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

Amy Perfors
Anne Fernald
Kalee Magnani
Virginia A. Marchman
Stanford University

Running Head: EFFICIENCY IN UNDERSTANDING

This research was supported by a grant from the National Institutes of Health to Anne Fernald (MH41511). Our thanks to Tomoko Wakabayashi, Alycia Cummings, and the research assistants at the Center for Infant Studies at Stanford for their many contributions, and to the parents and infants who participated in this study. Correspondence concerning this article should be addressed to: Amy Perfors, Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Rm NE20-388, 77 Massachusetts Ave, Cambridge, MA 02143; email: perfors@mit.edu.

Under review: Please do not cite without permission
Abstract

English-learning children (n=59) were observed longitudinally at 15, 18, 21, and 25 months in the “looking-while-listening” procedure as they looked at pictures while listening to speech naming one of the pictures. Analyses of the time course of eye movements revealed significant increases in the efficiency of children’s comprehension over this period. Parental-report data on productive vocabulary were gathered at 3-month intervals. Speed and accuracy in spoken word recognition at 25 months were correlated with numerous measures of lexical and grammatical development from 12 to 25 months. Analyses of nonlinear growth curves showed that children who were faster and more accurate in online comprehension at 25 months were those who had accelerated more rapidly in vocabulary growth across the second year.
Picking Up Speed in Understanding: Speech Processing Efficiency and Vocabulary Growth across the Second Year

Children in the early stages of learning a language are often credited with "acquiring" new vocabulary, as if words were things that 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. But we also speak of acquiring skills such as playing the piano, where the emphasis is on gradual mastery rather than possession. It is increasingly evident that learning to recognize, understand, and appropriately speak a new word is a gradual process. Not only do infants respond meaningfully to more and more words over the second year of life, 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 in very young language learners 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 psycholinguistic research with adults (Tanenhaus, Magnusen, Dahan, & Chambers, 2000), it is now possible to monitor the time course of spoken language understanding even in very young children. Using a looking-while-listening procedure with infants from 15 to 24 months of age, Fernald, Pinto, Swingley, Weinberg, & McRoberts (1998) found that both speed and accuracy in spoken word understanding increase dramatically over the second year. In that study, infants looked at pictures of objects while listening to speech naming one of the objects with a familiar word. Fifteen-month-olds responded inconsistently and shifted their gaze to the picture that had been named only after the end of the target word. In contrast, 24-month-olds were more reliable and much faster 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 the "vocabulary spurt", accelerating in the rate at which they add new words to their expressive vocabulary (Bloom, 1973), they are also becoming much more efficient in recognizing and interpreting familiar words in fluent speech.

How are these changes in children's receptive language abilities over the second year related to the rapid development in productive skill that typically occurs during the same time period? This was the overarching question that motivated the present research. To explore this question, we designed a longitudinal study of the development of receptive and productive language abilities in English-learning infants from 12 to 25 months of age. The first goal of this research was to replicate and extend the cross-sectional findings of the Fernald et al. (1998) study by documenting developmental changes in the speed and accuracy of spoken word recognition in the same infants across the second year. The second goal was to determine whether our recently developed online measures of speech processing could be used to study
Efficiency in understanding

individual differences in the development of speech processing abilities, asking whether infants who are relatively faster and more accurate in word recognition at one age are also relatively faster and more accurate at other ages as well. The third and most important goal of this research was to determine how the development of competence in spoken language understanding over the second year relates to development in other domains of linguistic competence such as non-linear patterns of growth in productive vocabulary and the emergence of grammatical abilities.

Efficiency in Spoken Word Recognition by Adults

The effortlessness with which adults can make sense of rapidly delivered speech in their native language belies the complexity of the task that initially faces the infant. To follow a typical conversation, skilled listeners must rapidly integrate acoustic information with linguistic and contextual knowledge, processing continuous strings of speech sounds at rates of 10 to 15 phonemes per second (Cole & Jakimik, 1980). Extensive research with adults has shown that the ability to process speech continuously is central to this remarkable efficiency. By making use of phonetic information as it becomes available, the listener can identify spoken words very rapidly, often before their acoustic offset (Marslen-Wilson, 1987; McClelland & Elman, 1986). Psycholinguistic studies using online measures of spoken word recognition reveal that listeners evaluate hypotheses about word identity incrementally, based on 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 is spoken. Research by Grosjean (1985), Marslen-Wilson (1987) and others has shown that the efficient processing of word-initial information can facilitate rapid decisions about the identity of many spoken words. Moreover, activation of the possible meanings of a spoken word typically begins within 150 ms of word onset (Zwitserlood, 1989) and adults can rapidly integrate the acoustic signal with linguistic knowledge and information from the visual scene in interpreting the spoken sentence as it unfolds in time (Allopenna, Magnuson, & Tanenhaus, 1998). 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. Fluent understanding of a speaker's meaning requires the ability to anticipate what is coming next in the stream of speech by integrating many different sources of linguistic knowledge with contextual information.

Efficiency in Speech Processing by Infants

Extensive research on the early emergence of speech processing skills shows that infants in the first year are attentive to speech patterns relevant to language structure. By the age of 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 fluent speech (Jusczyk, 1997). Even when familiarized only briefly with strings of nonsense syllables, 9-month-olds extract word-like units by noticing which syllables co-occur (Saffran, Newport, & Aslin, 1996). These studies show that over the first year infants become skilled listeners, able to make distributional analyses of phonetic
features of spoken language. Ten-month-olds listen preferentially to words likely to be familiar in the speech they hear (Halle & de Boysson-Bardies, 1994, 1996), indicating that they have some kind of acoustic-phonetic representation for frequently heard sound patterns. Although such accomplishments are often cited as evidence for early "word recognition", this selective response to familiar words can occur without any association between particular sound patterns and meanings, reflecting pattern detection abilities prerequisite for recognizing words in continuous speech.

By the end of the first year, children typically do begin to associate sound patterns with meanings, speaking a few words and appearing to understand many more. In the following months their progress in understanding is revealed through increasingly differentiated behavioral and verbal responses to speech. However, because the processes involved in comprehension are only partially and inconsistently apparent through the child's responses to speech in everyday situations, developments in receptive language competence are less accessible to observation than developments in speech production. Until recently, developmental studies of early comprehension have had to rely on off-line measures, responses made after the offset of the speech stimulus that do not tap into the real-time properties of spoken language. Studies of incremental processing by adults use more sensitive online measures that monitor the time course of the listener's response in relation to key points in the speech signal (Marslen-Wilson, 1987; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1996). Because comprehension occurs rapidly and automatically without time for reflection, it is important to examine the listener's interpretation during speech processing and not just afterward. By measuring infants' eye movements as they look at pictures while listening to speech, we can now track their understanding from moment to moment as the sentence unfolds. The Fernald et al. (1998) study provided the first evidence for age-related changes in the speed and accuracy of spoken word recognition in very young language learners. Other studies using this looking-while-listening procedure have shown that 2-year-olds are able to monitor speech continuously and attend to word-initial information in identifying words (Swingley, Pinto, & Fernald, 1999), and that by 18 months infants can already associate a familiar spoken word with the appropriate picture after hearing only the first 300 ms of the word (Fernald, Swingley, & Pinto, 2001). Moreover, even younger infants encode the phonological forms of familiar words in detail and find mispronunciations harder to recognize than correct pronunciations (Swingley & Aslin, 2002).

Relation of Receptive and Productive Competence in Early Language Development

In the 18th century diary study that provided the first systematic observations of a child's early language development, Tiedemann (1787/1927) observed that his son appeared to recognize a few words at 8 months of age, several months before he began to speak. Darwin (1877) made a similar observation a century later, noting that at 7 months his child understood a single word, his nurse's name, five months before uttering his first word. More recent studies based on larger samples and more reliable observational measures have confirmed these early estimates, showing that the first signs of comprehension are generally evident around 8 months of age with speech production typically beginning around the first birthday (Bloom, 1973; Snyder, Bates, & Bretherton, 1981). The research methods have changed considerably, but children evidently have not. The most extensive study of the rate of lexical development in children learning American English was designed to provide norms for the MacArthur Communicative Development Inventory [CDI], a set of parental report questionnaires used to
assess productive vocabulary size and various measures of grammatical competence in children from 8 to 36 months of age (Fenson et al., 1994). Parents were also asked to report on the words understood by the child, but only through the age of 16 months, at which point it becomes very difficult to obtain reliable estimates of the child's receptive vocabulary. Fenson et al. found that by the time children are able to speak 50 words, around the age of 16 months on average, their comprehension vocabulary is typically reported to be about four times as large.

The rate of learning to speak new words varies considerably among children in this age range, and research on the relation of growth in productive vocabulary to receptive language development has not yielded consistent results. Links between early comprehension and production have been investigated primarily through observational studies showing small correlations between the numbers of words understood and spoken by children in the second year (Bates, Bretherton, & Snyder, 1988). In a longitudinal study of language use by six children from 6 to 18 months, Harris, Yeeles, Chasin, & Oakley (1995) noted that words that were contextually flexible in production tended also to be so in comprehension, and words that were context-bound also tended to be so in both modalities. However, they also found substantial individual differences in the rate at which comprehension developed and in the extent of the lag between comprehension and production. The limited evidence for associations between early comprehension and later expressive abilities revealed in these studies only hint at a potential connection between fluency in understanding and speaking. And because estimates of receptive language competence have relied largely on observational methods and parental report, they reveal nothing about the cognitive processes that might underlie such a connection.

