Research Report

RAPID GAINS IN SPEED OF VERBAL PROCESSING BY INFANTS IN THE 2ND YEAR

Anne Fernald, John P. Pinto, Daniel Swingley, Amy Weinberg, and Gerald W. McRoberts
Stanford University

Abstract—*Infants improve substantially in language ability during their 2nd year. Research on the early development of speech production shows that vocabulary begins to expand rapidly around the age of 18 months. During this period, infants also make impressive gains in understanding spoken language. We examined the time course of word recognition in infants ages 15 to 24 months, tracking their eye movements as they looked at pictures in response to familiar spoken words. The speed and efficiency of verbal processing increased dramatically over the 2nd year. Although 15-month-old infants did not orient to the correct picture until after the target word was spoken, 24-month-olds were significantly faster, shifting their gaze to the correct picture before the end of the spoken word. By 2 years of age, children are progressing toward the highly efficient performance of adults, making decisions about words based on incomplete acoustic information.*

The language abilities of human infants improve dramatically during the 2nd year of life (e.g., Bates, Bretherton, & Snyder, 1988). Most 12-month-old infants are just beginning to speak, and acquire new words at a slow pace. At around 18 months of age, many infants show a "vocabulary burst,'' shifting to a much faster rate of acquisition (Bloom, 1973; Goldfield & Reznick, 1990), and by age 24 months, infants can typically produce 200 to 500 words (Fenson et al., 1994). Research on the early development of linguistic competence has focused much more on what infants can say than on what they can understand, in part because speech production is directly observable whereas comprehension is not. However, during this period of rapid expansion in productive vocabulary, infants also make impressive gains in their receptive language abilities. Here we provide the first fine-grained analysis of the time course of spoken-word understanding by infants ages 15 to 24 months. By monitoring the ongoing process of word recognition by infants in a behaviorally relevant context, we found that both the speed and the accuracy of speech processing increase steadily over the 2nd year.

Comprehension is a mental event that must be inferred from overt behavior and the context in which it occurs. One common method for assessing the early development of comprehension skills is to ask parents what words they think their child understands (Fenson et al., 1994). Other laboratory methods test comprehension by asking infants to choose a named object from among alternatives (e.g., Woodward, Markman, & Fitzsimmons, 1994). Such parental-report and object-choice measures allow inferences about which words an infant understands, but not about the actual process of word recognition. In adults, much more sophisticated methods are available for studying temporal aspects of spoken-language understanding. Research using various on-line measures of speech comprehension has shown that word recognition by adults is extremely fast and efficient, occurring as soon as there is sufficient acoustic information to disambiguate alternative word candidates (e.g., Altmann, 1990; Grosjean, 1980; Warren & Marslen-Wilson, 1987). However, it has not been possible to examine the speed and efficiency of speech comprehension in infants using such on-line behavioral techniques, because the task demands are too difficult and because infants cannot be instructed to respond quickly and accurately.

Studying the process of word recognition in infants requires a technique that has minimal task demands and that can be used in a naturalistic context. Because infants tend to look at a familiar object when it is named, longer looking at a named target object than at an unnamed distractor object has been used as a measure of word recognition (e.g., Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Reznick, 1990). In previous research using a preferential looking procedure, we found that 19-month-old infants oriented more reliably to pictures of familiar named objects than did 12- or 15-month-old infants (Fernald, McRoberts, & Herrera, 1992). We also found an age-related increase in the flexibility of speech processing: Although 15-month-old infants could recognize a familiar word in sentence-final position ("Over there there's a *ball*"), they failed to recognize the same word in a perceptually more challenging context, when it was embedded in the middle of the sentence ("There's a *ball* over there"). However, 19-month-old infants performed well regardless of the position of the word in the sentence. Thus, during the period of rapid expansion in productive vocabulary in the 2nd year, infants also develop greater proficiency in extracting lexical information from continuous speech.

