

An infant-based assessment of early lexicon acquisition

MARGARET FRIEND and MELANIE KEPLINGER
San Diego State University, San Diego, California

The majority of research on the acquisition of spoken language has focused on language production, due to difficulties in the assessment of comprehension. A primary limitation to comprehension assessment is maintaining the interest and attention of younger infants. We have developed an assessment procedure that addresses the need for an extensive performance-based measure of comprehension in the 2nd year of life. In the interest of developing an engaging approach that takes into account infants' limited attention capabilities, we designed an assessment based on touchscreen technology. This approach builds upon prior research by combining standardization and complexity with an engaging infant-friendly interface. Data suggest that the touchscreen procedure is effective in eliciting and maintaining infant attention and will yield more extensive and reliable estimates of early comprehension than do other procedures. The software to implement the assessment is available free of charge for academic purposes.

The majority of research on the acquisition of spoken language has focused on language production, due to difficulties in the assessment of comprehension. However, the study of language comprehension is a route to the concepts that underlie production. The concurrent study of comprehension and production promises to provide a more complete picture of the language acquisition process than does the study of production alone.

Over the past decade, a standardized measure of parent estimates of infant language comprehension, the MacArthur Communicative Development Inventory: Words and Gestures (CDI:WG; Fenson et al., 1993), was developed and extended to approximately 20 languages (<http://www.sdsu.edu/cdi/foreign.html>). The CDI:WG provides extensive data, but its reliability, validity, and utility for predicting the course of language development remain controversial (Feldman et al., 2000; Yoder, Warren, & Biggar, 1997). As a result, it is desirable to supplement parent report with performance-based measures of comprehension in order to address the concerns raised with regard to parent report and to obtain a more complete picture of child lexical knowledge.

This research was supported by a Blasker grant to the authors from the San Diego Foundation. Gratitude is extended to Larry Fenson and Kelly Diaz for comments on project development, to Tom Kaufeld for software development, to Nahil Berhane for assistance in the acquisition of stimulus materials, to Gaddy Bergmann for allowing us to use her beautiful turtle photograph in our assessment, and to Brian MacWhinney, Jonathan Vaughan, and two anonymous reviewers for their comments on a previous version of this manuscript. The first author has applied for international patent protection of this assessment procedure; however, the software is free for academic use. Correspondence concerning this article should be addressed to M. Friend, San Diego State University, 6363 Alvarado Court, Suite 103, San Diego, CA 92120 (e-mail: mfriender@sunstroke.sdsu.edu).

Two primary laboratory approaches to performance-based assessment of child language comprehension are object/picture selection and preferential looking. Object selection requires a volitional response (selecting, manipulating, or pointing) that makes this approach more appropriate for toddlers than for younger infants (Bates et al., 1988; Oviatt, 1980; Snyder, Bates, & Bretherton, 1981). Preferential-looking procedures facilitate the assessment of comprehension in younger infants by eliminating the volitional response component. In preferential-looking tasks, visual fixation is taken as a dependent measure of the infant's association of a lexical item with an object or action (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996; Naigles & Gelman, 1995; Schafer, 1998).

A primary limitation of these approaches is that, typically, very few items are presented in any single assessment, so that an extensive estimate of early comprehension is not possible. An exception to this limitation is the Comprehension Book, based on the object and picture selection approaches, developed by Ring and Fenson (2000) and extended by Rodrigue (2001). The original Comprehension Book is a set of 42 picture pairs organized in a picture book format. The lexical items represented by the exemplars in the book constitute a subset of items drawn from the CDI:WG and the CDI: Words and Sentences (CDI:WS). The items consist of nouns, verbs, and adjectives, representing a range of difficulty based on lexical development norms (Dale & Fenson, 1996). The Comprehension Book provides an extensive behavioral measure of comprehension; however, the response component (referential pointing) and format make infant compliance difficult to achieve before 20 months of age (Friend & Diaz, 2001; D. Thal, personal communication, July 2002).

The difficulty in maintaining the interest and attention of younger infants may lead to an attenuation of estimates

of vocabulary knowledge and of the relation between child performance and parent report. Friend and Diaz (2001) found that controlling for infant inattention in performance on the Comprehension Book leads to a significant increment in the variance accounted for by parent report. These data highlight the importance of developing measures of early vocabulary acquisition that maximize infant interest and attention.