Several researchers have suggested ways in which perceptual and cognitive components of speech processing skills might be related to early vocabulary growth, although there is disagreement about the likely direction of influence. Some have argued that infants' early representations of word forms are holistic and imprecise and that it is vocabulary expansion that forces a shift to more segmentally-based processing, as the child faces the need for more efficient differentiation and storage of new words (Charles-Luce & Luce, 1990; Walley, 1993). Stager & Werker (1997) agree that younger infants have difficulty in distinguishing phonetic detail when trying to map new word forms onto meanings but argue that this limitation is short-lived; thus by the time of the vocabulary spurt infants develop additional cognitive capacities that give them the resources to discriminate and learn words more quickly. Coming from another direction, Bloom (1973) suggests that children's growing ability to use more diverse cues to retrieve words in memory is a critical aspect of cognitive development leading to more rapid lexical growth. Then as vocabulary expands, the child is increasingly able to exploit linguistic cues in recalling stored words. Although speech processing is just one aspect of the developments Bloom describes, a clear implication of her argument is that enhanced efficiency in the information-processing skills underlying comprehension would accelerate growth in expressive language.

One way to investigate possible links between early speech processing skill and lexical development is to ask whether differences among children in the emergence of abilities essential for spoken language understanding are associated with differences in the size of their productive lexicons, either concurrently or at a later time. Although studies of speech perception skills emerging in the prelinguistic period now number in the hundreds (Aslin, Jusczyk, & Pisoni, 1998; Aslin, Pisoni, & Jusczyk, 1983), these studies have focused almost exclusively on group
data rather than individual differences, and thus have not attempted to relate infants' performance on speech perception tasks to other measures of linguistic development. One recent exception is a study by Werker, Fennell, Corcoran, & Stager (2002) investigating the relation of infants' vocabulary size to their ability to learn phonetically similar novel words. They found that 14-month-olds confused similar-sounding words under conditions when they were attempting to map the novel words onto novel objects, although under less taxing conditions the same words were readily discriminated by infants at this age. Werker et al. interpret this finding as evidence that younger infants have insufficient cognitive resources to attend to fine phonetic detail in a word-learning situation. When they examined performance on this task in relation to productive vocabulary size, they found that those 14-month-olds who were successful in word learning were those with relatively larger vocabularies. However, at 18 months infants had less difficulty overall in learning the novel words and there was no relation between performance and vocabulary size. Werker et al. conclude that productive 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 only in the earliest stages of building a vocabulary.

Three recent studies of spoken word recognition provide the only experimental evidence to date for a relation between speech processing skills and vocabulary size in infants around the time of the vocabulary spurt. Fernald et al. (2001) used the looking-while-listening paradigm to demonstrate that 18- and 21-month-old infants could recognize words based on partial phonetic information. When familiar words were presented both as whole words in intact form and as partial words in which only the first 300 ms of the word was heard, infants at both ages recognized partial words as quickly and reliably as whole words. Further analyses revealed that infants with > 100 words in their productive vocabulary were more accurate in identifying familiar words than were infants with < 60 words. Grouped by response speed, infants with faster mean reaction times were more accurate in word recognition and had larger productive vocabularies than infants with slower response latencies. Fernald (2002) also found convergent results in a study of online understanding of verbs by 26-month-old children; 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 exploring a related question, Zangl, Klarman, Thal, Fernald, & Bates (in press) presented infants between 12 and 24 months with words that were either naturally spoken or perceptually degraded through filtering or time compression. Using a variant of the looking-while-listening procedure with different methods of data analysis, Zangl et al. also found accuracy and speed in word recognition to be correlated with productive vocabulary size.

The results of these three studies provide preliminary evidence that increased speed and efficiency in understanding spoken language may be associated with lexical growth in the age range when vocabulary typically starts to expand more rapidly. It is important to note, however, that two other recent studies using eye-movement measures to track word recognition found no association between vocabulary size and success in online speech processing. When Swingley & 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 expressive vocabulary development. The inconsistencies in these findings could perhaps be explained by differences in study design and in the age of participants. By following a large sample of children longitudinally and examining their emerging speech processing skills in
Efficiency in understanding

8

relation to measures of their lexical and grammatical development at multiple time points, the present research aims to provide some clarification.

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 quite different perspectives. One goal was to replicate the earlier finding that efficiency in online speech processing increases steadily across the second year of life (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 in this procedure 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 reaction time was in fact related to linguistic processing, or whether infants simply make faster eye movements as they grow older. This latter explanation seemed unlikely given research showing that by 12 months of age infants’ mean visual reaction time is approaching adult values (Canfield, Smith, Brezsnayak, & Snow, 1997; Rose, Feldman, Jankowski, & Caro, 2002). However to explore this possibility, we added a control measure to assess visual reaction time in a non-linguistic task, a procedure in which infants tracked visual images alternating in a random pattern from side to side in silence. We expected to find an increase in accuracy and a decrease in reaction time in the word recognition task across the four time points, with no appreciable change in visual reaction time in the non-linguistic task.

A second goal of this research was to investigate the stability and independence of our measures of speed and accuracy in online speech processing. Since all previous studies 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, nor about the extent to which they are intercorrelated. Do those infants who respond most reliably at any given age also respond most rapidly? And are those infants with the fastest mean reaction times at one age among the fastest at later ages? Exploring such questions will help 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 and most central goal of this research was to investigate the validity of the online measures used here and in other recent studies, by asking how gradual developments in the efficiency of spoken language processing between 15 and 25 months relate to progress in other aspects of language competence emerging across the second year. In particular, to what extent are speed and accuracy in receptive language processing associated with growth in productive vocabulary as well as with grammatical development? This question was addressed by examining relations between measures of infants’ performance in online word recognition at each age with measures of their vocabulary size and grammatical competence, as reported by parents at each age on the MacArthur CDI. As a convergent measure of lexical knowledge, the Peabody Picture Vocabulary Test was administered to children at 25 months.
Method

Participants

The initial sample of participants consisted of 63 full-term infants recruited at the age of 8 months through a university hospital. Thirty-seven of the infants (58.7%) were male. The study sample was from a predominantly middle class population, with 85.7% of families Caucasian, 12.7% mixed Asian, and 1.6% mixed Hispanic/Caucasian. The highest educational level attained by at least one of the parents in the family was: some graduate-level education (62%), an undergraduate degree (35%), and some college (3.2%). All infants were from families in which English was the primary spoken language.

Participants visited the laboratory for testing at the ages of 15, 18, 21, and 25 months. Data for some subjects at each age were not included in the analyses for that particular sample point, for the following reasons: (1) fussiness and refusal to sit through the looking-while-listening procedure (15 months: n = 2; 18 months: n = 4; 21 months: n = 3; 25 months: n = 3); (2) inattentiveness, defined as not fixating one of the stimulus pictures at target word onset on at least 70% of trials (15 months: n = 12; 18 months: n = 6; 21 months: n = 6; 25 months: n = 6); (3) experimenter error (18 months: n = 1); (4) missed session due to travel (18 months: n = 1). At 18 months, four of the participants dropped out of the study permanently due to illness (n=1) or moving away from the area (n=3), and 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 both valid 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 of the participants who completed the word-recognition procedure did not complete the Peabody Picture Vocabulary Test due to fussiness or fatigue. Unless indicated otherwise, all analyses include the data from all infants who completed that portion of the procedure.

Parental Report Measures of Lexical and Grammatical Development

When their infants reached the ages of 12, 15, 18, 21, and 25 months, parents were mailed the appropriate version of the MacArthur Communicative Development Inventory (CDI) (Fenson et al., 1994). At 12 and 15 months, they were asked to fill out the “infant version” of the checklist, CDI: Words and Gestures, reporting on the words comprehended and produced by the child, as well as their child's use of communicative gestures. At the later ages, parents were asked to fill out the “toddler version” of the checklist, CDI: Words and Sentences, reporting on the child's productive vocabulary and various measures of grammatical proficiency. At the 12-month time point, parents returned the completed CDI to the laboratory by mail; at later ages they brought the completed CDI form with them when they visited the laboratory with their child for testing.

The CDI yields several different measures of vocabulary size and/or grammatical skill: Receptive vocabulary. The measure of receptive vocabulary was the number of words understood by the child according to parental report on the infant version of the CDI. Because this measure of comprehension becomes increasingly less reliable beyond the age of 16 months, it is not included on the toddler version of the CDI and thus was limited to the 12- and 15-month samples in this study.
Productive vocabulary. The measure of productive vocabulary was the number of words spoken by the child, as reported by parents on the CDI at 12, 15, 18, 21, and 25 months.

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 toddler version of the CDI at 18, 21, and 25 months.

Grammatical complexity. The grammatical complexity measure on the toddler CDI is designed to assess the child’s use of word combinations and closed class morphemes. Parents are asked to choose one of two utterances typical of their child's speech productions, e.g., *I go away* vs. *I went away* or *Truck table* vs. *Truck on table*. 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.

Peabody Picture Vocabulary Test (PPVT)

At 25 months, the PPVT-Revised (Dunn & Dunn, 1981) was administered following the looking-while-listening procedure. Standard procedures for administration of the PPVT were followed\(^1\). The child sat on the parent's lap at a low table, across from the experimenter. A video camera recorded the test stimuli and the child's responses. The experimenter began by showing the child some pictures and 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.* If the child did not touch one of the pictures, the experimenter asked the parent to demonstrate, demonstrated herself, or asked the child to use an animal or toy to touch the picture. Once the child learned the task, the actual procedure began. The child was then shown a series of black-and-white illustrations arranged four per page. One picture on each page was the designated target, with the location of the target picture counterbalanced across pages. When the child had looked at the page for about 10 sec, the experimenter called attention to one of the pictures by saying "Point to the [target]". If the child failed to respond by pointing to one of the pictures within 10 sec, the instruction was repeated again. Testing was continued until the child chose the incorrect picture on 6 out of 8 consecutive pages or refused to participate further. The child's score was their ceiling score minus errors.