These findings motivated us to monitor in greater detail the actual time course of word recognition by infants during this period. A promising approach used with adults has been to observe their gaze patterns as they look at an array of objects while listening to speech related to the objects. By examining sequences of eye movements time-locked to particular words, it is possible to monitor the rapid mental processes involved in understanding spoken language (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Fine-grained analyses of eye movements have also been used to measure infants' response to visual stimuli, although only in nonlinguistic contexts (cf. Haith, Hazan, & Goodman, 1988; Hood & Atkinson, 1993; Johnson, Posner, & Rothbart, 1994). By increasing the resolution of our coding system, we were able to track infants' eye movements as they looked at pictures while listening to speech naming one of the pictures. To examine the time course of word recognition, we identified the point at which infants initiated a shift in gaze toward the named picture, measuring their eye movements from the beginning of the spoken target word. Our goal was to assess developmental changes in the speed and accuracy of word recognition by infants from 15 to 24 months of age.

METHOD

Subjects

Subjects were 72 infants from monolingual English-speaking families in a predominantly middle-class population. We tested 24 infants in each of three age groups: 15 months (range: 14.5–15.5), 18 months (range: 17.5–18.5), and 24 months (range: 24.0–25.0). An additional 41

Address correspondence to Anne Fernald, Department of Psychology, Stanford University, Stanford, CA 94305; e-mail: fernald@psych.stanford.edu.

infants were tested but not included in the final sample for the following reasons: fussiness during testing (20), failure to look at pictures during one or more of the trials (6), eyes difficult to see during coding (6), parental interference during testing (4), and experimenter error or equipment failure (5). Parents were administered the MacArthur Communicative Development Inventory (see Fenson et al., 1994); only infants whose parents reported that they were familiar with all four target words were included in the analysis. Informed consent was obtained from the parents of all infants participating in this study.

Stimuli

The auditory stimuli consisted of sentences containing four target words, *doggie, baby, ball,* and *shoe,* typically among the first object names acquired by English-learning infants. The four target words were similar in duration (752–794 ms) and were always presented in the same two carrier frames ("Where's the ___? See the ___?"). These sentences were spoken in infant-directed speech style, in which words are typically longer in duration than in adult-directed speech (Fernald et al., 1989). The visual stimuli consisted of four colored pictures of familiar objects, one matching each of the target words. The pictures were presented in pairs (ball-shoe or dog-baby), with side of presentation of the objects in each pair varied across trials, and order of object presentation counterbalanced across subjects.

Procedure

Each subject was observed in a three-sided testing booth located in a darkened, sound-treated room (see Swingley, Pinto, & Fernald, in press, for details). The infant sat on his or her caregiver's lap 72 cm from two adjacent computer monitors, which were positioned at the infant's eye level. During each trial, one object picture was shown on each monitor. As a guard against bias, the caregiver's view of the pictures was blocked by an opaque curtain. After a 4-s silent familiarization period, the infant heard two sentences naming one of the objects. Each of the four objects served as the target on two trials and the distractor on two trials. Eight 6-s test trials were interspersed with three filler trials showing different pictures, to maintain the infant's interest. Visual and auditory presentations were controlled by a Macintosh computer located in an adjacent room.

Infants' patterns of looking to the two pictures were recorded with a video camera concealed between the monitors. A time code accurate to 0.1 s was superimposed on the video record. Eye movements were analyzed off-line frame by frame from the videotapes by highly trained coders who were unaware of the side of the target object on each trial. For every 100 ms, coders assessed whether the child was looking at the target, at the distractor, or at neither object. These measurements of the looking pattern on each trial were then analyzed in relation to the onset and offset of the first spoken target word, located by visual inspection of the acoustic waveform. A subset of the trials was coded independently by two or more coders, to determine interobserver reliability in measurement of eye movements. Kappas ranged from .85 to .99 (M = .92).