We have developed a new assessment, based on the work of Ring and Fenson (2000) and Rodrigue (2001), that addresses the need for an extensive performance-based measure of comprehension in the 2nd year of life. Like Ring and Fenson and Rodrigue, we present pairs of high-quality images in a forced-choice format. The lexical targets that the images represent are selected from the CDI:WG and the CDI:WS and consist of nouns, verbs, and adjectives that vary in frequency of occurrence in infants' receptive lexicons at 16 months of age (Dale & Fenson, 1996).

In contrast to Ring and Fenson (2000) and Rodrigue (2001), we have extended the original Comprehension Book in several ways and have introduced an innovative method of presentation appropriate to the assessment of younger infants. First, we randomized word difficulty across trials. In earlier Comprehension Books, a disproportionate number of easy items were among the first items presented. Since children who do not complete the task usually do not participate after the first few items, in earlier versions of this assessment, inattention was confounded with lack of knowledge. Second, we incorporated equal proportions of easy, moderate, and difficult words, rather than biasing the assessment toward early-appearing words. In this way, we have extended the ceiling of the assessment to make it appropriate across a wider range of ages. Third, we modified the images in order to produce better within-pair matches on salience (size, color, and brightness). Fourth, we replaced some images in order to provide more prototypical and contemporary exemplars. Finally, we incorporated new lexical items and images in order to produce better matches on frequency of occurrence. The most significant and unique contribution of our research is the introduction of a program and method of administration that facilitates the assessment of infants younger than 20 months of age.

In the interest of developing an engaging approach to infant language comprehension that takes into account infants' limited attention capabilities, we designed a program based on touchscreen technology. In this application, the appearance of colorful attractive images on the screen engages infant attention. Infants touch images on a 17-in. kiosk-enclosed screen in response to auditory prompts from an experimenter in which target vocabulary items are embedded (e.g., "Where is the *shoe*?" "Touch the *shoe*"). Touching the target image produces a reinforcing auditory signal that maintains interest and motivates task compliance. The reinforcing signal is presented only *after* a touch to the target image and is not presented if the infant or toddler fails to touch the target image. In this

way, these signals serve solely to reinforce touches to the target and are not confounded with lexical comprehension. Anecdotally, children in our laboratory have been observed to dance when the auditory stimuli are presented and to touch the screen multiple times in succession to produce them.

The combination of attractive visual and auditory stimuli activated in response to infant touch is a mainstay of educational toys for infants in the 1st and 2nd years of life. The present approach builds upon prior research (Ring & Fenson, 2000; Rodrigue, 2001) and introduces the familiar and engaging approaches employed in educational toys to the assessment of early language by combining standardization and complexity with an engaging infant-friendly interface.

HARDWARE REQUIREMENTS

The Computerized Comprehension Task (CCT) was developed using Macromedia's Director 7.5 and is designed for PC platforms running Windows 95, 98, and 2000 operating systems. It is not necessary to purchase Director unless the user wishes to modify the program. Alternatively, users may wish to write modifications of the existing program in other experiment generator software, with the permission of the first author and proper citation of the present documentation. Screen resolution must be set to $1,280 \times 1,024$ ppi to accommodate the image files, and the computer must be equipped with a speaker for the presentation of reinforcing auditory stimuli. The program is designed for use with monitors equipped with a touchscreen but can be implemented on standard CRT or LCD monitors, and all program features can be activated with mouse clicks. For example, using a standard monitor, an experimenter can generate the reinforcing auditory stimuli and write data to the output file by clicking a mouse when the infant touches the screen. Pilot data suggest that this procedure maintains infant attention but imposes increased complexity of administration on the experimenter. Since a key feature of the approach is the contingency between infant touches and the auditory reinforcers, a touchscreen is recommended. We have found that both capacitive and resistive touchscreens can be effective in testing young children. Capacitive screens record the disruption of an electrical field by the introduction of a finger or hand and can record very light touches. In contrast, resistive screens record pressure and can be less sensitive to light touches. Surface-acoustic-wave screens record the disruption of an acoustic signal by a finger, hand, or probe. These screens are so sensitive that the presence of dust or dirt on the screen can be sufficient to register as a touch. In our research, capacitive screens provide the best balance between sensitivity and durability. When set to their highest level of sensitivity, these screens record young children's touches reliably and accurately. We have successfully administered the CCT with the KDS Model P3-800 capacitive and the KDS Model 613PMARU resistive touchscreens.