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

Speech stimuli. In designing the stimulus sets to be used at different ages, we had two competing goals. On the one hand, it was important to include the same or comparable tokens at each age for purposes of controlled comparison of age-related differences in the efficiency of speech processing. In the Fernald et al. (1998) study, it was for this reason that infants were tested at all three ages on the same stimulus set, consisting of four words already familiar to the youngest infants. Because this earlier study was cross-sectional in design, there was no need to worry about effects of repeated testing. On the other hand, it was also important to make the stimulus sets increasingly challenging as the children grew older in the present longitudinal study, to keep them interested and engaged in the task. Moreover, because we were exploring individual differences in speech processing abilities, it was critical to include some types of stimuli known to be difficult for children at each age in order to avoid ceiling effects.
The speech stimuli at all ages consisted of prerecorded sentences containing target words occurring in final position 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 and maintain interest. 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 only partial phonetic information develops between 18 and 21 months of age and is also related to level of lexical development Fernald et al. (2001). In addition to these 16 trials, 13 filler trials were included. At 21 months, doggie, baby, birdie, and kitty, were presented both as whole words and as truncated words (daw, bei, ber, kl) 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 constrained (drink, eat) or semantically neutral (take, look at). Four filler trials were included at this age. Finally, at 25 months the standard set of “easy” target words (doggie, baby, ball, car) was augmented by target words that are typically not learned until later in the second year (monkey, cow, flower, tree, animal), all presented as whole words. To diversify the sentence frames, adjectives (nice, pretty) preceded target words on 8 trials. Four filler trials were included at 25 months.

To prepare these stimuli, a female native speaker of English produced several tokens of each sentence, matching them as closely as possible in intonation contour. These candidate stimulus sentences 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 based on comparability of the acoustic measurements of the 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 images of brightly colored objects corresponding to the target words, taken from photographs in children's books and magazines. Two different object tokens were used for each target word at every age. All images were closely matched in size and brightness, and were presented on a gray background on 15" diagonal computer monitors.

**Apparatus.** The looking-while-listening procedure was conducted in a testing booth in a sound-treated room. The testing booth was constructed of three panels covered in cloth and open on the fourth side. The side panels of the testing booth measured 1 x 2 m, and the front panel measured 1 x 1.2 m. Mounted in the front panel were two computer monitors and a loudspeaker for presentation of visual and auditory stimuli. During testing the infant sat on the parent's lap facing the monitors. A curtain suspended between the two sides of the booth hung down behind the infant's head, obstructing the parent's view of the monitors while still allowing infant access to the parent during the test session. The monitors were separated horizontally by 60 cm and were positioned at the infant's eye level. The loudspeaker was located on the floor centered between the monitors. A video camera mounted behind the front panel focused on the infant's face. The camera was connected to a VCR in the adjacent control room where the computer controlling the experiment was also located.
Procedure. Each visit to the laboratory began with a 20-min familiarization period prior to the test session. During this time the experimenter talked with the parents, obtained informed consent, and interacted with the child. When parent and child both appeared comfortable they were seated in the testing booth. The lights in the room were dimmed as identical pictures appeared on the two computer monitors in the booth to attract the infants' attention. A second experimenter located in the control room spoke briefly to the child over the loudspeaker to acquaint her with the sound source and the positions of the monitors. When the child was relaxed and attentive, the experimental session began. Trial types were presented in a quasi-random 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 sec prior to presentation of the speech stimulus, continuing for 1 sec after the offset of the sound stimulus. During the 1-sec intertrial interval the monitor screens were black. The duration of the test session was about 4 minutes.

Coding eye movements. During the testing session a digital time-code accurate to 33 ms was recorded onto the videotape, along with a visual marker indicating the onset of the speech stimulus on each trial. The tape for each session was digitized by Adobe Premiere and coded off-line by highly trained observers using custom software on a Macintosh computer. Coding was done without sound and observers were blind to the trial type and side of the target picture. For each trial, the highly trained 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 then 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.

Inter-observer reliability checks were conducted routinely for all coders. The reliability analysis was designed to focus on those trials in which shifts in gaze from one picture to the other occurred most frequently, for two reasons: first, on these trials there was greater potential for disagreement among coders, as opposed to trials on which the infants continued to fixate on one picture only; and second, trials involving shifts are also most crucial for the analysis of speed of processing. To prepare for the reliability analysis, tapes for 20% of the infants at each age were prescreened by an observer who was otherwise uninvolved in the analysis. This observer viewed the tape in real time to identify those trials on which shifts most frequently occurred, assigning this subset of trials to be independently coded by the reliability coder. For the 15-month sample, 7 trials were coded for each of 13 infants; on 90.4% of these trials, all 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 all shifts were within one frame of each other. At 21 months, 8 trials for each of 14 children were coded twice; 91.4% of all shifts were within one frame of each other. And at 25 months, 8 trials for each of 13 children were coded twice; 90.8% of trials were within one frame of each other.

Measures of Accuracy and Speed in Online Speech Processing

Because children cannot know at the beginning of a trial 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 by chance already be looking at the target picture (target-initial trials). On distracter-initial trials, the correct response is to shift from the distracter to the target picture, while on target-initial trials, the correct response is to continue
looking at the target picture without shifting away. Thus, a child with perfect accuracy across trials would shift gaze to the target picture 100% of the time on distracter-initial trials, and would never shift away from the target picture on target-initial trials. Shifts on all distracter- and target-initial trials were assessed within the time window from 300-1800 ms following target word onset. This window was determined by Fernald et al. (2001) based on detailed analyses of the distributions of shifts for 18- to 21-month-olds. 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 1800-3000 ms after target word onset were excluded because these delayed shifts were considered to be outliers that were less clearly in response to the target word.

**Proportion of correct shifts to target picture.** This measure of correct shifts represents the number of first shifts to the target picture that occurred within the 300-1800 ms window following target word onset on distracter-initial trials, as a proportion of all first shifts within that time window on distracter- and target-initial trials combined.

**Proportion of incorrect shifts to distracter picture.** This measure of incorrect shifts represents the number of first shifts to the distracter picture that occurred within the 300-1800 ms window following target word onset on target-initial trials, as a proportion of all first shifts within that time window on distracter- and target-initial trials combined.

**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 spends fixating the target picture as a percentage of total time fixating either picture during the 300-1800 ms window from target word onset.

**Reaction time (RT).** Speed of response to the spoken word was calculated based on distracter-initial trials on which a correct shift occurred within the 300-1800 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 for this is that children at all ages differed in how frequently they by chance happened to be looking at the distracter picture at target word onset, and thus differed in the numbers of trials on which reaction time could be calculated. Another reason is that there were fewer trials to begin with at younger ages, and in general younger infants shifted less reliably than at older ages. 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 the distracter to the target picture. In contrast, a mean RT for the same infant at 21 or 25 months might reflect performance on 6 - 10 or more trials. However, at all ages the mean RT scores for each child was based on at least two shifts, since the data from those with only a single shift were excluded from the relevant analyses.

**Visual Reaction Time (VRT) Procedure**

To control for both individual differences and age-related changes in infants’ ability to shift from one picture to another when no language processing was involved, reaction time to a peripheral visual stimulus in a nonlinguistic context was assessed longitudinally. We used a modified version of the Visual Expectation Paradigm developed by Haith, Hazan, & Goodman.
Efficiency in understanding

(1988), a procedure designed to measure infants' raw reaction time 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 visual reaction time (VRT) procedure. The infant sat on the caregiver’s lap at a distance of 72 cm from a single monitor computer (36 cm diagonal), positioned centrally at the infant’s eye level. To reduce the potential for influence by the caregiver, the parent holding the child was instructed to close her eyes during presentation of the stimulus. Infants viewed colorful, computer-generated dynamic pictures that appeared on the left and right sides of the computer monitor at approximately a 7.9° angle from the infant’s eyes.

Visual stimuli. The stimuli consisted of ten different pictures (smiley face, diamond, wheel, circle, pentagon, triangle, bullseye, flag, star, and square) each approximately 4 cm square and each with a different dynamic pattern (e.g., spinning, falling, blinking, expanding). These pictures were displayed sequentially for 700ms in random order on either the right or left side of the screen, with the constraint that the stimulus appeared no more than three times in a row on the same side. Each infant saw a total of 20 pictures in a random sequence with no interstimulus interval (i.e., 0 ms ISI). This procedure lasted approximately 2.5 minutes.

Coding eye movements. During the testing session a digital time-code accurate to 33 ms was recorded onto the videotape, along with a visual marker indicating the onset of the stimulus on each trial. The tape for each session was digitized by Adobe Premiere and coded off-line by highly trained observers using custom software on a Macintosh computer. For each trial, the 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. The child’s latency to orient to the visual stimulus was calculated on those trials on which the stimulus switched from one side to the other.

Visual reaction time (VRT). Mean VRT was defined as the average latency to begin an eye movement towards the stimulus after the stimulus appeared. Mean VRT was calculated based on only those eye-movement latencies greater than or equal to 133 ms (Canfield et al., 1997), with a cutoff of 700 ms corresponding to the duration of each picture (Dougherty & Haith, 1997). Thus the mean VRT measure for each infant represents the average of all response latencies between 133 and 700 ms.

Results

The results are presented in five sections related to the major questions addressed in this research. The analyses in the first section describe changes in speech processing efficiency between the ages of 15 and 25 months in our longitudinal sample of English-learning children. The second section focuses on the interrelation of these measures of speed and accuracy in spoken language comprehension and their stability over time. The third section explores the relation of speed of spoken word recognition to visual reaction time in a non-linguistic task. The fourth section examines 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 patterns of non-linear growth in the size of children's productive vocabulary are related to the development of efficiency in online speech processing at 25 months.

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

One goal of this research was to replicate and extend the cross-sectional findings of Fernald et al. (1998) using a longitudinal design. To compare speech processing efficiency in the same children at different ages, we analyzed measures of speed and accuracy at 15, 18, 21 and 25 months using repeated measures mixed model analyses of variance.

