Developmental Psychology
2012, Vol. 48, No. 1, 136–148

0012-1649/11/$12.00 DOI: 10.1037/a0025511

Evaluating the Predictive Validity of the Computerized Comprehension Task: Comprehension Predicts Production

Margaret Friend, Sara A. Schmitt, and Adrianne M. Simpson
San Diego State University

Until recently, the challenges inherent in measuring comprehension have impeded our ability to predict the course of language acquisition. The present research reports on a longitudinal assessment of the convergent and predictive validity of the MacArthur–Bates Communicative Development Inventories: Words and Gestures (CDI: WG; Fenson et al., 1994, 1993, 2007) and the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003, 2008). The CDI: WG and the CCT evinced good convergent validity; however, the CCT better predicted subsequent parent reports of language production. Language sample data in the 3rd year confirm this finding: The CCT accounted for 24% of the variance in unique word use. These studies provide evidence for the utility of a behavior-based approach to predicting the course of language acquisition into production.

Keywords: infant, comprehension, production

Understanding the developmental path to early literacy is a priority in psychological and educational research. Importantly, literacy is dependent upon earlier developing language comprehension and production and the activities that support it. Dovetailing with this focus is a long-standing focus on identifying those children at greatest risk for language delay early in development. These children are also at greatest risk for difficulties in school. One early potential indicator of developing literacy and school-readiness is the pace of language acquisition in late infancy and early toddlerhood.

The Computerized Comprehension Task (CCT; Friend & Keplinger, 2003, 2008) was developed to augment other language assessments in laboratory settings and to facilitate direct assessment in settings where resources for data coding and analysis may be limited. At present there is no standard direct measure of early receptive language that predicts language into the preschool period. The CCT builds on extant intermodal preferential looking (Hirsh-Pasek & Golinkoff, 1996; Hollich et al., 2000; Meints, Plunkett, & Harris, 1999) and picture book approaches (Ring & Fenson, 2000) and incorporates colorful images and reinforcing auditory stimuli that recruit attention. In a within-subjects assessment of word comprehension on the CCT and a picture book, Friend and Keplinger (2008) found that infants were significantly more attentive and attempted more trials on the CCT. The psychometric properties of the CCT are strong: Internal consistency and test–retest reliability have been demonstrated in toddlers who are acquiring English as a primary language (Friend & Keplinger, 2008; Robledo, personal communication, April 20, 2010). In addition, scores on the CCT correlate with parent report on the MacArthur–Bates Communicative Development Inventories: Words and Gestures (CDI: WG; Fenson et al., 1994, 1993, 2007; Friend & Keplinger, 2008). Because the procedure takes less than 10 min to administer and can be self-contained on any computer—even a laptop—with a touch sensitive screen, assessment can be made widely available in a variety of settings.

One goal of the present article is to assess the reliability and validity of the CCT in a sample whose demographic characteristics approximate the local population. A second goal of the present article is to predict language production in the third year as assessed by the MacArthur–Bates Communicative Development Inventories: Words and Sentences (CDI: WS; Fenson et al., 1994, 1993, 2007) and a language sample from comprehension scores on the CCT in the second year of life. We consider the relevant literature on demographic influences on language acquisition and on the relation between comprehension and production. Next, we review the recent work on language delay and consider the need for an independent measure of comprehension to facilitate the identification of children at risk. Finally, we evaluate the reliability and convergent and predictive validity of the CCT. One strength of this research is that it provides a tool for assessing language comprehension at the individual level that is independent of parent report and language sample data to clarify basic findings in language acquisition research.

Demographic Influences on Child Language

It is well-established that socioeconomic status (SES) in early childhood is positively related to child language and that it may have effects on parental reporting as well. For example, Feldman

This article was published Online First September 19, 2011.

Margaret Friend, Sara A. Schmitt, and Adrianne M. Simpson, Department of Psychology, San Diego State University.

Portions of this research were presented at the International Conference on Infant Studies (Baltimore, Maryland) in March 2010. Grateful acknowledgment is extended to Larry Fenson and Philip Dale for their suggestions on an earlier version of this article. Gratitude is also extended to the parents and children who participated in this research.

Correspondence concerning this article should be addressed to Margaret Friend, Department of Psychology, College of Sciences, San Diego State University, 6505 Alvarado Road, Suite 101, San Diego, CA 92120-4913. E-mail: mfriend@sciences.sdsu.edu

et al. (2000) reported higher scores on the CDI: WG for children whose mothers had completed high school or less or received Medicaid relative to children whose mothers had completed college or had private health insurance. One interpretation of this finding is that lower maternal education may lead to different expectations about developmental milestones relative to higher maternal education and that these expectations may be reflected in parent reports of early language (Feldman et al., 2000).

Demographic factors such as low income and maternal education may have broad consequences for early language acquisition beyond an influence on parent reports. Mothers with lower educational attainment engage in speech toward their children that is more directive and includes fewer questions than mothers with higher educational attainment. Mothers with less education also tend to talk less to their children and to use smaller vocabularies (Hoff, Laursen, & Tardif, 2002). Hart and Risley (1995) reported that children from lower SES families hear about 62,000 words a week in contrast to children from working class families who hear about 125,000 words and children from professional upper SES backgrounds who hear 215,000 words. Moreover, children of poor and working class parents appear to develop vocabulary more slowly than children of parents in professional occupations. Similarly Hoff and Tian (2005) reported that maternal education has direct effects on maternal language, and this, in turn, has effects on development: Both maternal vocabulary and length of utterance are positively associated with child language.

These findings are consistent with a growing literature that reports SES influences on parent–child interaction and child language outcomes (Beitchman et al., 2008; Vernon-Feagans et al., 2008; see Hoff, 2006, for a review) and underscore the importance of supplementing parent reports and language samples with independent behavioral measures to provide a clearer picture of the relationship between demographic factors and early language. One interest in the present article is whether SES effects similar to those reported for parent report and language samples obtain on the CCT. Additionally, we are interested in the utility of the CCT in predicting developmental outcomes. We turn now to a consideration of empirical findings on the relation between comprehension and production.

The Relation Between Comprehension and Production

The assertion that children's earliest facility with language should provide an estimate of future language development has a great deal of face validity. Nevertheless, extensive variability in language achievements at any one age coupled with similar variability in the rate of language acquisition has made the prediction of language production from comprehension a daunting task, and, a few decades ago, the topic generated considerable controversy. On the one hand, it was argued that the content and function of comprehension and production vocabularies are sufficiently distinct to suggest that they are dissociated processes (Benedict, 1979; Snyder, Bates, & Bretherton, 1981). Conversely, comprehension and production can be seen as related components of a single language process to the extent that overall levels of comprehension predict production. Bates, Benigni, Bretherton, Camaioni, and Volterra (1979) used detailed parent interviews and found that comprehension vocabulary was correlated with production vocabulary around the end of the first year of life. Snyder et al. (1981) used a similar procedure and reported an asymmetric relation between comprehension and production. Specifically, the correlation between comprehension and production was low, and comprehension outpaced production in some infants. However, Bates (1993) later concluded that the lack of sufficiently sensitive experimental measures may have contributed to the poor correspondence between comprehension and production.