PROGRAM DESIGN

The program presents 41 pairs of vibrantly colored digital images representing nouns, verbs, and adjectives. Two images appear simultaneously at left- and right-center screen when the experimenter clicks an invisible button at the bottom of the screen. An initial touch to the target image produces a unique reinforcing auditory signal and is coded as a “1” in the output file. An initial touch to the distractor image does not produce an auditory reinforcer and is coded as a “0” in the output file. Subsequent touches are not recorded in the output file, but touches to the target continue to produce the auditory reinforcer for the duration of the trial. If there is no touch during a trial, the trial is coded as missing (“.”) in the output file, and no auditory reinforcer is produced. It is possible to repeat a trial and overwrite the output for that trial if an erroneous touch to the screen occurs. Blue screens are interleaved between trials, and the pace of the assessment is controlled by the experimenter, with the exception that each trial has a maximum duration of 7 sec before the screen is refreshed and a blue screen appears. This maximum duration ensures that the experimenter maintains control of the cursor at the end of the trial. Without a fixed trial length, infants become so interested in touching the screen that it can be difficult to return control to the experimenter in order to advance to the next trial.

There is an equal representation of easy (comprehended by more than 66% of 16-month-olds), moderately difficult (comprehended by 33%–66% of 16-month-olds), and difficult (comprehended by less than 33% of 16-month-olds) word pairs. Word difficulty is defined a priori on the basis of normative parent report data from the CDI:WG (Dale & Fenson, 1996). The procedure calls for a forced-choice touch to one member of a pair of photographs in response to an experimenter’s query (e.g., “Where is the *diaper*?” “Touch *diaper*”). There are two forms of the procedure, and the member of each pair that serves as the target is counterbalanced across forms. Within forms, the level of difficulty is matched within pairs and is randomized across stimulus presentations. Target items appear with equal frequency on the right and the left sides of the screen. The side on which the target appears is randomized across presentations, with the restriction that targets appear no more than twice in succession on the same side. This arrangement is consistent with the well-researched design used in preferential-looking paradigms (Hirsh-Pasek, & Golinkoff, 1996) to reduce orientation bias effects. The program produces a log-on screen, three stimulus presentation phases (training, test, and reliability), and a data output file generated during the test and reliability phases.

Log-On Screen

Double-clicking/double-touching the CCT program icon produces a log-on screen for acquiring participant information prior to testing. The log-on screen contains the following data fields that accept keyboard data entry: participant, sex, date of birth, date of testing (automated),

form number (1–2), experimenter initials, and five additional blank fields.

Double-clicking/touching an invisible button at the bottom right corner of the screen prompts the user to enter a path and a filename for the output data file. Once a path and a filename have been specified, a dialogue box is produced with the following options: begin training, begin test, and exit. A single click/touch initiates the training or test phase or exits the program. When either the training or the test button is selected, a blue screen appears.

Double-clicking/touching the invisible button at the bottom left corner of the screen at any time during training or testing produces an “Exit Program?” dialogue box with three buttons: repeat, exit, and cancel. The repeat feature presents the most recent trial again and overwrites the output file for that item. The exit feature codes data for the current trial as missing (“.”), saves the output file under the specified path and filename, and terminates the program. The cancel feature removes the dialogue box and returns the user to a blank blue screen without modifying the output file. The program continues with presentation of the next trial in the sequence.

Visual Stimuli

Visual stimuli for the program include 90 high-quality digital images consisted of prototypical referents for the lexical targets assessed within the program. Each image exists as an individual jpeg format file. All files are 562×750 pixels, with a resolution of 160 ppi. The images are high-quality digital images selected from Art Explosion (1995–1998) and Freefoto (www.freefoto.com) and supplemented with images created in the laboratory. All the images were edited extensively using Adobe Photoshop 5.0 (1999) to maximize salience matching within image pairs. Image backgrounds match the screen color so that there are no lines of demarcation between the images and the blank screen areas. Sensitive screen areas are defined in such a way that a touch (with a finger or hand) to the target image or the edge of the screen adjacent to the target image results in a score of “1,” written to the output file. A touch to any other screen area results in a score of “0,” written to the output file. Because the target image accounts for less than 50% of the total sensitive screen area, random touches have a probability of less than 50% of being coded erroneously as “hits” in the output file. The inclusion of a reliability phase provides a further check on random responding.

Prompts

Verbal prompts for each trial are not automated in the program and must be produced by the experimenter administering the task. We have found it useful to use the following prompts for the word classes represented in the assessment:

“Where is the ___? Touch ___,” for nouns;

“Who is ___? Touch ___,” for verbs; and

“Which one is ___? Touch ___,” for adjectives.