Reaction Time. 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 following the onset of the target word. Mean reaction time decreased significantly with age, with an overall decrease of 223 ms across the 10-month period: 15 months ($M = 994$ ms, $SD = 263$ ms), 18 months ($M = 993$ ms, $SD = 245$ ms), 21 months ($M = 836$ ms, $SD = 235$ ms), 25 months ($M = 771$ ms, $SD = 128$ ms), $F(3,114) = 16.8, p < .001$. Planned contrasts indicated that decreases in reaction time were significant between 18 and 21 months ($p < .001$) and 21 and 25 months ($p < .05$). Note also that there was considerably less variance in RT performance (reported here as $SD$) at 25 months than at the previous three assessments, suggesting that indices of processing speed may be more stable at the older rather than younger ages. There was no significant decrease in mean RT from 15 to 18 months ($p > .10$). This latter finding is at odds with the Fernald et al. (1998) cross-sectional finding that 18-month-olds were on average 150 ms faster than 15-month-olds to orient to the target picture in response to a familiar spoken word, a discrepancy to be considered in the Discussion section.

Accuracy. Children’s accuracy in word recognition was assessed in two different ways. Because the behavior involved in a correct response differs fundamentally depending on where the child happens to be looking at the onset of the target word, we first evaluated distracter- and target-initial trials separately. For each individual at each age, the proportion of correct shifts and the proportion of incorrect shifts were calculated, based on distracter-initial trials and target-initial trials respectively. As shown in Figure 1a, the tendency to shift correctly to the target picture on distracter-initial trials increased significantly with age: 15 months ($M = .53, SD = 30$), 18 months ($M = .59, SD = .32$), 21 months ($M = .70, SD = .24$), 25 months ($M = .85, SD = .16$), $F(3,109) = 23.2, p < .001$. Planned contrasts indicated that this effect was due to significant increases in accuracy between 18 and 21 months ($p < .03$) and 21 and 25 months ($p < .001$), with no reliable increases observed between 15 and 18 months of age ($p > .25$). As also shown in Figure 1b, the tendency to shift incorrectly from the target to the distracter picture did not vary reliably with age: 15 months ($M = .29, SD = .23$), 18 months ($M = .35, SD = .30$), 21 months ($M = .31, SD = .20$), 25 months ($M = .36, SD = .19$). Thus, across the age range studied, children were able to respond correctly by maintaining their gaze when they happened to be already looking at the target picture at the onset of the target word. However, when children happened to be looking at the distracter picture at target word onset, they were more likely to shift correctly away from the distracter to the target picture at older ages than at younger ages. On this measure too the variability declined substantially with age. While this may reflect improved reliability of measurement, it may also be a consequence of the fact that by 25 months accuracy was approaching ceiling levels. At this age nearly 30% of the children correctly shifted to the named picture on 100% of codable trials.
The second approach to operationalizing accuracy was to measure the total amount of time the child fixated the target picture during the relevant time window following the onset of the target word. The proportion of looking time to the target picture increased significantly with age: ($M = .54, SD = .08$), 18 months ($M = .56, SD = .10$), 21 months ($M = .60, SD = .09$), 25 months ($M = .66, SD = .07$), $F(3,108) = 14.2, p < .001$. This finding suggests that older children tended to spend relatively more time fixating the target picture rather than the distracter picture, regardless of where they were looking at the start of the auditory signal.

Both of these accuracy measures are valid; however, because the proportion of looking time to target is based on data from both target- and distracter-initial trials combined, it is partially redundant with the proportion of correct shifts measure used in the first accuracy analysis. Moreover, the differential patterns of responding on distracter- and target-initial trials suggest that the ability to correctly shift gaze from distracter to target picture undergoes developmental change during this period, unlike correct responses that require the child merely to keep looking at the target without shifting away. Since the measure of proportion of looking time to the target picture includes these less informative target-initial trials, it is likely to be less sensitive to developmental changes in accuracy than the measure of proportion of correct shifts to the target picture on distracter-initial trials. For this reason we will use proportion of correct shifts as our measure of accuracy in the remaining analyses in this paper.

The Stability and Intercorrelation of Online Speech Processing Measures

Stability of online measures of speed and accuracy over time. To examine the stability over time of the online measures used in the looking-while-listening procedure, we asked whether a child’s accuracy and speed assessed at one time point was correlated with the same measure at the next time point, 3-4 months later. The month-to-month correlations for RT and proportion of correct shifts to target picture (accuracy) from 15 to 25 months are shown in Table 1. Although the RT measure was not stable between 21 and 25 months ($r = .21$, ns), correlations between mean RT at 15 and 18 months ($r = .35, p < .05$), and at 18 and 21 months ($r = .36, p < .05$), were both reliable. Accuracy was also moderately stable from one test session to the next. The correlations between the proportion of correct shifts at 15 and 18 months was not significant ($r = .24, ns.$), however, the associations between 18 and 21 months ($r = .48, p < .01$), and 21 and 25 months ($r = .43, p < .01$) were both reliable. These results indicate that in spite of differences in the particular stimulus sets used to assess spoken language comprehension at the different ages, general characteristics of children’s online processing was relatively consistent across this age range.

Intercorrelations of online measures of speed and accuracy. The next analyses examined the extent to which our online measures were correlated with each other at each age. Results from previous cross-sectional studies (Fernald et al., 1998; Zangl et al., in press) suggest that developmental changes in online processing abilities should be reflected in both faster and more accurate responding. To explore the relation of speed and accuracy, we asked whether the proportion of correct shifts was negatively correlated with mean RT for children at each age. The results indicated that at 15 months, measures of speed and accuracy were not significantly correlated ($r = - .23$, ns). However, at all later ages, the correlations between speed and accuracy in word recognition were significant: 18 months ($r = - .40, p < .01$), 21 months ($r = - .34, p < .05$), and 25 months ($r = - .53, p < .001$). Thus, efficiency in speed and in accuracy, two independent
components of online spoken language processing, go hand-in-hand toward the end of the second year.

Relation of RT in Online Speech Processing to VRT in a Non-Linguistic Task

The visual reaction time (VRT) measure was included in this study to assess children’s speed of visual orienting in a non-linguistic task. Because the VRT task also required infants to shift gaze rapidly from one picture to another, it was similar to the looking-while-listening procedure used here to assess spoken language comprehension. The question motivating this control measure was whether the dramatic increase in speed of word recognition over the second year could be accounted for in part by a more general increase in the speed of visual orienting. A comparison of mean VRT scores (in ms) at 15 months ($M = 342, SD = 45$), 18 months ($M = 335, SD = 62$), and 21 months ($M = 332, SD = 61$) revealed no significant change with age.

Although changes with age were minimal, mean VRT proved to be a highly stable measure over time. As shown in Table 1, correlations between children’s mean VRT scores at 15 and 18 months ($r = .56, p < .01$) and at 18 and 21 months ($r = .71, p < .001$) were both reliable. However, we found no relation between speed of visual orienting to a silent peripheral image in the VRT procedure and speed of orienting to a named picture in the looking-while-listening procedure. Correlations between children’s mean RT in the word recognition task and their mean VRT in the non-linguistic task at 18 months ($r = -.08$), 21 months ($r = -.19$), and 25 months ($r = -.11$) were haphazard and non-significant. This lack of relationship between the verbal and non-verbal measures suggests that the developmental increases in response speed observed in the looking-while-listening procedure reflect improved efficiency in interpreting visual stimuli in the context of linguistic reference, rather than increased speed in visual processing more generally.

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 competence in spoken language understanding over the second year relates to development in other domains of linguistic competence such as growth in productive vocabulary and the emergence of grammatical abilities. Table 2 presents the descriptive statistics for off-line measures of vocabulary and grammar from the MacArthur CDI and the PPVT-R. As expected, reported vocabulary comprehension increased substantially between 12 and 15 months of age, $t(55) = 11.5, p < .001$, and at both time points children were reported to understand more words than they produced ($p < .001$). Further, reported vocabulary production increased from 12 to 25 months of age, with children moving from fewer than 10 words at 12 months to an average of almost 400 words at 25 months, $F(4, 196) = 159.7, p < .001$. Improvements across the period from 18 to 25 months were also observed in both grammar measures: complexity: $F(2,106) = 53.7, p < .001$; M3L: $F(2,106) = 79.5, p < .0001$. To illustrate, at 18 months fewer than half of the children (44.6%) were reported to produce any word combinations (M3L = 1.0). However, by 25 months, nearly all of the children (91.1%) were doing so. At the same time, there were considerable individual differences in these measures, consistent with the degree of variation that is typical of children in this age range (Fenson et al., 1994).

Concurrent correlations between RT and accuracy scores assessed in the looking-while-listening procedure and offline measures of language abilities at 15, 18, 21 and 25 months are
presented in Table 3. Speed of processing was significantly correlated with all vocabulary and grammar measures as well as the PPVT at the 25-month time point. Consistent with previous studies (Fernald et al., 2001; Zangl et al., in press), this finding suggests that two-year-old children who responded more quickly to the labeled picture in the word recognition procedure were also more likely to have larger productive vocabularies, greater reported complexity of grammatical forms, and better scores on a standardized test of lexical knowledge. At the same time, concurrent relationships at the younger ages were much weaker, and none was statistically significant. Turning to the accuracy measure, Table 4 presents moderate and significant relations between online and offline measures at the 21-month ($r_s = .32$ to $.50$) and 25-month ($r_s = .35$ to $.60$) time points. Thus, the children at these ages who were more accurate in online spoken word recognition were also the children with larger productive vocabulary and grammar scores on the offline measures. Again, relations between online and offline measures were weak at the younger ages.