RESULTS

A preliminary analysis verified that infants in each age group demonstrated recognition of the target words. Examining the overall proportions of looking time to the target and distractor objects for the 4-s period following the offset of the target word, we found that infants looked significantly more at the target objects than at the distractor objects in response to naming. Performance was well above chance (50%) at all ages: 15 months, 65.4%, $t(23) = 5.5$, $p < .0001$; 18 months, 72.3%, $t(23) = 8.7$, $p < .0001$; and 24 months, 77.2%, $t(23) = 13.4$, $p < .0001$.

In the next analysis, we focused on developmental changes in the speed of word recognition, calculating infants' reaction time to seek out the correct picture in response to the spoken target word. Because the infants could not anticipate the side on which the target object would be presented, they were equally likely to be looking at the distractor or the target object at the beginning of a trial. At the onset of the target word, infants were looking at the distractor picture on 39.8% of the trials, at the target picture on 38.4% of the trials, and at neither picture on 21.8% of the trials, averaged across the three age groups. Using only those trials on which the infant was initially fixating the distractor object, we calculated reaction time by measuring the infant's latency to initiate a shift in gaze from the distractor to the target object when the target word was spoken. Mean response latency decreased significantly with age: 995 ms for 15-month-olds, 827 ms for 18-month-olds, and 679 ms for 24-month-olds, as shown in Figure 1. The overall age effect was reliable, $F(2, 63) = 6.83$, $p < .01$. Comparisons between age groups revealed that 18-month-olds were reliably faster than 15-month-olds, $t(41) = 1.75$, $p < .05$, and 24-month-olds were reliably faster than 18-month-olds, $t(43) = 2.04$, $p < .05$.

The final analysis compared correct shifts in gaze (distractor to target) and incorrect shifts in gaze (target to distractor) in response to the target word. This analysis included all trials in which the infant was on task, excluding those trials in which the infant was looking away from both pictures when the target word was spoken. In the graphs shown in Figure 2, the lines show the proportion of trials on which the infants shifted fixation from the picture they were initially looking at to the

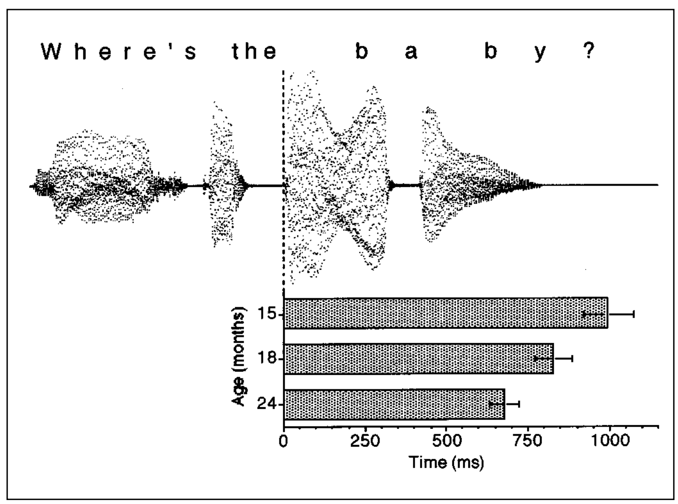

Fig. 1. Mean latencies to initiate a shift in gaze from the distractor picture to the target picture, measured from the beginning of the spoken target word, for 15-, 18-, and 24-month-old infants. This analysis included only those trials on which the infant was initially looking at the incorrect picture and then shifted to the correct picture when the target word was spoken. The graph is aligned with an amplitude waveform of one of the stimulus sentences.