Auditory Stimuli

All the images are associated with an auditory stimulus that is presented only *after* the infant touches the target image. Touches to the distractor or to other screen areas *do not* produce an auditory stimulus. Each stimulus consists of the target lexical item presented in child-directed speech, accompanied by a reinforcing sound. The reinforcing sounds were selected to be engaging and meaningfully related to the target words. For example, a touch to the target “ball” produces the word “ball” in child-directed speech, followed by a *bouncing* sound. In this way, infants are motivated to touch the target image to produce the reinforcing auditory stimulus. The stimuli are sampled at 22 kHz in stereo and are included in the program as .wav files. Sounds were selected from SoundEffects.com’s Flash eFX CD (soundeffects.com/flashefx.html) and were modified for this program and supplemented with sounds created in our laboratory.

Training Phase

In the training phase, infants are familiarized with the assessment procedure, and no data are written to the output file. The training phase can be repeated as necessary. There are four pairs of items in the training phase: diaper and shoe, cup and sock, bathtub and door, and book and car. Each item is comprehended by at least 85% of 16-month-olds, according to Dale and Fenson (1996), and should be easily comprehended by most infants in the 2nd year. After the four pairs of items have been presented, a dialogue box appears with the following options: repeat training and begin test. Single clicking/touching the repeat training button initiates repetition of the four training trials, beginning with a blank blue screen. Single clicking/touching the begin test button initiates the test trials and data collection, beginning with a blank blue screen. From this point on, all touches and missing data are coded in the designated output data file.

Test Phase

In the test phase, 41 pairs of images are presented. In our experience, completion of this phase takes about 10 min. The lexical items assessed, screen orientation, and level of difficulty are presented for each form of the procedure in Table 1. Sample test images are presented in Figure 1. Each image pair is preceded by a blank blue screen. Experimenters use this opportunity to engage the infant’s attention and to present the verbal prompt in order to assess comprehension (e.g., “Where is the *shoe*? Touch *shoe*”). Double-clicking/touching the invisible button at the bottom right of the screen produces the target image and a distractor. Touches to the screen are recorded as “1” if the first touch is to the target and as “0” if the first touch is to the distractor. Subsequent touches are not coded. If there is no touch and the experimenter moves to the next trial, a “.” is written to the output file to identify that trial as missing. At any point, it is possible to exit the program or to present the most recent trial a second time (overwriting the existing output for that trial) by double-clicking/touching the invisible button at the bottom left of the screen. Once

Table 1
Test Phase: Lexical Items, Screen Orientation, and Difficulty Level

Orientation		Target		Difficulty Level
Left	Right	Form 1	Form 2	
dog	bird	dog	bird	E
sliding	running	running	sliding	M
mouth	eye	mouth	eye	E
sheep	lion	sheep	lion	D
orange	green	green	orange	D
kissing	hugging	hugging	kissing	E
pulling	swimming	pulling	swimming	D
telephone	keys	telephone	keys	E
kicking	drawing	drawing	kicking	D
bus	firetruck	bus	firetruck	D
nose	foot	foot	nose	E
happy	sad	happy	sad	D
button	hat	button	hat	M
juice	banana	banana	juice	E
old	new	old	new	D
toothbrush	spoon	toothbrush	spoon	E
drinking	dancing	dancing	drinking	E
swinging	jumping	jumping	swinging	M
horse	cow	horse	cow	M
milk	cookies	cookies	milk	E
table	chair	table	chair	M
little	big	big	little	D
eating	throwing	eating	throwing	E
scissors	money	scissors	money	D
red	blue	blue	red	D
truck	airplane	airplane	truck	M
full	empty	full	empty	D
bicycle	train	train	bicycle	M
penguin	giraffe	penguin	giraffe	D
cheese	apple	apple	cheese	E
smiling	crying	smiling	crying	M
playing	sleeping	playing	sleeping	M
bottle	ball	ball	bottle	E
reading	washing	reading	washing	M
bubbles	doll	bubbles	doll	M
turtle	butterfly	butterfly	turtle	D
touching	riding	touching	riding	M
dirty	clean	clean	dirty	M
cat	duck	duck	cat	E
pig	fish	pig	fish	M
girl	boy	boy	girl	D

Note—For difficulty level, E (easy) = items comprehended by more than 66% of infants, M (moderate) = items comprehended by 33%–66% of infants, and D (difficult) = items comprehended by fewer than 33% of infants at 16 months of age.

a new trial has been presented, it becomes the most recent trial, and the output for the previous trial can no longer be overwritten. When the experimenter double-clicks/touches the button at the bottom right of the screen following the final trial in the sequence, a dialogue box appears with the following options: same and opposite. Single-clicking/touching one of these buttons initiates the reliability phase. Double-clicking/touching the invisible button at the bottom left of the screen enables the user to exit the program without collecting reliability data.