For both measures then, concurrent correlations between the online and offline measures were strongest when the children were older. These results provide further evidence that links between online measures of language processing and offline measures of language abilities are most evident in those children who are likely to be experiencing rapid growth in linguistic abilities. At the same time, the absence of correlations among these measures at younger ages suggests that early in acquisition speed and efficiency of online spoken understanding might be relatively independent of particular linguistic accomplishments. In light of the considerable individual differences that are evident during this period, the current longitudinal design has advantages over cross-sectional studies in its ability to disentangle these factors by tracking developmental changes across the period in the same group of children.

Given that both speed and accuracy in word recognition at 25 months were concurrently correlated with offline linguistic measures and performance on the PPVT, we next examined relations between speech processing skills at 25 months and reported measures of vocabulary and grammar assessed at all four of the previous time points. As shown in Table 5, those children with faster mean response speed at 25-months tended to be those who had had larger productive vocabularies at 12, 15, 18 and 21 months of age. Online RT at 25 months was also significantly related to off-line measures of grammar at 21 months. The relation between 25-month RT and grammatical measures at the 18-month time point was weak, not surprising given that children at this age are just beginning to produce complex grammatical forms. Table 6 presents parallel findings for the accuracy measure. Children who responded more reliably at 25 months were more likely to have had larger vocabularies at all previous assessments. Again, associations between 25-month online speech processing abilities and earlier grammatical accomplishments were most consistent at the 21-month time point. Taken together, these findings suggest that the skills involved in spoken language understanding at 25 months are related to both concurrent and prior levels of linguistic development.

To explore further the specificity of this link, the next series of analyses examined the relative contributions of age and level of language development in predicting children’s success in spoken word recognition. Given the strong intercorrelation between the lexical and grammatical measures ($r_s = .60$ - .80), language level was operationalized in terms of vocabulary size, as in previous studies (Bates & Goodman, 1997). At each age children were divided into
four groups according to the number of words in their reported production vocabulary on the CDI (0 - 99, 100 - 299, 300 - 499, and 500+ words). The proportions of children falling into each vocabulary group at each time point are presented in Table 7. Not surprisingly, at the 15 month time point all of the children spoke fewer than 100 words, whereas by 25 months only four children were still in the lowest vocabulary group.

Looking first at speed of word recognition, Figure 2 presents RT as a function of vocabulary level at the 15, 18, 21, and 25 month time points. This figure shows that reaction times decreased overall with age, consistent with the analysis reported above. At the same time, within each age group, reaction times also decreased as a function of vocabulary level. Because of the unequal distributions of participants in the age and vocabulary groups, the contributions of both factors were analyzed using a mixed model repeated measures ANOVA with both age (4) and vocabulary level (4) as within-subjects factors. We found that both factors were significantly related to the patterns of RT scores: age, \( F(3, 35) = 5.9, p < .002 \); vocabulary level: \( F(3, 34) = 6.1, p < .002 \). These findings suggest that the decreases in reaction time observed here were associated with both age-related and vocabulary-related changes that accompany language acquisition during this period.

We next examined accuracy scores as a function of vocabulary level within each of the four age groups, as shown in Figure 3. Again, it is clear that children tend to recognize words more accurately as they get older, but at each age accuracy also varies with vocabulary size. For example, at 21 months of age those children who still produced fewer than 100 words were also relatively inaccurate in spoken word recognition (\( M = .55 \)), while children at the same age who produced more than 500 words were highly accurate in identifying spoken words (\( M = .95 \)). These effects were analyzed using a repeated measures mixed model ANOVA with age group (4) and vocabulary level (4) as within-subjects factors. Results indicated a significant effect of vocabulary level on accuracy scores, \( F(3, 60) = 16.9, p < .0001 \). Unlike the results seen for reaction time, however, age was not a significant factor in the prediction of accuracy scores when vocabulary level was also taken into account, \( F(3, 25) = 1.1, ns \). This finding indicates that the significant increase in accuracy scores with age reported earlier was actually attributable to changes in vocabulary level rather than directly to maturational effects.

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 level of language development. Because age and language level were treated as within-subjects variables, the co-linearity in the variance observed at each time point was taken into account. However, these analyses provided no information about the observed variation in how individual children were changing over time; that is, there was no concept of trend. Clearly, one advantage of a longitudinal repeated measures design is that it enables us to assess not only group-level effects but also trajectories of change that characterize individual children across the period from 12 to 25 months of age.

However, to reveal these patterns of change it is necessary to use statistical techniques that evaluate the data in terms of a hierarchical or multilevel structure. At the first level (Level 1), repeated observations of individuals are assessed with respect to individualized growth
functions that can be 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 (12, 15, 18, 21 and 25 months) in the current design, it was also possible to explore the degree to which individual trajectories of vocabulary growth were characterized by increases in the rate of change (i.e., non-linear change or acceleration). Once these Level 1 parameters were defined for each individual, Level 2 analyses determined whether these parameters were predicted by other person-level (between-subjects) characteristics. The two Level 2 factors evaluated in our analysis were RT and accuracy as assessed at 25 months of age. These were chosen as predictor variables for two reasons. First, previous analyses had indicated that online measures assessed at 25 months were significantly correlated with both concurrent and prior linguistic accomplishments. Second, online performance at the 25 month time point appeared to provide more stable estimates of spoken word comprehension than at the earlier assessments, as indicated by reduced variability in performance across the sample. In sum, we used models that incorporated both Level 1 and Level 2 variables to examine the degree to which characteristics of individual trajectories of growth in vocabulary were predicted by the speed and accuracy of online spoken language understanding at 25 months of age.

The ability to examine higher-order characteristics of growth trajectories is particularly useful in the domain of lexical development. It is well known that vocabulary growth for many children increases in the rate of change toward the end of the second year, and that this phenomenon is best characterized in terms of both linear and non-linear functions (Goldfield & Reznick, 1990). Preliminary analyses of the current set of individual trajectories using regression-based curve estimation techniques reflected substantive linear and non-linear vocabulary growth. Statistical estimates of variance accounted for indicated that a linear model provided a strong degree of fit (mean $r^2 = .84$), averaging across individuals. However, linear models accounted for significantly less variance than both quadratic (mean $r^2 = .97$, $t(49) = 8.5$, $p < .001$) and logistic (mean $r^2 = .94$, $t(49) = 5.5$, $p < .001$) models. While it could be argued that logistic models might provide a more realistic estimation of intercept and end-point values (since logistic functions constrain the starting point to be zero and limit the upper-bound estimate), we chose to model nonlinear growth using a quadratic function. This was due to its relative mathematical simplicity, a comparable level of goodness of fit for most children, and the fact that few ceiling effects were observed in this particular sample of children; that is, few children reached the part of the curve that represents a leveling off or slowing down in rate of change.

Hierarchical growth curve modeling was conducted using the linear mixed models procedure in SPSS, with restricted maximum likelihood estimation used for all models. (This procedure is analogous to PROC MIXED in SAS and Hierarchical Linear Modeling [HLM], (Raudenbush & Bryk, 2002; Raudenbush, Bryk, Cheong, & Congdon, 2001)). When conducting growth curve analyses, models are evaluated in two stages. In the first stage, we examined Level 1 models with no Level 2 variables (i.e., the “unconditional” model). These models included as fixed effects an intercept (estimated vocabulary size at Time 1, which was 12 months of age in this case), slope (the estimated linear change over time) and acceleration (the estimated quadratic change over time). In order to capture the time structure of the multiple assessments of vocabulary across the five “waves” of data, the factor “time” was introduced as a 5-level (0-4) repeated measures factor. Exploratory evaluation of several within-subjects covariance structures indicated that the models with the best least squares estimates (goodness of “fit”
expressed as -2 Restricted Log Likelihood [2RLL] in “smaller-is-better” form) were those that assumed an unstructured within-person error covariance matrix.

Examination of this unconditional model demonstrated significant individual variation in intercept, \( F(1, 41) = 38.4, p < .001 \), and acceleration, \( F(1, 47) = 232.9, p < .001 \), but not slope, \( F(1, 41) = 0.4, ns \). In addition, both the intercept and the quadratic terms were significantly different from zero \( (p < .001) \). Parameter estimates indicated a mean intercept of 8.6 words and a mean curvilinear rate of change of 2.2. These estimates indicate that the average child began with about 9 words and increased the rate of vocabulary growth two-fold over each testing period. The results from this Level 1 model indicated substantial variability in initial status and nonlinear growth rates. Thus, in the second stage of this analysis, we examined whether the individual differences observed in children’s trajectories of vocabulary growth were related to individual differences in their speed and accuracy of online language comprehension (Level 2).

The first Level 2 variable examined was the speed of word recognition at 25 months. We found that mean RT at 25 months was significantly related to initial vocabulary size, \( F(1, 44) = 4.4, p < .05 \), and also to degree of acceleration in vocabulary growth between the ages of 12 and 25 months, \( F(1, 47) = 10.9, p < .002 \). That is, speed of spoken word recognition at 25 months was predictive of the number of words produced at 12 months of age; moreover, mean RT scores at 25 months were predictive of more rapid acceleration in vocabulary growth trajectories across the second year. Figure 4 depicts the relation between processing speed at two years of age and vocabulary growth from 12 to 25 months. Children were classified as either Fast (≤ 750 ms) or Slow (> 750 ms) in response speed by dividing the sample into two groups based on a median split of the mean RT scores at 25 months. As one can see, those children with faster reaction times at the age of 25 months had had substantially faster rates of curvilinear growth across the second year, as compared to children with slower reaction times.