More recently, in a large, demographically diverse sample, Feldman et al. (2000) reported modest but significant correlations between parent reports of comprehension on the CDI: WG and production on the CDI: WS across the period from roughly 1 to 2 years of age. Recent gains have also been made in predicting language outcomes from direct assessments as well. For example, Kuhl (2009) has shown that the perception of native and nonnative contrasts at 7 months of age is related to subsequent vocabulary size. Specifically, infants with high discrimination scores for native contrasts had larger productive vocabularies at 24 months than infants with low discrimination scores. Additionally, infants with high discrimination scores for non-native contrasts were at a disadvantage and had smaller vocabularies than infants with lower scores. These differences diminish after 24 months of age, suggesting that phonetic sensitivity to native speech contrasts best predicts the early phases of language acquisition. Kuhl (2009) suggested that phonetic sensitivity paves the way for word learning by increasing "neural commitment" (p. 226) to the native language and by enhancing attention to its statistical regularities.

Extending prediction beyond the second year, Marchman and Fernald (2008) used a modified preferential looking approach (the "looking-while-listening" task) and found that the speed with which children shift attention to a named referent at 25 months of age predicts language outcomes at 8 years. Marchman and Fernald concluded that the efficient processing of spoken language at 25 months facilitates subsequent language development by freeing cognitive resources to develop rich lexical representations.

Together, these recent studies reveal continuity between comprehension and production: Children's earliest facility with native words and sound patterns presages their facility with spoken language in spite of the considerable variability that exists between and within individual developmental trajectories. That is, the perceptual and cognitive abilities that scaffold language are good predictors of subsequent parent reports of language achievement. An independent, direct assessment of language comprehension prior to the second birthday is needed to extend these findings. Such a measure may enable us to assess risk for language delay sufficiently early in development to provide timely and effective interventions prior to school entry.

Risk for Language Delay

Children with persistent language delay exhibit difficulties in school as well as a host of other developmental problems (Snowling, Adams, Bishop, & Stothard, 2001; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006). Approximately 15%–16% of children at 2 years of age present with a language delay in the absence of known underlying sensory, cognitive, or neurological deficits (Desmarais, Sylvestre, Meyer, Bairati, & Rouleau, 2008; Law, Boyle, Harris, Harkness, & Nye, 2000), and those who are delayed in both language comprehension and production are at greatest risk for persistent language delay (Law et al., 2000). The

persistence of language delay among children delayed in production only is approximately 40%, whereas the persistence among children delayed in comprehension *and* production is approximately 75% (Law et al., 2000). Children who continue to show language delay at 5 years of age are significantly more likely to develop literacy difficulties than children whose language delay resolves by 3 or 4 years of age (Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998); furthermore, at 3 years of age, there is evidence for a direct link between oral language skills and first grade reading in typically developing children (National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network, 2005).

The fact that the persistence of language delay carries risk for literacy and other developmental difficulties and that children with delays in comprehension *and* production carry the highest risk argues strongly for predicting developmental risk on the basis of language comprehension. Yet, in Desmarais et al.'s (2008) review, only 5 of 217 studies on language delay assessed comprehension skills. Language comprehension is central to the development of productive language, literacy, and general cognitive skills but has received limited attention due in part to the difficulty of assessing comprehension early in development.

Two primary difficulties limit the measurement of early comprehension. First, infant attention is poor through the second year of life, making direct assessment of comprehension a challenge. Researchers have circumvented this problem by taking infant visual orientation to lexical targets as a measure of vocabulary comprehension (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Reznick, 1990; Schafer & Plunkett, 1998).

This approach has permitted exploration of a range of phenomena in early language acquisition, including the learning of novel words (Schafer & Plunkett, 1998); typicality effects in the comprehension of nouns, verbs, and prepositions (Meints et al., 1999; Meints, Plunkett, & Harris, 2002, 2008); and the early understanding of syntax (Golinkoff et al., 1987). However within-infant variability in preferential looking tasks is high, and convergent validity is often low (Hirsh-Pasek & Golinkoff, 1996; Hollich et al., 2000; Reznick, 1990; but see Killing & Bishop, 2008). For example, Reznick (1990) reported increases in visual fixation to lexical targets from 14 to 20 months of age. However, younger infants did not appear to understand the task, and data loss was high owing to the number of trials that had to be eliminated due to inattention. This speaks to the need to make word recognition tasks engaging and their contingencies transparent to ensure complete and meaningful data. These limitations make it difficult both to verify basic findings at the individual level and to evaluate existing individual findings from parent report.

A second difficulty is ease of administration. Basic research outside the laboratory as well as widespread screening for developmental risk requires an instrument that can be administered and scored easily in a range of environments without extensive data extraction and analysis (i.e., "portability"). One approach to these difficulties is the use of parent report. The MacArthur–Bates Communicative Development Inventories have been adapted into nearly 60 languages and for good reason: Parents observe their children in a wide range of contexts and are useful informants on the words their children understand and say. Nevertheless, it has been argued that memory demands (Stiles, 1994) and education (Feldman et al., 2000) may limit parent accuracy in detailing the potentially large set of words in an infant's receptive lexicon. In addition, parent reports may be overly based on context-specific knowledge (Tomasello & Mervis, 1994). This could lead to either under- or over-prediction of true vocabulary. For example, the parent may have limited knowledge of the full range of words that the child comprehends (under-prediction), or the child's comprehension may be limited to a small set of— or even a single— exemplar(s) (over-prediction). More recently, concern has been raised over the sensitivity and specificity of parent report (Law & Roy, 2008).

The present research evaluates the predictive validity of a computerized measure of early comprehension that is portable and overcomes the attention limitations of infants in the second year of life. Although it does not benefit from assessing the child's knowledge in multiple contexts as does parent report, it is a useful direct measure of early de-contextualized word knowledge. Indeed, it is potentially useful as a means of augmenting or clarifying parent report. Of interest is whether, at a summary level, comprehension as measured on the CCT predicts subsequent language production and, in particular, whether the CCT accounts for significant additional variance in language outcomes beyond that accounted for by parent reports of comprehension.

We present two studies, both of which are part of a larger longitudinal investigation of the relation between home environments and early language outcomes. In Study 1, we compare performance on the CCT at 16–21 months of age concurrently with parent reports on the CDI: WG in a small sample that approximates the demographic characteristics of the English-speaking population of San Diego County. In Study 2, we report on the predictive relation between comprehension and production in a subset of these participants 1 year later using the CCT and the CDI: WG comprehension scores at Time 1 as predictors of parent reports of production on the CDI: WS and children's performance in a language sample at Time 2.

Study 1

Method

Participants. Data were collected from 57 typically developing infants; data on seven participants were excluded, and the final data analyses were conducted on a sample of 50 infants. Exclusion was due to parental non-compliance (2), technical failure (2), and infant fussiness (3). The final sample consisted of 24 males and 26 females ranging in age from 16 to 21 months (M = 18;2 [months;days], range = 16;2–21;4). There were four infants at 16 months of age, 14 infants at 17 months of age, five infants at 18 months of age, seven infants at 19 months of age, seven infants at 20 months of age, and 13 infants at 21 months of age. Although we welcomed participation from both parents, more mothers (N = 48) than fathers (N = 2) participated. Thus, throughout this article, our discussions of "parent" and "parent report" refer to findings from a sample composed predominantly of mothers.