Reliability Phase

Single-clicking/touching the *same* button initiates the presentation of one third of the test items a second time in



Figure 1. Sample exemplar pair for the lexical terms *telephone* and *keys*.

the same relative left–right screen orientation. Single-clicking/touching the *opposite* button initiates the presentation of one third of the test items a second time in the opposite relative left–right screen orientation. These options permit the assessment of test–retest reliability on a subset of items and the evaluation of potential side-bias effects. The reliability items were selected to mirror the relative proportions of easy, moderate, and difficult items and of nouns, verbs, and adjectives in the full test. The reliability items are presented for each form of the procedure in Table 2.

Output File

The program generates a comma-delimited output file that contains all of the participant information entered on the log-on screen and data from the test and reliability phases. The file can be designated as an Excel file by adding the extension “.csv” when specifying the path and filename at the beginning of the assessment. If no extension is specified, it is saved as comma-delimited text and can be imported into a range of programs for analysis.

Scoring

A single comprehension score can be generated from the output file by summing infants’ correct touches across trials, congruous with the procedure for estimating vocabulary comprehension on the CDI:WG (Fenson et al., 1993). More detailed comprehension profiles can be obtained by creating separate summary scores based on word class (nouns, verbs, and adjectives) or difficulty (easy, moderate, and difficult). This can be accomplished by importing the output file into a statistical analysis software program and writing syntax to create these different groupings.

In scoring, it is important to be mindful of infants’ missing and ambiguous responses (e.g., touching both images simultaneously). These responses can be conceived as either comprehension or performance limitations. The

program does not distinguish a touch to the target followed immediately by a touch to the distractor (ambiguous) from a touch to the target alone (correct). We supplement the automated output file with a videotape record

Table 2
Reliability Phase: Lexical Items, Screen Orientation, and Difficulty Level

Orientation		Target		Difficulty Level
Left	Right	Form 1	Form 2	
Opposite Orientation Trials				
bird	dog	dog	bird	E
running	sliding	running	sliding	M
lion	sheep	sheep	lion	D
train	bicycle	train	bicycle	M
throwing	eating	eating	throwing	E
firetruck	bus	bus	firetruck	D
banana	juice	banana	juice	E
sad	happy	happy	sad	D
cow	horse	horse	cow	M
cookies	milk	cookies	milk	E
blue	red	blue	red	D
riding	touching	touching	riding	M
apple	cheese	apple	cheese	E
washing	reading	reading	washing	M
Same Orientation Trials				
dog	bird	dog	bird	E
sliding	running	running	sliding	M
sheep	lion	sheep	lion	D
bicycle	train	train	bicycle	M
eating	throwing	eating	throwing	E
bus	firetruck	bus	firetruck	D
juice	banana	banana	juice	E
happy	sad	happy	sad	D
horse	cow	horse	cow	M
milk	cookies	cookies	milk	D
touching	riding	touching	riding	M
cheese	apple	apple	cheese	E
reading	washing	reading	washing	M

Note—For difficulty level, E (easy) = items comprehended by more than 66% of infants, M (moderate) = items comprehended by 33%–66% of infants, and D (difficult) = items comprehended by fewer than 33% of infants at 16 months of age.

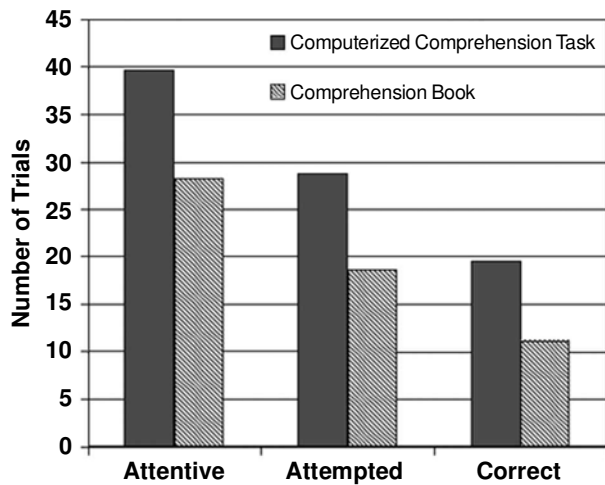


Figure 2. Attentive, attempted, and correct trials as a function of task. Infants attended to more trials than they attempted [$F(1,8) = 27.99, p < .05$] and attempted more trials than they correctly completed [$F(1,8) = 36.57, p < .05$].

and independent observations in order to track ambiguous responses and manually enter these responses in the output file. Our preliminary evaluation of test–retest reliability suggests that missing and ambiguous responses are best treated as errors of comprehension.