A similar pattern was observed for accuracy in spoken word recognition at 25 months. The analyses of this Level 2 model indicated that accuracy was related to initial vocabulary size, \( F(1, 43) = 5.3, p < .03 \), as well as acceleration in vocabulary growth, \( F(1, 47) = 10.8, p < .002 \). Parallel to the previous analysis, children were classified as either High (≥ 92% correct) or Low (< 92%) in accuracy based on a median split of the mean accuracy scores at 25 months. Figure 5 shows that those children who were more accurate in spoken word recognition at 25 months had accelerated significantly more quickly in their trajectories of lexical growth across the second year, as compared those children with lower accuracy scores at 25 months. Visual inspection of Figure 5 suggests vocabulary growth is related less strongly to accuracy than to response speed. However, this apparent difference is most likely a consequence of the ceiling effects in accuracy that were reported above.

In general then, these hierarchical growth curve models can be interpreted as suggesting strong relations between the course of vocabulary growth and measures of online speech processing efficiency. The Level 1 models indicated that individual growth trajectories are best characterized by functions that include a nonlinear (quadratic) component. Level 2 models indicated that individual differences in these trajectories were significantly related to individual differences in the speed and accuracy of spoken language comprehension. Thus, we found that those children who were faster and more accurate in spoken word recognition at 25 months were
Efficiency in understanding

the same children who had demonstrated more rapid changes in vocabulary growth across the second year. In addition, a comparison of the model fits (-2 RLL in “smaller-is-better” form) indicated a better overall fit for those models that incorporated the Level 2 factors of RT (-2 RLL = 2465) and accuracy (-2 RLL = 2466), as compared to the unconditional Level 1 model (-2 RLL = 2482) that did not include these additional factors. This suggests that a description of the individual variation in trajectories of vocabulary growth for these children was improved when indices of online performance were taken into account.

Discussion

This experimental study of English-learning children from 12 to 25 months of age provides the first longitudinal data on the emergence of efficiency in spoken language understanding over the second year of life, relating developmental changes in speech processing abilities to vocabulary development over the same period. Each child was observed four times in the looking-while-listening procedure, at 15, 18, 21, and 25 months, to obtain online measures of efficiency in recognizing familiar words in fluent speech. Children’s speed and accuracy in spoken word recognition increased significantly over this period. To explore the relation of these newly developed online measures of speech processing skill in young language learners to more traditional measures of linguistic development, we gathered parental reports of vocabulary growth and grammatical usage in the same children at 12, 15, 18, 21, and 25 months. We found that speed and accuracy in word recognition at 25 months of age were robustly correlated with an impressive range of measures of lexical and grammatical development from 12 to 25 months. The most important finding of this study was that the individual differences observed in children’s trajectories of vocabulary growth were related to individual differences in their speech processing abilities. An analysis of growth curves revealed that those children who were faster and more accurate in spoken word recognition at 25 months had accelerated relatively more rapidly in vocabulary growth across the second year.

Developmental Changes in the Efficiency of Spoken Language Processing

One goal of this research was to replicate previous cross-sectional results showing that efficiency in spoken language processing increases steadily over the second year of life. Observing separate groups of children at 15, 18, and 24 months, Fernald et al. (1998) found age-related changes in the reliability and speed of word recognition over the 9-month period. The longitudinal findings presented here were consistent with this pattern overall, documenting a significant increase in accuracy and decrease in reaction time over the four time points sampled. At the 15-month time point, the mean reaction time for infants in the present study (M = 994 ms) was essentially identical to the mean RT for 15-month-olds in the Fernald et al. study (M = 995 ms), an exact replication that was gratifying but not surprising given that the same speech stimuli were used in both. For 18-month-olds, however, there were differences between the two sets of results. In the earlier cross-sectional study the mean RT at 18 months (M = 827 ms) was significantly lower than at 15 months, whereas in our longitudinal sample the mean RT at 18 months (M = 993 ms) remained constant over this 3-month period. While in both studies the increase in response speed across the second year was highly significant, the overall decrease in reaction time was somewhat less in the present longitudinal study (M = 223 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 were quite different 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. Previous research had shown that 18-month-olds are able to recognize such partial words (Fernald et al., 2001); however, because success on partial word trials at this age is related to level of lexical development, this stimulus set may have been especially difficult for many of the infants at this age. An additional challenging feature of this stimulus set is that we introduced more filler trials than were included in the Fernald et al. (2001) study of partial word recognition, in an effort to reduce the redundancy of the frequently repeated stimuli. This tactic of increasing stimulus variability to maintain infants’ interest unfortunately had the opposite effect, with the result that infants completed on average only 55% of the trials at the 18-month test session, far fewer trials than they completed on average at 15 months (80%), 21 months (87%), or 25 months (94%). 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 word stimuli before. Despite the somewhat anomalous results at the 18-month time point, however, 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 the Fernald et al. (1998) cross-sectional study.

**Using Online Measures of Speech Processing to Study Individual Differences in Young Language Learners**

A second goal of this research was to investigate the stability of our measures of speed and accuracy in online speech processing. This question has both theoretical and methodological implications for our understanding of the competencies involved in the early development of efficiency in understanding. On the theoretical side, we want to be able to determine whether speed of processing or any other component of speech processing efficiency is a stable characteristic of infants from the earliest stages of lexical learning, and to what extent variation in processing capacities might predict variation in later language development. On the methodological side, exploring this question can contribute to the refinement and validation of newly developed online measures of spoken language understanding by young children, and also help us determine the effectiveness of these measures in assessing individual differences in the early development of speech processing abilities. Although eye-tracking techniques are now used extensively in psycholinguistic research with adults (Tanenhaus et al., 2000) and older children (Trueswell, Sekerina, Hill, & Logrip, 1999), they are not yet widely used in research with infants. Measures of proportion of total looking time to one visual stimulus over another are certainly common in infancy research, including studies using measures of auditory/visual matching (Halberda, 2003; Hollich, Hirsh-Pasek, & Golinkoff, 2000; Schafer & Plunkett, 1998). However, these looking-time measures do not rely on frame-by-frame analysis of eye movements, and thus they do not capture the time course of spoken language understanding with precision. Only a few studies to date have provided high-resolution measurements of infants’ response times in spoken word recognition of the sort used in the present research (Fernald et al., 1998; Swingley & Aslin, 2000; Zangl et al., in press). Since all previous developmental studies using such online measures have either focused on a single age group or used a cross-sectional design, nothing is known about the stability of these measures.
Our analysis of the month-to-month correlations for accuracy and RT revealed modest stability on both measures, although the patterns of results were somewhat inconsistent. Children’s mean accuracy scores were significantly correlated from the 18- to 21 and the 21- to 25-month time points, while mean RT scores were significantly correlated from the 15- to 18- and the 18- to 21-month time points, but not from 21 to 25 months. One likely explanation for the inconsistencies in these results is that any underlying stability was obscured by measurement error. As described above, unanticipated problems with the stimulus set used at 18 months resulted in a high rate of inattentiveness during the test session, which meant that the data from that session did not represent children’s best performance, consistent with the finding that speed and accuracy were both lower than expected at 18 months. However, it is important to note that RT at 15 months was strongly correlated with RT at 25 months. Since the stimulus set used at 15 months was more consistent and age-appropriate than that used at 18 months, the mean RT measure was arguably more representative at 15 than at 18 months.

Although measurement error may have compromised to some extent our assessment of stability across the entire age range in this study, the finding that response speed at 15 months predicted response speed at 25 months is striking, given that RT at 25 months was correlated with a host of measures of lexical and grammatical development. These findings indicate that online measures of speech processing efficiency could provide a useful tool for studying individual differences in the development of spoken language understanding, possibly with clinical applications. Of course, it is also possible 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 also depends on attentional factors. If variability in attentiveness is greater in younger infants than in 2-year-olds, then this could contribute to early lack of stability in speech processing performance. Further longitudinal research with carefully designed stimuli and multiple test sessions at each age is the only way to resolve these measurement issues. Such research could also contribute to the development of a potentially valuable diagnostic procedure for assessing receptive language skills in very young children.

Relations Between Efficiency in Spoken Language Comprehension and Growth in Vocabulary Across the Second year

The third and most important goal of this research was to evaluate relations between online measures speech processing efficiency and traditional offline measures of language development. Although some previous studies using online measures have found that infants more advanced in lexical development are also faster and more accurate in spoken word recognition (Fernald et al., 2001; Zangl et al., in press), others have not (Swingley & Aslin, 2000; Werker et al., 2002). The results of the present study clearly indicate that both speed and accuracy in the looking-while-listening procedure were correlated with a range of offline measures of lexical and grammatical abilities. Analyses of concurrent relations among these measures revealed that online speech processing skills were most strongly linked to offline indices of linguistic abilities at 25 months of age, with weaker links at earlier time points. This result is consistent with cross-sectional research using a similar online task that found the relation between speech processing efficiency and productive vocabulary development to be
more robust later in the second year, as compared to earlier in the second year when vocabulary growth is just getting started (Zangl et al., in press).

However, the longitudinal design of the current study allowed a broader examination of these links as children progressed from the beginning stages of word learning well into the period of rapid growth in expressive vocabulary and early use of grammar. It is interesting to note that children’s response speed and accuracy at 25 months were not only related to their concurrent vocabulary size but also to almost all prior measures of vocabulary and grammar starting at the age of 12 months. That is, those children who were faster and more accurate in spoken word recognition at two years of age were likely to be the same children who had shown relatively more sophisticated language skills across the entire one-year period studied. This finding suggests that the relations between online and offline measures are not restricted to a particular developmental period. Further multivariate analyses ruled out the possibility that these relations could be attributed to general age-related changes in processing skill, since reaction time and accuracy were significantly predicted by vocabulary level even when age was taken into account.

