, J., McDonough, L., & Brown, N. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development, 69,* 1330–1344.

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

Haith, M., Hazan, C., & Goodman, G. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development, 59,* 467–479.

Haith, M., Wentworth, N., & Canfield, R. (1993). The formation of expectations in early infancy. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 8, pp. 251–297). Westport, CT: Ablex.

Harris, M., Yeeles, C., Chasin, J., & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language, 22,* 1–17.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Brookes Publishing.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74,* 1368–1378.

Hurtado, N., Marchman, V. A., & Fernald, A. (2005). *Spoken word recognition by Latino children learning Spanish as their first language.* Manuscript submitted for publication.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27,* 236–248.

Jusczyk, P. (1997). Finding and remembering words: Some beginnings by English-learning infants. *Current Directions in Psychological Science, 6,* 170–174.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109,* 490–501.

Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology, 28,* 899–904.

Kail, R., & Salthouse, T. (1994). Processing speed as a mental capacity. *Acta Pscyhologica, 86,* 199–225.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992, January 31). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.

Kwong See, S., & Ryan, E. (1995). Cognitive mediation of adult age differences in language performance. *Psychology and Aging, 10,* 458–468.

Marslen-Wilson, W. (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.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in*

reading: Visual word recognition (pp. 148–197). Hillsdale, NJ: Erlbaum.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences, 5,* 309–315.

Polka, L., & Werker, J. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 421–435.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rose, S., Feldman, J., Jankowski, J., & Caro, D. (2002). A longitudinal study of visual expectation and reaction time in the first year of life. *Child Development, 73,* 47–71.

Saffran, J., Newport, E., & Aslin, R. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35,* 606–621.

Salthouse, T. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2,* 179–183.

Salthouse, T. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development, 69,* 309–320.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49,* 238–299.

Snyder, L., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8,* 565–582.

Song, H., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language, 52,* 29–57.

Stager, C., & Werker, J. (1997, July 24). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388,* 381–382.

Swingley, D., & Aslin, R. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76,* 147–166.

Swingley, D., & Aslin, R. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science, 13,* 480–484.

Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71,* 73–108.

Tanenhaus, M., Magnusen, J., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29,* 557–580.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1996). Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In T. Inui & J. McClelland (Eds.), *Perception and communication* (pp. 457–478). Cambridge, MA: MIT Press.

Thorpe, K., & Fernald, A. (in press). Knowing what a novel word is not: Two-year-olds "listen through" ambiguous adjectives in fluent speech. *Cognition.*

Tiedemann, D. (1927). Observations on the development of the mental faculties of children (S. Langer & C. Murchison, Trans.). *Pedagogical Seminary and Journal of Genetic Psychology, 34,* 205–230. (Original work published in 1787)

Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73,* 89–134.

Tsao, F., Liu, H., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75,* 1067–1084.

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

Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37,* 265–279.

Werker, J., Fennell, C., Corcoran, K., & Stager, C. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy, 3,* 1–30.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development, 6,* 179–208.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32,* 25–64.