PRELIMINARY DATA

The primary impetus for developing this program and method of administration was to facilitate the assessment of comprehension in children younger than 20 months of age. Previous data from our laboratory and others, using the Comprehension Books (Friend & Diaz, 2001; D. Thal, personal communication, November 2000), indicate that it is exceedingly difficult to maintain attention with younger infants. Even breaking Comprehension Book assessments into several sets of trials does not successfully overcome this obstacle. Rather, when infants younger than 20 months of age complete a few items and disengage, it is virtually impossible to re-engage them in the task. This seems to reflect a genuine limitation in the age appropriateness of picture book assessments for maintaining attention in younger infants, a limitation addressed by the present procedure. In order to evaluate empirically the utility of the CCT for assessing vocabulary comprehension in children under 20 months of age, a within-subjects comparison was conducted on Comprehension Book and CCT performance at 16 months of age.

Method

Participants. Twelve 16-month-old infants (M age = 16 months, 19 days, range = 16 months, nine days to 17 months, 8 days; 6 females and 6 males) recruited through advertisements in local parenting and entertainment magazines participated as part of a larger study of vocabulary comprehension. A \$10 gift certificate to a local toy store was provided as an incentive.

Procedure. Data were collected in a within-subjects design in two testing sessions scheduled 1 week apart. The first testing session was always scheduled within 2 weeks of the infant’s 16-month birthday. In each session, one vocabulary assessment (Comprehension Book or CCT) was administered. The Comprehension Book was identical in content to the CCT, and both assessments were administered under identical optimal testing conditions. During each assessment, the infant was seated in the parent’s lap, and the parent wore dark glasses, the lenses of which were covered in black cardboard, and a pair of headphones, over which music played. In this way, the parents were prohibited from influencing their infant’s performance in either assessment. The only difference between the tasks was in the method of administration (picture book or touchscreen). The order of tasks was counterbalanced across participants.

Results

An order (Comprehension Book or CCT first) \times form (1 or 2; see Table 1) analysis of variance, with task and response (attentive, attempted, and correct trials) as repeated measures, revealed no effects of order or form and significant main effects of task [$F(1,8) = 9.639, p < .05$] and response [$F(2,7) = 41.365, p < .05$]. The effect of task was due to the infants’ attending, attempting, and producing correct responses on more trials for the CCT relative to the Comprehension Book. The effect of response reveals that the infants attended to significantly more trials than they attempted (touched or pointed) and that there were significantly more attempted than correct trials. Figure 2 summarizes the number of trials during which the infants were engaged (attentive), on which the infants attempted a response (attempts), and on which they correctly identified the target referent across the two assessments. These data are described more fully below. In addition, we present individual data for both assessments and reliability data for the CCT.

Attentive trials. To assess the utility of the CCT for engaging infants in a vocabulary comprehension task, we compared the number of trials on which the infants remained attentive for the Comprehension Book and the

Table 3
Number of Trials Correct, Incorrect, and Not Attempted as a Function of Task

Infant	Correct		Incorrect		No Attempt	
	CCT	CB	CCT	CB	CCT	CB
1	25	23	15	11	1	7
2	18	9	7	9	14	23
3	24	16	12	8	5	17
4	11	0	1	0	28	41
5	17	13	9	10	26	23
6	11	0	1	0	29	41
7	10	2	8	7	23	32
8	16	0	8	0	9	41
9	21	18	13	16	7	7
10	26	25	10	8	5	8
11	21	10	11	13	8	18
12	35	18	3	7	3	16
Mean	19.58	11.17	8.17	7.42	12.25	22.42
SD	7.34	9.12	4.55	5.16	9.66	13.21

Note—CCT, computerized comprehension task; CB, comprehension book.