A majority of parents reported that their infants were exposed exclusively to English (62%); however, a substantial minority (38%) was exposed to a second language. Maternal ethnicity is reported for consistency with the census data. Participants were 72% Caucasian, 14% Hispanic, 4% Asian, 4% Black, and 6% mixed race (consisting of 4% Caucasian and Hispanic, and 2%

Caucasian and East Indian). These proportions are similar to those reported in the 2000 census of the English-speaking population in San Diego County. The average education level of parents was 15 years of formal education and ranged from 9 to 18 years, which was somewhat higher than the general population of the county: 95% of parents in our sample had attained a high school diploma, in contrast to 82.6% in the county, and 55% had attained a bachelor's degree or higher, in contrast to 29.5% in the county. Similarly, reported family income clustered about the mean (96,381) for the county. Although this sample approximates the local population, it must be acknowledged that it remains predominately middle-class. Selected demographic characteristics of the sample are presented in Table 1.

Participants were recruited through advertisements in local parenting magazines; community-based Internet resources; a free weekly newspaper; and flyers distributed at local libraries, daycare facilities, and low-income nutrition centers. Parents received a $10 gift certificate to a local store to compensate for travel to the laboratory, and infants received a small gift, either a toy or book, for their participation.

Measures.

Demographic measure. Parents reported demographic characteristics by completing a brief questionnaire at home, which they brought with them to the laboratory. All of our participants provided data on maternal education and ethnicity, 42 participants provided data on approximate income, and 49 participants provided data on second language exposure. These data were collected to characterize our sample and to permit an analysis of SES effects on CCT scores because maternal education, in particular, has been linked to reporting differences on the CDI (Feldman et al., 2000), to parent language practices (Hart & Risley, 1995; Hoff & Tian, 2005; cf. Vernon-Feagans et al., 2008), and to child language outcomes (Beitchman et al., 2008; Hoff & Tian, 2005).

CDI: WG. The CDI: WG is a well-established measure on which parents are queried on their child's understanding and production of a wide range of words, phrases, and gestures (Fenson et al., 1993, 2007). Of particular interest in the present study were parent reports of comprehension on the vocabulary checklist. Fenson et al. (1994) reported test–retest reliability for comprehension scores on the vocabulary checklist in the .80 range across a 1–3 month interval for infants older than 12 months of age. Percentile rankings for the CDI have recently been extended to 18 months of age; however, in our experience, the instrument can be extended productively to a majority of children at 20–21 months of age.

CCT. The CCT addresses the need for an extensive performance-based measure of comprehension in infants younger than 20 months of age. It contains vibrantly colored, high-quality digital images presented in a forced-choice format using lexical targets selected from the CDI: WG and the CDI: WS. Images are prototypical and contemporary exemplars of lexical targets matched within trials for brightness, size, color, word class, and word difficulty. Infants are prompted to touch images on a touch-sensitive screen in response to experimenter prompts in which target vocabulary items are embedded (e.g., "Where's the *dog*? Touch *dog*!"). When a correct touch is executed, a congruous

Table 1
Distribution of Selected Demographic Characteristics of Participants in Study 1

Characteristic	Boys (n = 24) n (%)	Girls (n = 26) n (%)	Total (N = 50) n (%)	% in county[a]
Maternal education				
High school or less	2 (8.3)	5 (19.2)	7 (14.0)	34.0
Some college	3 (12.5)	5 (19.2)	8 (16.0)	31.0
College graduate	10 (41.7)	6 (23.1)	16 (32.0)	21.1
Post-baccalaureate	9 (37.5)	10 (38.5)	19 (38.0)	12.7
Approximate income				
$15,000–$34,999	1 (4.2)	4 (15.4)	5 (10.0)	17.5
$35,000–$49,999	2 (8.3)	4 (15.4)	6 (12.0)	12.9
$50,000–$74,999	4 (16.7)	4 (15.4)	8 (16.0)	17.8
$75,000–$99,999	5 (20.8)	7 (26.9)	12 (24.0)	13.8
$100,000–$150,000	8 (33.3)	2 (7.7)	10 (20.0)	15.7
$150,000+	1 (4.2)	2 (7.7)	3 (6.0)	13.0
Maternal ethnicity[b]				
Asian	1 (4.2)	1 (3.8)	2 (4.0)	4.9
Black/not Hispanic	1 (4.2)	1 (3.8)	2 (4.0)	5.0
Hispanic[c]	2 (8.3)	5 (19.2)	7 (14.0)	15.5
White/not Hispanic	18 (75.0)	18 (69.2)	36 (72.0)	71.2
Mixed race	2 (8.3)	1 (3.8)	3 (6.0)	3.8
Second language				
No	16 (66.7)	15 (57.7)	31 (62.0)	64.8
Yes	7 (16.7)	5 (19.2)	18 (38.0)	35.2

Note. Adapted from "A Longitudinal Assessment of the Home Literacy Environment and Early Language," by S. A. Schmitt, A. M. Simpson, and M. Friend, 2011, *Infant and Child Development,* Table 1. Copyright 2011 by John Wiley & Sons, Ltd.

[a] Source is U.S. Census Bureau, 2006–2008 American Community Survey at http://factfinder.census.gov/. [b] For Asian and Hispanic categories, county values are corrected to reflect the proportion of adults who "speak English very well." [c] Hispanic ethnicity is not independent from race.

auditory stimulus is played. For example, prompted touches to the image of a dog elicit the sound of a dog barking. These sounds maintain interest and motivate compliance and occur only *after* a touch to the target image to avoid confounding responses to the sounds with lexical comprehension.

There are four training trials, 41 test trials, and 13 reliability trials in which images are presented in opposite left–right orientation on the screen. In each trial, two images appear simultaneously at left- and right-center screen. Blue screens are interleaved between trials, and the pace of the assessment is experimenter-controlled with the exception that trials have a maximum duration of 7 s before the screen is refreshed and a blue screen appears. This is necessary to ensure that the experimenter maintains control of the cursor at the end of the trial because children, having executed a correct response, often touch the target repeatedly. Extensive pilot testing indicates that this trial duration is optimal, enabling toddlers to execute a touch to the screen without unduly taxing attention. The maximum duration for the entire assessment is 41 trials $\times$ 7 s = ~5 min. An additional 1.5 min is required for reliability testing. Thus, the time required for administration of the entire assessment, including training trials, is less than 10 min.

There are equal numbers of easy (comprehended by more than 66% of 16-month-olds), moderately difficult (comprehended by 33%–66%), and difficult (comprehended by less than 33%) words defined a priori based on parent report data from the CDI: WG at 16 months of age (Dale & Fenson, 1996). The distribution of test items by word class and difficulty is presented in the Appendix. Like the CDI: WG, the CCT also has a high ceiling and, in our experience, can be productively extended to children at 20–21 months of age.

There are two forms of the test, such that the image that serves as the target is counterbalanced across forms, and form is counterbalanced across participants. Within forms, level of difficulty is matched within pairs and randomized across trials rather than presented in order of increasing difficulty. This organization helps to keep children engaged in the task because items understood by the majority of 16-month-olds are interspersed throughout the procedure, and correct touches to these targets produce auditory reinforcement. Targets appear with equal frequency on the right- and left-screen. Side is randomized with the restriction that targets appear no more than twice in succession on the same side consistent with the approach developed by Hirsh-Pasek and Golinkoff (1996) to reduce orientation-bias effects.