The results from the growth curve analysis further underscore the continuity in language abilities that is apparent across the second year. These analyses revealed that speech processing skills are related to characteristics of individual trajectories of growth in productive vocabulary. Consistent with previous research (Bates & Goodman, 1997), the children in this study varied significantly in both initial vocabulary size and the degree to which vocabulary growth was characterized by non-linear increases. Both of these individual difference variables (i.e. intercept and acceleration) were significantly predicted by response speed and accuracy in spoken word recognition. That is, those children who responded more quickly and reliably in the looking-while-listening task at 25 months were those who had produced more words at 12 months of age and who had demonstrated faster rates of productive language growth across the second year. Thus it was not just that those children with more developed speech processing abilities also had more fully developed productive language skills at any given time. Rather, greater efficiency in spoken language comprehension was related more broadly to the overall pattern of children’s developmental trajectory in language learning.

**Why is Faster Processing Speed Associated with More Rapid Lexical Learning?**

Although these findings clearly indicate that children who are more efficient in speech processing are also able to add new words to their productive vocabularies more rapidly, the nature and direction of this relation is far from clear. Does having a larger vocabulary enable children to recognize familiar words more quickly? Or conversely, does the ability to identify familiar spoken words more efficiently somehow facilitate vocabulary growth? Although this chicken-or-egg version of the question is oversimplified, we will first consider each of these possibilities on its own and then suggest ways in which processing speed may interact with vocabulary size in ways that could influence the rate of lexical and grammatical learning.

One approach to understanding the relation of speech processing skill to lexical learning starts with the assumption that differences in the size of children’s productive vocabularies have nothing to do with processing capabilities but rather depend primarily on the amount and quality of the input. Research has shown that the sheer amount of talk directed to preschool children
varies substantially across families, and that these differences in early language input are
associated with differences in various measures of the child’s later language development (Hart
& Risley, 1995; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). Thus children may
develop larger vocabularies because they hear familiar words much more frequently and also
have more opportunities to learn new words. Since children with richer input have more
experience overall in hearing and making sense of speech, they may develop greater facility in
spoken language understanding for the same reason. In this case faster response speed in
recognizing words could be seen as a kind of practice effect. A different argument for the role
of vocabulary size in driving the development of speech processing skills is that young children
who have larger vocabularies may require more refined and efficient word recognition skills in
order to distinguish among greater numbers of potentially confusable representations in the
mental lexicon. Charles-Luce & Luce (1990) and Walley (1993) have made this argument,
suggesting that vocabulary growth itself leads to the development of more efficient processing
strategies and new forms of phonological organization.

An alternative approach to understanding how skill in spoken word recognition relates to
vocabulary development draws on research on the role of processing speed in cognitive
functioning. A substantial literature on processing speed in children and adults has shown that
reaction time decreases with age across a range of cognitive tasks. In a meta-analysis of results
from 72 studies, Kail (1991) examined mean response times on tasks including mental rotation,
letter matching, the Stroop task, reading, and visual search. While some of these tasks assessed
linguistic competence, others were clearly non-linguistic. Kail concluded that age differences in
processing speed reflect some general (i.e. non-task-specific) component that matures rapidly
during childhood. In addition to these age-related changes in response speed, for both children
and adults there is substantial variability in mean reaction time at any given age. These well-
documented individual differences in response speed predict differences in performance on
numerous cognitive tasks (Kail & Salthouse, 1994). Adult processing speed is correlated with
fluid intelligence, short-term memory, analogical reasoning, performance on Raven’s matrices,
memory, and language performance (Kail, 1992; Kail & Salthouse, 1994; Kwong See & Ryan,
1995; Salthouse, 1991, 1996). Indeed, these correlations are so robust that many researchers
believe that differences in processing speed account fundamentally for age-related differences in
cognition (Birren, 1965; Salthouse, 1996). Kail and Salthouse argue that faster processing speed
even enhances unspeeded reasoning and memory, domains that are sometimes seen as
independent from processing speed.

This literature on the development in processing speed suggests that individual differences in
mean reaction time in the age range we studied could be associated with cognitive skills that are
not specific to language learning. In the looking-while-listening procedure used here, the correct
responses that reveal linguistic knowledge presuppose that the child is also proficient in several
other cognitive processes that are not exclusively linked to language. On each trial the child
must first encode the visual image, parse the sentence correctly to identify the target word, 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, however, the child must reject the fixated 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 to the task, rapid encoding of the visual images,
association of an appropriate picture with the target word, integration of visual and auditory input, mobilization of oculomotor responses, and the ability to disengage quickly from one picture in order to attend to another. These are only a few of the perceptual, motor, and cognitive processes that could influence children’s performance in this word recognition procedure, and all entail abilities that are not specific to linguistic processing. Thus one explanation for the relation we found between efficiency in online word recognition and vocabulary growth over the second year is that faster processing speed is associated with greater competence in a wide range of abilities that may facilitate but are not limited to language learning. If this were the case, we might expect to find that speed of visual orienting in a non-linguistic task is related to some degree to online and offline measures of speech processing efficiency, although no such relation was observed in the current study.

One can also imagine a more direct link between rapid online responses to speech and acceleration in vocabulary growth. Salthouse (1996) and others have argued that a major benefit of faster processing speed is that it can free up additional cognitive resources, a factor that may be of special importance in the early stages of language learning. Even for adults who are fluent in a language, efficiency in comprehension requires incremental processing as the speech signal unfolds, rather than waiting until the end of a word or sentence to interpret what has been said (Marslen-Wilson, 1987; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). For infants the challenge of understanding fluently spoken sentences is immeasurably greater. Not only do they need to attend to each word as it comes along, but they must also develop the ability to remember and relate non-adjacent words in the sequence in order to appreciate the long-range dependencies in spoken sentences that are crucial for mastering syntax. To the extent that the child can identify known spoken words as quickly as possible, more resources will be available for attending to subsequent words in the sequence. This could lead to greater effectiveness in learning unfamiliar words encountered later in the sentence as well as tracking distributional information about structural relationships among the words. For these reasons children with faster response speed might advance more quickly in both lexical and grammatical development.

Each of these possible mechanisms for explaining the relations we found between online speech processing efficiency and lexical development can be described as if it functioned in isolation. However, while variations in early language environment and individual differences in perceptual and cognitive abilities related to success in speech processing may indeed contribute independently to early language development, it seems likely that these factors also interact synergistically in ways that affect the rate of learning. For example, child-directed speech functions as “language input” only to the extent that the child can at some level actually process what is heard. Although the word *dog* may be spoken with equal frequency in two environments, the child who is a little faster and more efficient in accessing the meaning of *dog* from the acoustic signal will be able to identify that word more reliably whenever it is spoken, as compared to a child less efficient in speech processing who will fail to identify the word on some occasions. For this reason faster recognition of familiar words may enable a child to build up more robust lexical representations in less time because these words are processed as meaningful lexical items more reliably, and thus, in effect, relatively more frequently. Whether or not particular words are actually spoken more frequently in the environment of the child, the ability to recognize them more rapidly and reliably might convey some of the processing advantages of “frequency effects” widely observed in research on adult comprehension (Monsell, 1991).
Efficiency in understanding

As a result of such positive-feedback processes, there could be cascading advantages for a child with a larger number of more robust lexical representations that would contribute to subsequent success in learning new words. It seems likely that rapid and reliable access to the meanings of familiar words in continuous speech would increase the capacity of the child to notice and remember unfamiliar words that follow. This would be advantageous because attentional resources are critical for learning new words. Moreover, by the end of the second year children increasingly make use of contextual information from known words to infer the meanings of unknown words (Goodman, McDonough, & Brown, 1998). Thus a slight initial advantage in speed and reliability of spoken word recognition could be enhanced through such positive-feedback processes, eventually leading to acceleration in vocabulary growth as well as further increases in the efficiency of spoken language understanding.

It is intriguing to note that a common property of systems with positive-feedback loops is the kind of rapid nonlinear change that is so robustly observed in vocabulary growth. Our findings show that speed of spoken word recognition was more highly correlated with the quadratic term than the linear term in the trajectories of vocabulary growth for individual infants. What this means is that at any given age, children’s mean reaction time tells us not how fast they are able to acquire words at that particular age, but rather how rapidly they are increasing their ability to learn new words. This accords well with the suggestion by Elman et al. (1996) that the nonlinearity in vocabulary growth indicates that "the more words you know, the easier it is to accumulate more." A positive-feedback loop between efficiency in processing familiar words and the ability to learn new words is one mechanism that can explain this insight.

It is clearly impossible to determine whether some children in our study were more efficient in spoken word recognition because they were developing productive language at a faster rate, or conversely, whether they were developing language at a faster rate because they were more proficient at the component skills involved in understanding spoken language in real time. What the data do suggest, however, is that the skills involved in both of these domains are closely intertwined and intricately related throughout the course of acquisition. These results also indicate that the relations between receptive and productive measures demonstrated here should not be characterized as a ‘snapshot’ of a child’s linguistic accomplishments at a particular point in development, e.g. when children’s vocabularies are undergoing the most rapid growth. A more reasonable assumption is that the relations we found between online and offline measures indicate that these different aspects of language competence share important representations, mechanisms, and cognitive resources that work together as the child builds up a system of working language knowledge.
References


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), 1-30.


Efficiency in understanding

Footnotes

1 The version of the Peabody Picture Vocabulary Test used here (PPVT – Revised, Dunn & Dunn, 1981) was normed for children from 3 to 9 years of age, and thus not suitable for purposes of clinical assessment with younger children. We used the PPVT with 25-month-olds only as an additional vocabulary test to complement our other online and offline measures of language competence, with no interest in evaluating the performance of individual children with reference to PPVT norms.

2 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 engaging and challenging to children at each age. For these reasons we did not conduct separate analyses of each sub-type. Thus reported means for speed and accuracy were calculated over responses to all the stimuli in the stimulus set (not including fillers) at each age.