Table 4
Proportion of Items Correctly Identified
by Each Participant as a Function of Difficulty Level

Infant	Easy		Moderate		Difficult	
	CCT	CB	CCT	CB	CCT	CB
1	.92	.46	.36	.71	.57	.50
2	.85	.38	.14	.14	.36	.14
3	.85	.62	.64	.43	.29	.14
4	.62	.00	.14	.00	.07	.00
5	.46	.38	.57	.21	.21	.36
6	.62	.00	.21	.00	.00	.00
7	.46	.00	.07	.07	.21	.07
8	.54	.00	.29	.00	.36	.00
9	.31	.38	.57	.50	.64	.43
10	.77	.62	.71	.57	.43	.64
11	.62	.38	.57	.14	.36	.21
12	.85	.54	.79	.43	.93	.36
Mean	.65	.31	.42	.27	.37	.24
SD	.19	.25	.25	.25	.25	.22

Note—CCT, computerized comprehension task; CB, comprehension book.

CCT. The infants were considered attentive if they looked at the images regardless of whether a touching or pointing response was elicited. The infants attended for a significantly greater number of trials for the CCT ($M = 39.58, SD = 4.91$) than for the Comprehension Book [$M = 28.25, SD = 17.39; t(11) = 2.538, p < .05$]. On average, the infants attended to about 70% of the trials for the Comprehension Book and over 95% of the trials for the CCT. The CCT elicits a significant increment in infant attention.

Attempted trials. The number of trials attempted is a measure of the motivation of infants to produce a response. Attempted trials were those trials on which the infant touched or pointed to an image regardless of whether that image constituted the target. The infants attempted significantly more trials for the CCT ($M = 28.75, SD = 9.67$) than for the Comprehension Book [$M = 18.58, SD = 13.21; t(11) = 4.336, p < .05$]. On average, the infants responded to the experimenter’s prompt by touching or pointing to an image on about 30% of the trials for the Comprehension Book, in contrast to 70% of the trials for the CCT. The CCT elicits a significant increment in infant responsiveness.

Number correct. We looked at the total number of touches or points to the target referent across the two assessments, as well as the proportion correct as a function of a priori item difficulty. The infants correctly identified the target referent significantly more often for the CCT ($M = 19.58, SD = 7.34$) than for the Comprehension Book [$M = 11.17, SD = 9.12; t(11) = 5.612, p < .05$].

To the extent that child performance reflects normative parent report of comprehension, we expected the children to be more likely to identify target referents for *easy* words than for *moderately difficult* or *difficult* words. Both tasks contained 14 easy words (comprehended normatively by at least 66% of 16-month-olds), 13 moderate words (comprehended by 34%–65% of 16-month-olds), and 14 diffi-

cult words (comprehended by fewer than 33% of 16-month-olds; Dale & Fenson, 1996). To assess the consistency of child performance with this classification, we looked at proportion correct for each difficulty level in each task. For the CCT, the infants correctly identified the target referent for 65% of the easy words ($SD = .19$), 42% of the moderate words ($SD = .25$), and 37% of the difficult words ($SD = .25$). For the Comprehension Book, in contrast, the infants correctly identified the target referent for 31% of the easy words ($SD = .25$), 27% of the moderate words ($SD = .25$), and 24% of the difficult words ($SD = .22$). The proportion of items for which the infants correctly identified the target referent for the CCT largely mirrored the a priori difficulty classification based on parent report. Infant performance in the Comprehension Book was attenuated, however, and was not consistent with a priori word difficulty.

Individual data. We examined individual patterns of response and trends in the comprehension of words at different difficulty levels across the two tasks. Data on correct responses, incorrect responses, and nonattempts (no touching or pointing response produced) for the individual participants are presented in Table 3. The majority of the infants were consistent in their response patterns across tasks. When an item was missed, 5 infants were more likely to produce no response at all than an incorrect response, and 2 infants were more likely to produce an incorrect response than no response. Five infants responded asymmetrically across tasks: 3 infants were more likely to produce an incorrect response than none at all for the CCT, but not for the Comprehension Book, and 2 infants showed the opposite pattern. In general, these response patterns appear to be more a function of individual differences than of task.

Individual variability was observed in the order of mastery of easy, moderate, and difficult words, suggesting that normative parent report data provide only a rough estimate of the order of acquisition of specific words for individual infants (see Table 4). Once sufficient data are accumulated, it may be appropriate in future revisions to base a priori word difficulty on normative child performance, rather than on parent report. Even so, individual variation in language acquisition may prove to be a limi-

Table 5
Number of Items Correctly Identified
for Each Participant in the Test and Reliability Phases

Infant	Test Phase	Reliability Phase
1	10	11
2	10	9
3	13	12
4	5	1
5	6	9
6	6	3
Mean	7.08	3.75
SD	2.87	4.94

Note—Values are number of reliability targets correctly identified out of a possible 14.

tation to the application of group level data to individual performance in this as in other extant procedures.