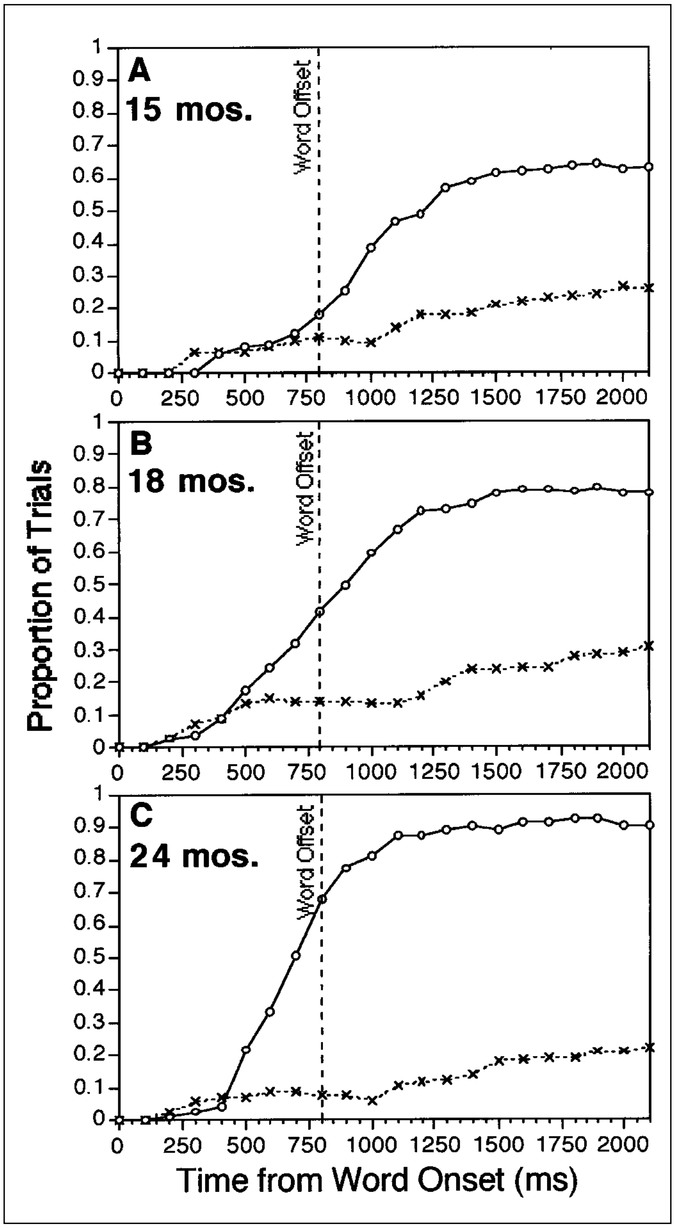

Fig. 2. Shifts in gaze from one picture to the other in response to the target word, for infants ages (a) 15 months, (b) 18 months, and (c) 24 months. Solid lines represent correct responses, shown as the proportion of trials on which infants were initially fixating the distractor object at the onset of the target word and then shifted to the target picture (distractor-to-target shifts); dashed lines represent incorrect responses, shown as the proportion of trials on which infants were initially fixating the target object and then shifted to the distractor picture (target-to-distractor shifts).

other picture, calculated every 100 ms from the beginning of the target word. Solid lines represent the proportion of correct responses, when the infants started out looking at the distractor picture and then shifted to the target, and dashed lines represent the proportion of incorrect responses, when the infants started out looking at the target picture and then shifted to the distractor. As more of the word was heard, correct shifts outnumbered incorrect shifts at all three ages. However, the lines diverge earlier and to a greater extent at 24 months, indicating that older infants were not only faster, but also more accurate in matching familiar spoken words with their referents.

DISCUSSION

These findings reveal that infants, like adults, respond spontaneously to a familiar spoken word by rapidly fixating the visual referent that matches the word they hear. Moreover, during their 2nd year, infants make dramatic gains in speed and efficiency in understanding familiar words in continuous speech. Although 15-month-old infants demonstrated recognition of the four target words by reliably looking at the correct picture in response to naming, they did not initiate a shift in gaze to the correct picture until after the target word had been spoken. In contrast, 24-month-olds were 300 ms faster, on average, shifting their gaze to the correct picture before the end of the spoken word. These results demonstrate that by 2 years of age, children are rapidly progressing toward the highly efficient performance of adults, making decisions about words based on incomplete acoustic information (e.g., Tanenhaus et al., 1995; Warren & Marslen-Wilson, 1987).