The CCT has strong immediate (.70) test–retest reliability, has moderate reliability over a 4-month delay (.61), and converges with parent report on the CDI: WG (.64). In addition, word class and word difficulty analyses reveal the predicted patterns of performance: Children are most accurate for nouns as contrasted with verbs and adjectives and for easy as contrasted with moderately difficult and difficult words (Friend & Keplinger, 2008; Friend & Zesiger, in press; Robledo, personal communication, April 20, 2010). Finally, exclusion rates for fussiness are low relative to many other infant-based tasks including Preferential-Looking. Test–retest reliability, convergence with parent report, and predicted word class and word difficulty effects indicate that the CCT is a valid and reliable measure of early word comprehension and that children's responses are independent and non-random.

Procedure. Participants initiated contact with the laboratory via telephone or e-mail. Once a participant contacted the laboratory, an initial screening interview was conducted over the telephone. Participants who met the screening criteria were scheduled to visit the laboratory. Selection criteria for infant participation were as follows: reported English exposure greater than or equal to 75% of total weekly language exposure, full-term pregnancy, and normal hearing and vision. Participants were also screened on the basis of parental education and occupation and to ensure a sample as representative as possible of the local population.

Approximately 1 week prior to the appointment, a basic demographic information sheet, the CDI: WG, and instructions for completing it were sent to the participants to be completed prior to their appointment. Parents were encouraged to consult with others who spent time with the child when filling out the CDI. They were also instructed to give credit for words with the same meaning and for mispronunciations but not for imitations when reporting vocabulary production.

Upon arrival to the laboratory, infants were given time to warm-up and play with toys in the lobby. As part of the warm-up, all infants engaged in "finger-painting" on a portable computer screen using a large brush setting in Microsoft Paint™ to become familiar and comfortable with the touch-sensitive screen on which the CCT is administered.

The CCT was administered on a kiosk situated on the floor of a separate testing room dimly lit by star- and moon-shaped night lights. Infants were seated in front of the kiosk with parents seated behind. The kiosk permits the toddler to sit in a low, cushioned chair or on their parent's lap to complete the procedure. A touch sensitive screen is enclosed in the kiosk and positioned 30 cm from the child and 16 cm above the floor at a 60° angle. Parents wore opaque glasses to prohibit them from guiding their infants' responses to experimenter's prompts. Parents were instructed not to repeat the experimenters' questions or interact with their child for the duration of the testing session. Responses were coded as correct if the child touched/pointed to the target image and were coded as incorrect if the child (1) touched/ pointed to the distractor, (2) provided an ambiguous response (e.g., a touch to the center of the screen), or (3) did not touch or point in response to the prompt.

Children participated in four training trials containing pairs of items that parents report are known by 85% of 16-months-olds (Dale & Fenson, 1996): diaper and shoe, cup and sock, bathtub and door, and book and car. These trials familiarized participants with the contingencies of the testing situation (i.e., colorful images will appear, and touching the one that corresponds to the experimenter's prompt will elicit an interesting sound). Once the child successfully completed these trials, testing commenced and proceeded until all of the trials were completed or until the child failed to respond on four consecutive trials. The proportion of correct responses to trials presented before the procedure was terminated was taken as the final comprehension score on the CCT. Only those children who could not be engaged in the task during the training phase were excluded from further participation ($N = 3$).

In each trial, target and distractor images appeared on the touch sensitive screen following the experimenter's prompt. The prompt was dependent upon word class:

For nouns, "Where's the ___? Touch ___";

For verbs, "Who's ___? Touch ___"; and

For adjectives, "Which one is ___? Touch ___."

If a child failed to respond on two consecutive trials, the experimenter reintroduced the contingency between a correct re-

sponse and an auditory reinforcement by touching the target image. If a child attempted to touch the target image but failed to do so within the 7-s window, the image pair was re-presented. Following the presentation of 41 test trials, 13 image pairs were re-presented in opposite left-right orientation for reliability testing. We administered the reliability assessment only to those infants who remained alert and attentive throughout the testing session ($N = 31$).

Results and Discussion

The average vocabulary comprehension score on the CDI: WG was 234.92 ($\sigma = 94.83$) for girls and 217.88 ($\sigma = 107.97$) for boys. Scores ranged from the 1st to the 99th percentile. Only three of our 50 participants reached ceiling (one 18-month-old and two 21-month-olds). The average vocabulary comprehension score on the CCT was 17.92 ($\sigma = 7.90$) for girls and 16.71 ($\sigma = 7.17$) for boys. Scores ranged from four to 32 correct responses (out of a possible 41). As expected, immediate test–retest reliability at the summary level on the CCT was strong, $r(48) = .69, R^2 = .58, p < .001$. To explore the consistency of responses to specific exemplars, we calculated item-level agreement. On average, children responded consistently across test and reliability phases on 67% of the items. However the median (71%) and mode (79%) better capture the central tendency of the distribution of agreement scores. This strong reliability for items presented in opposite left–right orientation from test indicates that side-bias effects were negligible. Both the CDI: WG and CCT vocabulary comprehension scores were approximately normally distributed. There was good item-level agreement between parent reports and child performance on the CCT ($M = .54$), and parent–child agreement scores were normally distributed. Age was significantly related to CDI: WG scores, $r(48) = .52, p < .001$, and CCT scores, $r(48) = .29, p = .03$.

We assessed children's use of static images as referents by comparing CCT responses as a function of word class to parent report on the CDI: WG. Children's knowledge of nouns, verbs, and adjectives on the CCT was significantly correlated with parent reports of noun, verb, and adjective knowledge: $r(48) = .412, p = .003$; $r(48) = .357, p = .01$; and $r(48) = .415, p = .003$, respectively. The average item-level agreement was .54 for nouns, .42 for verbs, and .55 for adjectives. Children appeared to recognize the static images as referents; however, there was a slight attenuation in performance on verbs. There was no effect of test form on CCT scores.

Demographic analyses. Descriptive data on vocabulary comprehension scores on the CDI: WG and the CCT are presented as a function of demographic characteristics in Table 2. We conducted preliminary correlations to assess the influence of age, sex, maternal education, approximate family income, and exposure to a second language (in hours per week) on comprehension vocabulary scores on the CDI: WG and the CCT. No significant relation was observed between any demographic variable and comprehension vocabulary scores.

We conducted two regression analyses using only maternal education and age as predictors first on the CDI: WG and then on the CCT. For the CDI: WG, this model was significant, $F(2, 47) = 8.50, p = .001$, and age was the only significant predictor, $t(47) = 4.08, p < .001, B = .52$, with the effect of maternal education not approaching significance. For the CCT, the model was also significant, $F(2, 47) = 4.18, p = .01$. Both maternal education and age were significant predictors of comprehension scores on the CCT, $t(47) = 2.23, p = .03, B = .31$; $t(47) = 2.61, p = .01, B = .36$, respectively. As can be seen in Table 2, the effect of maternal education reflects an advantage for children of college-educated mothers, such that they are reported to know more lexical items than children of mothers without a college degree. The age effect is consistent with reports on two previous studies of this instrument (Friend & Keplinger, 2003, 2008).