Reliability. Reliability data were collected on a subset of our sample ($n = 6$) for one third of the items for the CCT displayed in opposite left–right orientation, relative to the test phase. Missing data and ambiguous responses (e.g., simultaneous touches to both the target and the distractor) were coded as incorrect. The number of reliability items correct (out of 14) for each participant in the test and reliability phases is presented in Table 5. Performance in the test phase was significantly correlated with that in the reliability phase ($r = .828, p < .05$). This result is encouraging and suggests that the CCT provides reliable estimates of comprehension at the summary level. The significant increase in engagement elicited by the CCT does not appear to come at the expense of measurement reliability. However, the sample for the reliability assessment is quite small, and the stability of this correlation is uncertain.

These data suggest that the CCT is effective in eliciting and maintaining infant attention, motivating performance, and yielding a broad estimate of early comprehension. This represents a substantial improvement over previous methods, because attention is not confounded with vocabulary knowledge. Finally, our preliminary evaluation suggests that the CCT yields reliable estimates of comprehension vocabulary at 16 months of age.

RECOMMENDED APPLICATION

The software and procedure were designed specifically for the assessment of early vocabulary comprehension before 20 months of age. The psychometric properties of the assessment remain to be determined, and there are no data at this time on its clinical utility. It is currently recommended for use in experimental settings. To optimize the effectiveness of the stimulus images for eliciting infant attention, it is our experience that it is best to conduct the assessment in a quiet room with soft lighting.

AVAILABILITY

The software to run the vocabulary assessment is available free of charge on the CHILDES Web site at <http://childes.psy.cmu.edu>. Users are requested to share data obtained with this procedure with the first author. The source code may be obtained for the purpose of modification by writing to the first author. The source code and procedure may be used exclusively for academic purposes.

REFERENCES

Adobe Photoshop (Version 5.0) [Computer Software]. (1999). San Jose, CA: Adobe Systems Incorporated.

Art Explosion [Computer Software]. (1995–1998). Calabasas, CA: Nova Development Corporation.

BATES, E., BREThERTON, I., SNYDER, L., BEEGHLY, M., SHORE, C., McNEw, S., CARLSON, V., WILLIAMSON, C., GARRISON, A., & O'CONNELL, B. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. New York: Cambridge University Press.

DALE, P. S., & FENSON, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, **28**, 125-127.

FELDMAN, H. M., DOLLAGHAN, C. A., CAMPBELL, T. F., KURS-LASKY, M., JANOSKY, J. E., & PARADISE, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development*, **71**, 310-322.

FENSON, L., DALE, P. S., REZNICK, J. S., THAL, D., BATES, E., HARTUNG, J. P., PETHIK, S., & REILLY, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: Singular.

FRIEND, M., & DIAZ, K. (2001, December). *Relation between parent report and child performance on vocabulary comprehension*. Paper presented at the Conference on Early Lexicon Acquisition: Normal and Pathological Development, Lyon, France.

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.

HIRSH-PASEK, K., & GOLINKOFF, R. M. (1996). The intermodal preferential looking paradigm: A window into emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 105-124). Cambridge, MA: MIT Press.

NAIGLES, L. G., & GELMAN, S. A. (1995). Overextensions in comprehension and production revisited: Preferential-looking in a study of dog, cat, and cow. *Journal of Child Language*, **22**, 19-46.

OVIATT, S. L. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, **51**, 97-106.

RING, E. D., & FENSON, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language*, **20**, 141-159.

RODRIGUE, S. R. (2001). *Assessment of language comprehension and production in children from low-income and middle-income homes: Parent report versus child performance*. Unpublished master's thesis, San Diego State University.

SCHAFFER, G. (1998). *Word learning in infancy*. Unpublished doctoral dissertation, University of Oxford.

SNYDER, L. S., BATES, E., & BREThERTON, I. (1981). Content and context in early lexical development. *Journal of Child Language*, **8**, 565-582.

YODER, P. J., WARREN, S. F., & BIGGAR, H. A. (1997). Stability of maternal reports of lexical comprehension in very young children with development delays. *American Journal of Speech Language Pathology*, **6**, 59-64.

(Manuscript received January 16, 2002; revision accepted for publication December 18, 2002.)