J. Child Lang. **26** (1999), 177–185. Printed in the United Kingdom
© 1999 Cambridge University Press

NOTE

Comparing productive vocabulary measures from the CDI and a systematic diary study*

BYRON F. ROBINSON AND CAROLYN B. MERVIS
University of Louisville

(*Received* 22 *December* 1997. *Revised* 9 *July* 1998)

ABSTRACT

Expressive vocabulary data gathered during a systematic diary study of one male child's early language development are compared to data that would have resulted from longitudinal administration of the MacArthur Communicative Development Inventories spoken vocabulary checklist (CDI). Comparisons are made for (1) the number of words at monthly intervals (9;10.15 to 2;0.15), (2) proportion of words by lexical class (i.e. noun, predicate, closed class, 'other'), (3) growth curves. The CDI underestimates the number of words in the diary study, with the underestimation increasing as vocabulary size increases. The proportion of diary study words appearing on the CDI differed as a function of lexical class. Finally, despite the differences in vocabulary size, logistic curves proved to be the best fitting model to characterize vocabulary development as measured by both the diary study and the CDI. Implications for the longitudinal use of the CDI are discussed.

INTRODUCTION

The MacArthur Communicative Development Inventories (CDI; Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993) are instruments designed to assess the early language skills of children through parental report. The CDI can be completed in approximately 30 minutes and requires no special training to administer. There are two forms of the CDI. The CDI/Words and Gestures measures comprehension and production vo-

[*] We thank John Pani and Cindy Mervis for help with data collection. Cindy Mervis and Kathy Johnson performed much of the data reduction and organization of the original diary. Data collection was supported by National Science Foundation Grant BNS 84-19036. Analysis and preparation of this article were partially supported by National Institutes of Child Health and Human Development Grant HD 29957. Address for correspondence: Byron F. Robinson, Department of Psychology, Louisville, KY 40292, USA. e-mail: bfrobi01@homer.louisville.edu

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

cabulary and the use of gestures of individuals functioning in the 0;8 to 1;4 age range. The second form, the CDI/Words and Sentences, measures vocabulary production as well as some aspects of grammar and syntax for individuals functioning in the 1;4 to 2;6 age range. Research performed during the development of the CDI indicates high levels of reliability and relatively good concurrent validity for most of the measures (Dale, Bates, Reznick & Morisett, 1989; Dale, 1991; Fenson *et al.*, 1993; Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994). Given the high cost and time required to administer alternative methods of estimating language functioning (e.g. language sampling, standardized tests, diary studies), the CDI represents a breakthrough in instrument development for the assessment of language in large samples of infants and children (Tomasello & Mervis, 1994).

Of particular interest to our research into early lexical development is the expressive vocabulary inventory of the CDI. The CDI measures productive vocabulary via an extensive checklist of words commonly spoken by young children. From this list, the parent marks each of the words his or her child says. As would be expected from any reasonably sized checklist, not all words that all children produce could possibly be included. As stated in the MacArthur Communicative Development Inventories User's Guide and Technical Manual (Fenson *et al.*, 1993), CDI measures should not be viewed as a comprehensive listing of a child's vocabulary:

> ... we regard the vocabulary totals as an index of a given child's vocabulary knowledge, rather than an exhaustive atlas. An atlas of child vocabulary would require a list of extraordinary length. Although the present index might approach the status of an atlas for the younger children, it becomes an increasingly smaller subset of vocabulary for older children. (p. 40)

The authors of the CDI expect it to underestimate older children's vocabulary. It is currently unknown, however, by how much vocabulary is underestimated, and how the extent of underestimation is related to age. Further, given that different word classes comprise different proportions of the English lexicon, it may be the case that the proportion of words the CDI underestimates differs by lexical class.

Another issue of interest is the extension of the CDI to the longitudinal setting. In contrast to longitudinal data, cross-sectional studies of language development are sometimes prone to averaging effects in which interesting paths to development can be lost by lumping results from heterogeneous groups of children into single summary figures. This is especially true of growth curve studies in which different forms of growth, such as linear, logistic or stepwise are evidenced by individual children in the sample. For instance, when single time points from cross-sectional sampling are combined to form a group growth curve, the original heterogeneity in the sample may be obscured. In fact, the resulting combined growth curve may not represent

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

the curve form of any single child in the sample (Robinson & Mervis, 1998). When growth growth curves are of interest, diary studies are often the best method of collecting valid data at dense enough time points to allow valid analyses. However, diary studies are costly, time consuming, and usually include only a single child. If an efficient form of assessment, such as the CDI, could be used at each time point in a longitudinal study, the cost and time savings would be significant. As a result, longitudinal studies with large Ns would be feasible. To our knowledge, however, the CDI has not been validated for longitudinal use.

The present study compares longitudinal outcome measures from the CDI to data gathered during a systematic diary study. Although diary study data are also prone to errors of omission, the comprehensive nature of a well executed systematic diary study is more likely to approximate a child's actual vocabulary than a checklist or other form of language sampling. Therefore, comparisons of diary wordlists and data gathered through CDI methods provide an index of validity for the CDI. Using data generated from a systematic diary study of early lexical development of one male child, we examine three comparisons in particular: (1) What proportion of words in the diary study appear on the CDI checklists? (2) Do the proportions differ by age and/or word class? (3) Are data gathered using the CDI format at monthly intervals able to generate accurate vocabulary growth curves?