Table 2
Mean Language Comprehension (and σ) as a Function of Selected Demographic Characteristics in Study 1

	CDI: WG	CCT
Characteristic	*M* (σ)	*M* (σ)
Maternal education		
High school or less	224.83 (95.90)	15.83 (6.52)
Some college	224.33 (130.52)	13.33 (9.37)
College graduate	262.00 (90.51)	20.06 (6.43)
Post-baccalaureate	198.79 (93.12)	17.42 (7.95)
Approximate income		
$15,000–$24,999	190.00 (100.93)	13.60 (11.97)
$35,000–$49,999	237.00 (154.42)	17.33 (9.99)
$50,000–$74,999	283.13 (87.97)	21.00 (5.85)
$75,000–$99,999	189.50 (86.77)	15.75 (4.82)
$100,000–$150,000	261.60 (101.84)	17.80 (9.33)
$150,000+	226.00 (74.65)	21.33 (1.53)
Maternal ethnicity		
Asian	287.00 (154.15)	15.00 (11.31)
Black/not Hispanic	241.00 (46.67)	21.00 (2.83)
Hispanic	210.00 (130.29)	16.00 (9.07)
White/not Hispanic	229.42 (90.84)	17.58 (8.03)
Mixed race	184.00 (179.02)	16.67 (2.31)
Second language		
No	232.68 (99.16)	16.90 (7.24)
Yes	217.05 (104.93)	18.05 (8.70)

Note. CDI: WG = MacArthur–Bates Communicative Development Inventories: Words and Gestures; CCT = Computerized Comprehension Task.

Convergent validity. To assess the convergent relation between scores on the CDI: WG and the CCT, we evaluated the correlation between these scores after partialling out the effect of age. Of interest was whether parent reports of vocabulary comprehension account for significant variance in CCT scores beyond that accounted for by the age of the participants. Vocabulary comprehension scores on the CDI: WG accounted for a small but significant proportion of unique variance in performance on the CCT above that accounted for by age, *partial* $r(47) = .361, p < .01, R^2 = .13$.

Consistent with previous research on English- (Friend & Keplinger, 2003, 2008) and French-speaking (Friend & Zesiger, in press) samples, vocabulary comprehension scores on the CCT converge with parent reports on the CDI: WG. The present results confirm the test–retest reliability and convergent validity of the CCT in a sample whose demographic characteristics approximate the local population. In addition, an analysis of responses on the CCT indicated that infants accurately identify the referent for a word on the majority of trials that they attempt. Thus, children do

not appear to be responding randomly; rather, their responses appear tied to word knowledge. However, the modest correlation between scores on the CDI: WG and the CCT reveal that these instruments are capturing different aspects of that knowledge.

One question that remains is whether the vocabulary comprehension scores on the CCT can predict subsequent language development. Friend and Keplinger (2008) reported stability in comprehension scores on the CCT from 16 to 20 and from 20 to 24 months of age. What is unknown is whether comprehension scores on the CCT can predict language production. If so, this would suggest that direct measures of early comprehension, particularly the CCT, have utility for predicting the trajectory of language acquisition as it transitions into production.

Study 2

Method

Participants. A sample of 25 children (14 girls, 11 boys) between the ages of 24 and 41 months (M = 29;3) and their primary caregivers from Study 1 returned to the laboratory to participate in Study 2. The distribution of selected demographic characteristics can be found in Table 3. Of this sample, 80% of mothers were Caucasian, 12% were Hispanic, 4% were Asian, and 4% were of mixed race. The average maternal education level was 16 years of formal education and ranged from 9 to 18 years.

The interval between the first and second laboratory visit ranged from 6 to 20 months (M = 10;11). One participant did not contribute a language sample on the second visit due to equipment failure. Thus, the final sample for the language sample analyses was 24 participants, 14 girls and 10 boys between the ages of 24 and 41 months (M = 28;27), with a mean interval of 10 months between first and second testing (range = 6–20 months).

Measures.

CDI: WS. Like the CDI: WG, the CDI: WS is a well-established parent report measure of child language designed for use with children from 16 to 30 months of age. The CDI: WS is designed to elicit parent estimates of productive vocabulary as well as utterance length and sentence complexity.

Language sample. During their second visit, parents and their children participated in 10 min of toy play and 10 min of shared book reading. The entire 20-min session was recorded onto Digital Audio Tape (DAT) from an Audio Technica AT898 Subminiature Cardioid Condenser microphone. The DAT tapes were transcribed independently by two coders. Interrater agreement was established on an utterance by utterance basis at a level equal to or exceeding .80 for 10% of the participants (N = 3) across the full 20-min language sample. Disagreements were resolved by discussion.

For each language sample, we obtained two measures: MLU (the mean length of utterance in morphemes—a measure of syntactic complexity) and NDW (number of different words, or unique word roots—a measure of lexical diversity). For the MLU analyses, all complete and intelligible utterances were sampled. For the NDW analyses, we obtained two estimates: one from the full set of complete and intelligible utterances (to capture the full variability of children's language production) and one from the first 50 complete and intelligible utterances (in the interest of standardizing the sample). Due to the low verbal productivity and intelligibility of several participants, we combined transcripts across the toy play and book reading situations to obtain 50 complete and intelligible utterances for the NDW analyses. Even so, there were three participants for whom the number of complete and intelligible utterances did not reach this threshold

Table 3
Distribution of Selected Demographic Characteristics of Participants in Study 2

	Boys (n = 11)	Girls (n = 14)	Total (N = 25)	
Characteristic	n (%)	n (%)	n (%)	% in county
Maternal education				
High school or less	1 (9.1)	2 (14.3)	3 (12.0)	34.0
Some college	1 (9.1)	3 (21.4)	4 (16.0)	31.0
College graduate	5 (45.5)	3 (21.4)	8 (32.0)	21.1
Post-baccalaureate	4 (36.4)	6 (42.9)	10 (40.0)	12.7
Approximate income				
$15,000–$24,999	0 (0)	3 (21.4)	3 (12.0)	16.5
$35,000–$49,999	0 (0)	3 (21.4)	3 (12.0)	12.9
$50,000–$74,999	1 (9.1)	2 (14.3)	3 (12.0)	17.8
$75,000–$99,999	7 (63.6)	2 (14.3)	9 (36.0)	13.8
$100,000–$150,000	3 (27.3)	1 (7.1)	4 (16.0)	15.7
$150,000+	0 (0)	3 (21.4)	3 (12.0)	13.0
Maternal ethnicity				
Asian	0 (0)	1 (7.1)	1 (4.0)	4.9
Black/not Hispanic	0 (0)	0 (0)	0 (0)	5.0
Hispanic	1 (9.1)	2 (14.3)	3 (12.0)	15.5
White/not Hispanic	9 (81.8)	11 (78.6)	20 (80.0)	71.2
Mixed race	1 (9.1)	0 (0)	1 (4.0)	3.8
Second language				
No	8 (72.7)	7 (50.0)	15 (60.0)	64.8
Yes	3 (27.3)	7 (50.0)	10 (40.0)	35.2

(one 26-month-old, one 28-month-old, and one 29-month-old). The completed transcripts were analyzed using the Systematic Analysis of Language Transcripts (SALT) software (Miller & Iglesias, 2008).