Infants' response latency to a spoken word in this task was influenced not only by the time required for linguistic processing, but also by the time required to program a shift in gaze to the correct picture. It is important to note that our reaction time measure was based on the time it took the infant to initiate an eye movement from one picture to the other and did not include the actual shift in gaze. An analysis of the duration of shifts in gaze from one picture to the other revealed no significant age differences in speed of executing eye movements. It might still be the case that the longer response latencies of the younger infants were due to the fact that they required more time to program an eye movement. This explanation seems unlikely, however, because eye movement control in infants is more adultlike than any other motor system (Haith, Wentworth, & Canfield, 1993). In infants as young as 3 months of age, the minimum latency to initiate a shift in gaze to a peripheral picture in a nonlinguistic task is about 150 ms, approaching adult reaction time in such situations (Canfield, Smith, Brezsnyak, & Snow, 1997). Thus, the most plausible explanation for the age differences in infants' response latency to spoken words is that they reflect differences in speed of linguistic processing rather than maturation of the oculomotor system.

Our data show that during the period of rapid expansion in speaking vocabulary at the end of the 2nd year, infants also get much faster at understanding the words they hear. Neuropsychological research on infants' responses to spoken words during the same period reveals a major shift in cerebral organization that appears to be closely linked to language knowledge (Mills, Coffey-Corina, & Neville, 1993). Such developmental changes in neural responses to known words may underlie the increase in speed and efficiency revealed in our on-line behavioral measures of verbal processing. Our results also show that under experimental conditions that are both well controlled and naturalistic, eye movements can be used effectively to monitor the time course of word recognition by infants. This paradigm will enable investigation of a wide range of questions related to the early development of competence in understanding spoken language.

Acknowledgments—We thank Praveen Menon-Johansson, Andrea Heberlein, and the undergraduate research assistants at the Center for Infant Studies for their dedication and skill in testing infants and coding eye movements. We also thank the parents and infants who participated in this study. This work was supported by National Institutes of Health Grant MH41511 (A.F.), National Science Foundation graduate fellowships (J.P.P., D.S.), and an equipment grant from Apple Computer.

REFERENCES

Altmann, G.T.M. (Ed.). (1990). *Cognitive models of speech processing.* Cambridge, MA: MIT Press.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar.* New York: Cambridge University Press.

Bloom, L. (1973). *One word at a time.* The Hague, Netherlands: Mouton.

Canfield, R.L., Smith, E.G., Brezsnyak, M.P., & Snow, K.L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 62*(2, Serial No. 250).

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Fernald, A., McRoberts, G.W., & Herrera, C. (1992, May). *Prosodic features and early word recognition.* Paper presented at the 8th International Conference on Infant Studies, Miami, FL.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*, 477–501.

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

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

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

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

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

Hood, B.M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development, 16*, 405–422.

Johnson, M.H., Posner, M.I., & Rothbart, M.K. (1994). Facilitation of saccades toward a covertly attended location in early infancy. *Psychological Science, 5*, 90–93.

Mills, D.L., Coffey-Corina, S.A., & Neville, H.J. (1993). Language acquisition and cerebral specialization in 20-month-old infants. *Journal of Cognitive Neuroscience, 5*, 317–334.

Reznick, S.J. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics, 11*, 145–166.

Swingley, D., Pinto, J.P., & Fernald, A. (in press). Assessing the speed and accuracy of word recognition in infants. In C. Rovee-Collier, L.P. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol.12). Norwood, NJ: Ablex.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken-word recognition. *Perception & Psychophysics, 41*, 262–275.

Woodward, A.L., Markman, E.M., & Fitzsimmons, C.M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30*, 553–566.

(RECEIVED 7/7/97; ACCEPTED 10/30/97)