3 A methodological insight to be gained from this 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 to recognize the same target word in the same frame when other the other stimuli in the sequence are highly variable, and susceptibility to such interference varies with age.
Table 1. Month-to-month correlations for three online measures

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Reaction Time (RT)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Accuracy&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Visual Reaction Time (VRT)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-to-18</td>
<td>.35*</td>
<td>.24</td>
<td>.56**</td>
</tr>
<tr>
<td>18-to-21</td>
<td>.36*</td>
<td>.48**</td>
<td>.71**</td>
</tr>
<tr>
<td>21-to-25</td>
<td>.21</td>
<td>.41**</td>
<td>na</td>
</tr>
</tbody>
</table>

Note:  
* p < .05; ** p < .01; na = not available  
<sup>a</sup> Mean response latency (msec) to shift to the target picture on distracter-initial trials  
<sup>b</sup> Mean proportion of correct shifts to the target picture on distracter-initial trials  
<sup>c</sup> Mean response latency (msec) of visual orienting to a silent peripheral image
Table 2. Descriptive statistics (M and SD) for the offline vocabulary and grammar measures from the CDI and PPVT-R

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Words understood(^a)</th>
<th>Words produced(^b)</th>
<th>Grammatical complexity(^c)</th>
<th>Mean 3-longest(^d)</th>
<th>PPVT-R(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>82.5 (62.3)</td>
<td>9.6 (10.8)</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>15</td>
<td>166.2 (88.7)</td>
<td>28.9 (30.7)</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>18</td>
<td>na</td>
<td>96.6 (101.4)</td>
<td>1.1 (4.4)</td>
<td>1.8 (1.0)</td>
<td>na</td>
</tr>
<tr>
<td>21</td>
<td>na</td>
<td>208.5 (158.9)</td>
<td>5.2 (7.4)</td>
<td>3.3 (3.0)</td>
<td>na</td>
</tr>
<tr>
<td>25</td>
<td>na</td>
<td>391.7 (176.8)</td>
<td>13.5 (11.5)</td>
<td>5.3 (3.2)</td>
<td>13.7 (5.4)</td>
</tr>
</tbody>
</table>

Note: * p < .05; ** p < .01; na = not available

\(^a\) Number of words reported as “understands” on the CDI: Words & Gestures

\(^b\) Number of words reported as “understands and says” on the CDI: Words & Gestures (12 and 15 months) or CDI: Words & Sentences (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

\(^d\) Mean length of the 3 longest utterances (M3L) reported on the CDI: Words & Sentences.

\(^e\) Standard score from the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981)
Table 3. Concurrent correlations between reaction time\(^a\) and off-line vocabulary and grammar measures

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Words understood(^b)</th>
<th>Words produced(^c)</th>
<th>Grammatical complexity(^d)</th>
<th>Mean 3 longest(^e)</th>
<th>PPVT-R(^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>-.22</td>
<td>-.14</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>18</td>
<td>na</td>
<td>-.26</td>
<td>-.26</td>
<td>-.23</td>
<td>na</td>
</tr>
<tr>
<td>21</td>
<td>na</td>
<td>-.17</td>
<td>-.13</td>
<td>-.17</td>
<td>na</td>
</tr>
<tr>
<td>25</td>
<td>na</td>
<td>-.38**</td>
<td>-.40**</td>
<td>-.36*</td>
<td>-.60**</td>
</tr>
</tbody>
</table>

Note: *\(p < .05\); **\(p < .01\); na = not available

- \(^a\) Mean response latency (msec) to shift to the target picture on distracter-initial trials
- \(^b\) Number of words reported as “understands” on the CDI: Words & Gestures
- \(^c\) Number of words reported as “understands and says” on the CDI: Words & Gestures (12 and 15 months) or CDI: Words & Sentences (18, 21, and 25 months)
- \(^d\) Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words & Sentences
- \(^e\) Mean length of the 3 longest utterances (M3L) reported on the CDI: Words & Sentences
- \(^f\) Standard score from the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981)
Table 4. Concurrent correlations between accuracy\(^a\) and off-line vocabulary and grammar measures

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Words understood(^b)</th>
<th>Words produced(^c)</th>
<th>Grammatical complexity(^d)</th>
<th>Mean 3-longest(^e)</th>
<th>PPVT-R(^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>.26</td>
<td>.13</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>18</td>
<td>na</td>
<td>.04</td>
<td>-.17</td>
<td>-.02</td>
<td>na</td>
</tr>
<tr>
<td>21</td>
<td>na</td>
<td>.50**</td>
<td>.32*</td>
<td>.41**</td>
<td>na</td>
</tr>
<tr>
<td>25</td>
<td>na</td>
<td>.49**</td>
<td>.44**</td>
<td>.35*</td>
<td>-.60**</td>
</tr>
</tbody>
</table>

Note:  
\(^a\) Mean proportion of correct shifts to the target picture after the onset of the target word  
\(^b\) Number of words reported as “understands” on the CDI: Words & Gestures  
\(^c\) Number of words reported as “understands and says” on the CDI: Words & Gestures (12 and 15 months) or CDI: Words & Sentences (18, 21, and 25 months)  
\(^d\) Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words & Sentences  
\(^e\) Mean length of the 3 longest utterances (M3L) reported on the CDI: Words & Sentences.  
\(^f\) Standard score from the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981)
Table 5. Correlations between reaction time\textsuperscript{a} at 25 months and off-line vocabulary and grammar measures at 12, 15, 18 and 21 months

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Words understood\textsuperscript{b}</th>
<th>Words produced\textsuperscript{c}</th>
<th>Grammatical complexity\textsuperscript{d}</th>
<th>Mean 3 longest\textsuperscript{e}</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>-.45**</td>
<td>-.39**</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>15</td>
<td>-.36*</td>
<td>-.36*</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>18</td>
<td>na</td>
<td>-.37**</td>
<td>-.13</td>
<td>-.23</td>
</tr>
<tr>
<td>21</td>
<td>na</td>
<td>-.45**</td>
<td>-.35*</td>
<td>-.42**</td>
</tr>
</tbody>
</table>

Note: * p < .05; ** p < .01; na = not available

\textsuperscript{a} Mean response latency (msec) to shift to the target picture on distracter-initial trials

\textsuperscript{b} Number of words reported as “understands” on the CDI: Words & Gestures

\textsuperscript{c} Number of words reported as “understands and says” on the CDI: Words & Gestures (12 and 15 months) or CDI: Words & Sentences (18, 21, and 25 months)

\textsuperscript{d} Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words & Sentences

\textsuperscript{e} Mean length of the 3 longest utterances (M3L) reported on the CDI: Words & Sentences.
### Table 6. Correlations between accuracy\(^a\) at 25 months and off-line vocabulary and grammar measures at 12, 15, 18 and 21 months

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Words understood(^b)</th>
<th>Words produced(^c)</th>
<th>Grammatical complexity(^d)</th>
<th>Mean 3 longest(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>.30(^*)</td>
<td>.34(^*)</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>15</td>
<td>.30(^*)</td>
<td>.33(^*)</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>18</td>
<td>na</td>
<td>.39(^**)</td>
<td>.17</td>
<td>.30(^*)</td>
</tr>
<tr>
<td>21</td>
<td>na</td>
<td>.42(^**)</td>
<td>.32(^*)</td>
<td>.39(^**)</td>
</tr>
</tbody>
</table>

Note:  
\(^a\) Mean proportion of correct shifts to the target picture on distracter-initial trials  
\(^b\) Number of words reported as “understands” on the CDI: Words & Gestures  
\(^c\) Number of words reported as “understands and says” on the CDI: Words & Gestures (12 and 15 months) or CDI: Words & Sentences (18, 21, and 25 months)  
\(^d\) Number of times the parent chose the second (more complex) example on the complexity section of the CDI: Words & Sentences  
\(^e\) Mean length of the 3 longest utterances (M3L) reported on the CDI: Words & Sentences.
Table 7. Percentage of children at each age in each vocabulary grouping\(^a\)

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>0-99 words</th>
<th>100-299 words</th>
<th>300-499 words</th>
<th>≥500 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>100.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>58.7</td>
<td>39.1</td>
<td>2.2</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>34.9</td>
<td>30.2</td>
<td>32.6</td>
<td>2.3</td>
</tr>
<tr>
<td>25</td>
<td>6.1</td>
<td>24.5</td>
<td>32.7</td>
<td>36.7</td>
</tr>
</tbody>
</table>

Note: \(^a\) Based on the number of words reported as “understands and says” on the CDI: Words & Gestures (15 months) or CDI: Words & Sentences (18, 21, and 25 months)
Figure 1a. Mean proportion (and standard error) of correct shifts to target picture on distracter-initial trials at each age.
Figure 1b. Mean proportion (and standard error) of incorrect shifts to distracter picture on target-initial trials at each age.
Figure 2. Mean reaction time (RT) by age and vocabulary level.

Note: Error bars represent standard errors. RT = Mean response latency (msec) to shift to the target picture on distracter-initial trials from the onset of the target word

Based on the number of words reported as “understands and says” on the CDI: Words & Gestures (15 months) or CDI: Words & Sentences (18, 21, and 25 months). All vocabulary groupings are not represented at 15 and 18 months (see Table 7).
Figure 3. Mean accuracy by age and vocabulary level.a

Note: Error bars represent standard errors. Accuracy = Mean proportion of correct shifts to target picture on distracter-initial trials.

a Based on the number of words reported as “understands and says” on the CDI: Words & Gestures (15 months) or CDI: Words & Sentences (18, 21, and 25 months). All vocabulary groupings are not represented at 15 and 18 months (see Table 7).
Figure 4. Mean trajectories of growth in vocabulary production as a function of reaction time at 25 months$^a$

Note: Error bars represent standard errors.

$^a$ Based on median split of mean response latencies (msec) at 25 months
Figure 5. Mean trajectories of growth in vocabulary as a function of accuracy at 25 months

Note: Error bars represent standard errors

a Based on median split of mean proportion of correct shifts to target picture at 25 months