Procedure. Participants were contacted by telephone. Approximately 1 week prior to their scheduled visit, parents were mailed the CDI: WS to complete at home, and they were asked to bring it to the laboratory on the day of the appointment. The researcher and dyads were introduced during a warm-up period of 10 min during which children played with toys in the lobby. After the warm-up period, dyads were guided into the testing room for observation. Adjacent to this room was an observation room equipped with a one-way mirror behind which a DAT recorder was stationed. Attached to the DAT recorder was a microphone housed discreetly in a conduit between the testing and observation rooms. For the first 10 min, caregivers were asked to play with their child in the same manner as they would at home with a toy farmhouse that included moving parts, animal sound effects, and a set of toy animals. During the second 10 min, caregivers were asked to read with their child just as they would at home from a selection of three age-appropriate picture books: *Biscuit's Picnic* by Alyssa Satin Capucilli, *Where Is the Big Red Doggie?* by Norman Bridwell, and *Mighty Tugboats* by Noeline Cassetari. Pilot-testing revealed that the transition from the warm-up period to the observational session was smoother when the toy play situation preceded the book-reading situation. For this reason, toy play preceded book reading for all participants.

Results and Discussion

Descriptive data. There was no difference between girls and boys in average vocabulary production scores on the CDI: WS or in language sample MLU and NDW (family-wise $\alpha = .05$). Parents reported an average of 455.32 ($\sigma = 144.15$) words produced on the CDI: WS. Scores ranged from the 8th to the 99th percentile for girls and from the 25th to the 65th percentile for boys. Because three of our participants were older than the reference sample for the CDI: WS, we conducted a preliminary analysis to determine whether vocabulary production was related to age and whether there were any statistical outliers. There was no effect of age on the CDI: WS (or on any language measure) and there were no outliers. In view of this, all participants were included in subsequent analyses involving the CDI: WS.

Because the toy play situation (MLU = 2.55, $\sigma = 0.67$) and the book reading situation (MLU = 2.35, $\sigma = 0.77$) were not counterbalanced, we report analyses using summary MLU across the two situations ($M = 2.48$, $\sigma = 0.68$). Also, in the interest of obtaining a sufficiently large number of utterances for the NDW analyses, the number of unique word roots was calculated across the entire 20-min observation. The average NDW for the full transcripts was 89.29 ($\sigma = 36.02$). There were no striking demographic effects, and, given the small sample size, we did not conduct additional demographic analyses. Descriptive data on language measures as a function of demographic characteristics are presented in Table 4.

Table 4
Mean Language Production (and σ) as a Function of Selected Demographic Characteristics in Study 2

	CDI: WS	Language sample	
	Production vocabulary	MLU	Unique words
Characteristic	*M* (σ)	*M* (σ)	*M* (σ)
Maternal education			
High school or less	469.33 (186.85)	2.00 (0.21)	75.33 (16.65)
Some college	409.00 (257.88)	2.31 (0.56)	80.50 (48.74)
College graduate	482.75 (31.49)	3.07 (0.47)	112.71 (42.56)
Post-baccalaureate	447.70 (101.43)	2.27 (0.71)	80.60 (30.95)
Approximate income			
$15,000–$24,999	481.67 (113.73)	1.80 (0.19)	62.00 (14.73)
$35,000–$49,999	634.67 (48.34)	2.75 (0.46)	112.00 (16.37)
$50,000–$74,999	375.67 (276.49)	2.52 (0.64)	89.00 (58.59)
$75,000–$99,999	424.22 (139.78)	2.59 (0.78)	88.62 (43.65)
$100,000–$150,000	421.75 (96.04)	2.68 (0.81)	97.75 (38.70)
$150,000+	466.00 (5.66)	2.64 (0.34)	86.50 (3.53)
Maternal ethnicity			
Asian	462.00	2.88	84.00
Hispanic	552.33 (81.83)	2.13 (0.26)	91.33 (20.13)
White/not Hispanic	447.75 (151.22)	2.45 (0.69)	86.68 (38.36)
Mixed race	309.00		
Second language			
No	460.87 (126.12)	2.42 (0.60)	87.93 (31.31)
Yes	447.00 (174.77)	2.55 (0.81)	91.20 (43.50)

Note. Unique word roots are reported across the full language sample, although both the full sample and the first 50 utterances were analyzed. CDI: WS = MacArthur–Bates Communicative Development Inventories: Words and Sentences; MLU = mean length of utterance in morphemes.

Scatterplots of the CDI: WG subscales and the CCT were examined with respect to their relation to language outcomes. Both the CDI: WG and CCT comprehension vocabulary scores exhibited acceptable range and serve as predictors. Of interest is whether vocabulary comprehension performance on the CCT predicts significant variance in language production beyond that predicted by parent reports on the CDI: WG.

Prediction to CDI: WS production. To assess the predictive relation between parent reports of vocabulary production and vocabulary comprehension scores obtained during participants' initial visit to the laboratory 6–20 months earlier, we first examined the zero-order correlations between these measures. As can be seen in Table 5, both the CDI: WG and CCT scores are strongly correlated with reported production on the CDI: WS. Next, we examined the prediction of CDI: WS scores from each of the comprehension measures in turn, to determine the unique variance in production vocabulary accounted for by the CDI: WG and the CCT. The partial correlation between the CDI: WG and the CDI: WS was not significant when controlling for CCT scores. In contrast, the correlation between comprehension scores on the CCT and production score on the CDI: WS remained strong when controlling for CDI: WG scores. We repeated these analyses controlling for age at second testing and found that this basic pattern of findings was maintained.

The CCT significantly predicts parent reports of language production on the CDI: WS in the third year. Moreover, the CCT accounts for greater variance in subsequent language production than parent reports of comprehension on the CDI: WG obtained during the same testing session. Of greater importance, however, is the extent to which estimates of language comprehension predict the language sample.

Prediction to the language sample. Zero order correlations reveal that the relation between the comprehension measures in Study 1 and productive lexical diversity in Study 2 as measured by the number of unique word roots is strong and positive (see Table 5). The partial correlations of each predictor with children's NDW scores clarified these relationships. Because the sample size is small, we focus on relative effect sizes across our two predictors. The CCT accounted for 24% of the variance in children's unique word use and 13% of unique variance, controlling for the CDI: WG (see Figure 1). In contrast, the CDI: WG accounted for 16% of the variance in children's unique word use but only 4% of unique variance. This pattern was maintained when we repeated these analyses controlling for age at second testing. Earlier performance on the CCT accounts for unique variance over and above parent reports of comprehension on the CDI: WG. Parent reports on the CDI: WG do not account for significant unique variance over and above scores on the CCT.

Next, we evaluated the prediction from comprehension to language sample MLU. Parent report on the CDI: WG was superior to child performance on the CCT as a predictor of sentence complexity in the third year. The CDI: WG accounted for 20% of unique variance in MLU after controlling for the CCT, whereas the CCT accounted for 10% of unique variance after controlling for the CDI: WG. When we repeated these analyses controlling for age at second testing, both measures accounted for comparable variance. Whereas the CCT accounted for greater variance in the overall size of children's spontaneous vocabulary as indexed by the number of unique words used in a language sample, parent reports on the CDI: WG and child performance on the CCT were roughly equal indicators of subsequent MLU.

General Discussion

Demographic Influences on Child Language

The demographic variables selected for inclusion in this study, with the exception of maternal education, did not account for significant variance on either the CDI: WG or the CCT. Consistent with previous research indicating an influence of SES and maternal education on parent language input and child outcomes (Beitchman et al., 2008; Hart & Risley, 1995; Hoff et al., 2002; Vernon-Feagans et al., 2008), maternal education accounted for a small but significant proportion of the variance in scores on the

Table 5
Zero-Order and Partial Correlations for Comprehension and Production Measures

	Production measures		
Comprehension measures	CDI: WS	Language sample MLU	Language sample NDW
CDI: WG			
Zero-order	.503*	.605**	.397
Controlling for CCT	.278	.450*	.190
Controlling for CCT and age	.151	.350	.058
CCT			
Zero-order	.645**	.539**	.490*
Controlling for CDI: WG	.500*	.329	.361
Controlling for CDI: WG and age	.520*	.325	.359

Note. CDI: WS = MacArthur–Bates Communicative Development Inventories: Words and Sentences; MLU = mean length of utterance in morphemes; NDW = number of different words, or unique word roots; CDI: WG = MacArthur–Bates Communicative Development Inventories: Words and Gestures; CCT = Computerized Comprehension Task. MLU and NDW refer to the total unique word roots for this table taken across all complete and intelligible utterances in the language sample. Degrees of freedom are 22 for the CDI: WS partial correlations and 21 for all other partial correlations.
* $p < .05$. ** $p < .01$.

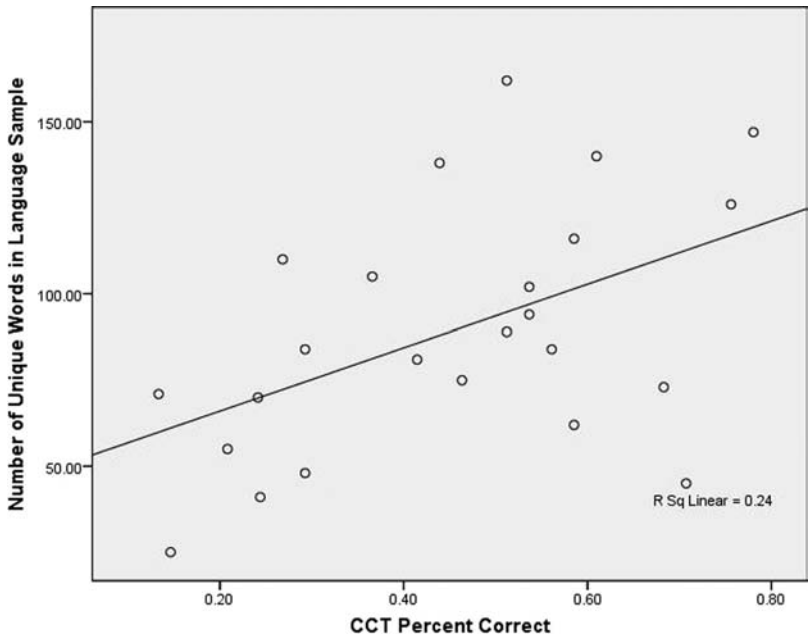

Figure 1. Relation between scores on the Computerized Comprehension Task (CCT) and children's unique word use in a language sample.

CCT indicating that children of college-educated mothers were more successful than children of mothers who had not completed college at identifying word–referent relationships.

Vernon-Feagans et al. (2008) argued that the relation between demographic characteristics and child language outcomes is mediated by maternal language. That is, demographic characteristics (such as maternal education and income) may have their effects on child language indirectly through their impacts on maternal vocabulary and language use. It must be noted that the variance in language status accounted for by SES can be small, if significant, particularly early in development. For example, Hoff and Tian (2005) found that SES accounted for 5% of the variance in vocabulary size in a sample of children in the United States roughly the same age as the present sample. Consistent with this finding, maternal education accounted for 5% of the variance in vocabulary comprehension scores on the CCT, suggesting that the CCT captures established variability in child language as a function of SES. At the same time, SES influences do not appear to have a disproportionate effect on CCT scores.

The Relation Between Comprehension and Production

In the present sample, both parent report on the CDI: WG and child performance on the CCT predicted parent report on the CDI: WS. This finding is consistent with Feldman et al. (2000), who found that CDI: WG comprehension at 12 months of age predicted CDI: WS production at 24 months. However, the CDI: WG did not account for unique variance in parent reports of language production beyond that accounted for by child performance on the CCT.

The CDI: WG was a stronger predictor than the CCT of language sample MLU 6–20 months later, accounting for about 20% of the variance. However, the CCT emerged as a stronger predictor of children's vocabulary as measured by the number of unique words used in the language sample, accounting for 24% of the variance in children's unique word use (or 13% of the unique variance when controlling for the influence of scores on the CDI: WG). Parent report and direct assessment provide different windows on early language comprehension. Parents, on the one hand, have a wide range of experience with the child, in different settings, from which they can draw when assessing their child's word knowledge. However, it is possible that children evince knowledge of some words under very specific circumstances and with very specific exemplars. That is, some of the word knowledge that parents report may be context-dependent. To the extent that this is the case, we might expect parents to err on the side of over-estimating early word knowledge by treating children's context-dependent responses as generalized and flexible word comprehension. Alternatively, if the parent does not see their child in a wide range of contexts, under-prediction may occur. In contrast to parent report, direct assessments such as the CCT assess word knowledge without the supporting context to which children may be accustomed, and the exemplars likely vary from the ones with which children are most familiar. In addition, there is the added burden of touching or pointing on cue. Identifying word-referent relations on the CCT likely requires rather stable word representations and direct assessments like the CCT most likely err on the side of under-estimating early word knowledge. Yet this arguably conservative estimate predicts parent reports of vocabulary production and children's spontaneous word use several months later. These findings reveal, consistent with other recent reports, that we can predict the course of language acquisition into production at a group level from early direct measures of comprehension.

The CCT is the first assessment to predict language development from the second through the third year of life. It allows us to assess the continuity of language acquisition behaviorally with a high participant retention rate and provides an estimate of vocabulary knowledge that converges with parent report. However, unlike parent report, it requires a de-contextualized, volitional response from the child. By combining the two approaches, we may be better able to define the typical range of early word knowledge.

The finding that an easy-to-administer behavioral task can predict children's language as it is transitioning from comprehension to production also has implications for clinical work. Toddlers with combined deficits in comprehension and production have been shown to be at the greatest risk for persistent clinically significant language delay (Law et al., 2000), and they remain at risk for deficits in academic and social functioning as they mature (Snowling et al., 2001, 2006).

The present findings suggest that the CCT is a potential new approach to this long-standing problem. A chief advantage of the CCT is that it is a portable, easy-to-administer assessment appropriate to the attention limitations of children in the second year of life. The CCT has revealed itself to be a reliable and valid measure of comprehension vocabulary in the second year and a significant predictor of production in the third year of life. These findings hold across children exposed only to English and those who heard another language at home. However, it is important to note two limitations of the present research: a small sample size and a variable measurement interval between first and second testing. These limitations argue for replication in larger samples with a consistent measurement interval. In addition, predicting production at the group level is not evidence of clinical utility. Future research to establish the sensitivity and specificity of the CCT in identifying children who may be at risk for language delay and other developmental problems is needed.

References

Bates, E. (1993). Comprehension and production in early language development. *Monographs of the Society for Research in Child Development, 58*(3–4), Serial No. 233.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York, NY: Academic Press.

Beitchman, J. H., Jiang, H., Koyama, E., Johnson, C. J., Escobar, M., Atkinson, L., . . . Vida, R. (2008). Models and determinants of vocabulary growth from kindergarten to adulthood. *Journal of Child Psychology and Psychiatry, 49,* 626–634. doi:10.1111/j.1469-7610.2008.01878.x

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6,* 183–200. doi:10.1017/S0305000900002245

Dale, P., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28,* 125–127. doi:10.3758/BF03203646

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders, 43,* 361–389. doi:10.1080/13682820701546854

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. J., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development, 71,* 310–322. doi:10.1111/1467-8624.00146

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.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., . . . Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA: Singular Publishing Group.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur–Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Brookes.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers, 35,* 302–309. doi:10.3758/BF03202556

Friend, M., & Keplinger, M. (2008). Reliability and validity of the Computerized Comprehension Task (CCT): Data from American English and Mexican Spanish infants. *Journal of Child Language, 35,* 77–98. doi:10.1017/S0305000907008264

Friend, M., & Zesiger, P. (in press). A systematic replication of the psychometric properties of the CCT in three languages: English, Spanish, and French. *Enfance*.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical syntactic comprehension in a new paradigm. *Journal of Child Language, 14,* 23–45. doi:10.1017/S030500090001271X

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

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26,* 55–88. doi:10.1016/j.dr.2005.11.002

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 2: Biology and ecology of parenting* (2nd ed., pp. 231–252). Mahwah, NJ: Erlbaum.

Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders, 38,* 271–278. doi:10.1016/j.jcomdis.2005.02.003

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. B., Brown, E., Chung, H. L., . . . Rocroi, C. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65*(3), Serial No. 262.

Killing, S. E. A., & Bishop, D. V. M. (2008). Move it! Visual feedback enhances validity of preferential looking as a measure of individual differences in vocabulary in toddlers. *Developmental Science, 11,* 525–530. doi:10.1111/j.1467-7687.2008.00698.x

Kuhl, P. K. (2009). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In J. Colombo, P. McCardle, & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research disorders* (pp. 213–244). New York, NY: Psychology Press.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language & Communication Disorders, 35,* 165–188. doi:10.1080/136828200247133

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health, 13,* 198–206.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11,* F9–F16. doi:10.1111/j.1467-7687.2008.00671.x

Meints, K., Plunkett, K., & Harris, P. L. (1999). When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology, 35,* 1072–1078. doi:10.1037/0012-1649.35.4.1072

Meints, K., Plunkett, K., & Harris, P. L. (2002). What is "on" and "under" for 15-, 18-, and 24-month-olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology, 20,* 113–130. doi:10.1348/026151002166352

Meints, K., Plunkett, K., & Harris, P. L. (2008). Eating apples and houseplants: Typicality constraints on thematic roles in early verb learning. *Language and Cognitive Processes, 23,* 434–463. doi:10.1080/01690960701726232

Miller, J., & Iglesias, A. (2008). Systematic Analysis of Language Transcripts (SALT), Student Version 2008 [Computer software]. Middleton, WI: SALT Software.

National Institute of Child Health and Human Development Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428–442. doi:10.1037/0012-1649.41.2.428

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics, 11,* 145–166. doi:10.1017/S0142716400008742

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. doi:10.1177/014272370002005902

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

Snowling, M. J., Adams, J. W., Bishop, D. V. M., & Stothard, S. E. (2001). Educational attainments of school leavers with a preschool history of speech-language impairments. *International Journal of Language & Communication Disorders, 36,* 173–183. doi:10.1080/13682820010019892

Snowling, M. J., Bishop, D. V. M., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry, 47,* 759–765. doi:10.1111/j.1469-7610.2006.01631.x

Snyder, L. S., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8,* 565–582. doi:10.1017/S0305000900003433

Stiles, J. (1994). On the nature of informant judgments in inventory measures: And so what is it you want to know? *Monographs of the Society for Research in Child Development, 59,* 180–185.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41,* 407–418.

Tomasello, M., & Mervis, C. B. (1994). This instrument is great but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development, 59,* 174–179.

U.S. Census Bureau. (2010). *2006–2008 American Community Survey*. Retrieved from http://factfinder.census.gov/

Vernon-Feagans, L., Pancsofar, N., Willoughby, M., Quade, A., Cox, M., & The Family Life Key Investigators. (2008). Predictors of maternal language to infants during a picture book task in the home: Family SES, child characteristics, and the parenting environment. *Journal of Applied Developmental Psychology, 29,* 213–226. doi:10.1016/j.appdev.2008.02.007

Appendix

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

Orientation		Target		
Left	Right	Form 1	Form 2	Difficulty level
Dog	Bird	Dog	Bird	Moderate
Sliding	Running	Running	Sliding	Moderate
Mouth	Eye	Mouth	Eye	Easy
Sheep	Lion	Sheep	Lion	Difficult
Orange	Green	Green	Orange	Difficult
Kissing	Hugging	Hugging	Kissing	Easy
Pulling	Swimming	Pulling	Swimming	Difficult
Telephone	Keys	Telephone	Keys	Easy
Kicking	Drawing	Drawing	Kicking	Difficult
Bus	Fire truck	Bus	Fire truck	Difficult
Nose	Foot	Foot	Nose	Easy
Happy	Sad	Happy	Sad	Difficult
Button	Hat	Button	Hat	Moderate
Juice	Banana	Banana	Juice	Easy
Old	New	Old	New	Difficult
Toothbrush	Spoon	Toothbrush	Spoon	Easy
Drinking	Dancing	Dancing	Drinking	Easy
Swinging	Jumping	Jumping	Swinging	Moderate
Horse	Cow	Horse	Cow	Moderate
Milk	Cookies	Cookies	Milk	Easy

(*Appendix continues*)

Appendix (*continued*)

Orientation		Target		
Left	Right	Form 1	Form 2	Difficulty level
Table	Chair	Table	Chair	Moderate
Little	Big	Big	Little	Difficult
Eating	Throwing	Eating	Throwing	Easy
Scissors	Money	Scissors	Money	Difficult
Red	Blue	Blue	Red	Difficult
Truck	Airplane	Airplane	Truck	Moderate
Full	Empty	Full	Empty	Difficult
Bicycle	Train	Train	Bicycle	Moderate
Penguin	Giraffe	Penguin	Giraffe	Difficult
Cheese	Apple	Apple	Cheese	Easy
Smiling	Crying	Smiling	Crying	Moderate
Playing	Sleeping	Playing	Sleeping	Moderate
Bottle	Ball	Ball	Bottle	Easy
Reading	Washing	Reading	Washing	Moderate
Turtle	Butterfly	Butterfly	Turtle	Difficult
Touching	Riding	Touching	Riding	Moderate
Dirty	Clean	Clean	Dirty	Moderate
Cat	Duck	Duck	Cat	Easy
Pig	Fish	Pig	Fish	Moderate
Girl	Boy	Boy	Girl	Difficult

Note. For difficulty level, Easy = items comprehended by more than 66% of infants, Moderate = items comprehended by 33%–66% of infants, and Difficult = items comprehended by less than 33% of infants at 16 months of age (Dale & Fenson, 1996).

Received June 15, 2010
Revision received February 24, 2011
Accepted April 15, 2011 ■