METHOD

Participant

The vocabulary productions of one male child (Ari) provided the raw data for this study. Ari was observed for most of his waking hours except for approximately 15 hours a week when he was in daycare. The portion of the diary reported here began when Ari was aged 9;10.15 and ended at 2;0.15.

Materials

Descriptions of Ari's vocal productions were recorded, at the time they were uttered, either in writing or on a handheld audiotape recorder by one of a number of trained caregivers. The trained caregivers included Ari's mother, father, and maternal aunt. All three observers had training in research methodology and two of the observers have had training in phonetics and phonology. The utterance descriptions were transcribed, coded and entered into a computer database. Almost complete records of all of Ari's productions were recorded until the age of 1;4. After that point, the observers recorded instances of new words, overextensions, underextensions, and words that appeared to be in transition. This procedure yielded more than 38925 entries containing the word produced, the phonetic form of the word, the context in which the word was used, and the adult response to the word. (See Mervis,

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

1987 and Mervis, Mervis, Johnson & Bertrand, 1992 for a more comprehensive description of the diary study method and coding system.)

Procedure

Cumulative monthly word counts were generated from the database of Ari's lexical productions. It was assumed that any word produced once remained in the vocabulary, a fact that the child's mother confirmed for virtually all cases. Each word was coded for lexical class: (1) noun, (2) predicate (verbs and words such as adjectives that can be used in a predicative function), (3) closed class, and (4) 'other' (e.g. animal sounds, games and routines, etc.) Starting with Ari's first word at 0;10.15, monthly counts of vocabulary size were calculated from the daily word totals by tallying counts from the 16th day of one month to the 15th of the next month. For each month, the CDI/Words and Sentences vocabulary checklist was scored based on the words appearing in the diary study.

RESULTS

The diary method of vocabulary size estimation yielded a final cumulative count of 1265 words at 2;0. Nouns comprised the bulk of Ari's vocabulary; with the number of nouns = 851, predicates = 302, closed class = 67, and other = 45 words. Of the words contained in the diary study, 543 also appeared on the CDI checklist. The CDI words included 306 nouns, 166 predicates, 53 closed class words, and 18 words falling in the 'other' category. Figure 1 presents Ari's vocabulary size at monthly intervals for the diary

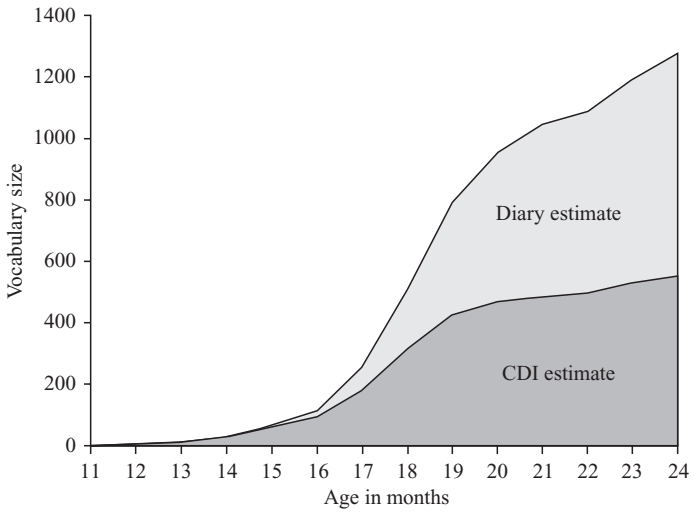


Fig. 1. Monthly diary study and CDI estimates of total vocabulary size.

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

study and the CDI estimates and Figure 2 presents the proportion of Ari's diary vocabulary broken down by word class. Ari's vocabulary size based on the CDI estimates placed him at or above the 90th percentile relative to the CDI norms for boys (Fenson *et al.*, 1993) from 1;5 to 2;0. His final vocabulary size of 543 words at 2;0 (percentile rank = 93%) is high, but within normal variation.

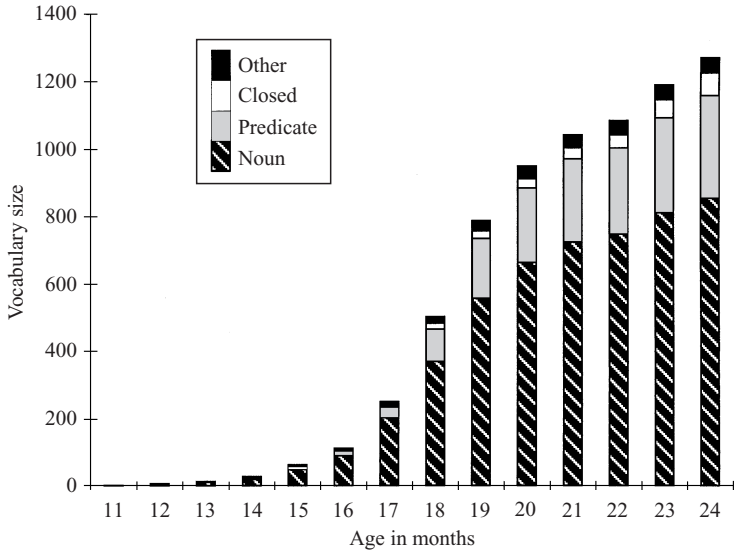


Fig. 2. Monthly diary study estimates of the proportion of words in each lexical class.

Until 1;4, the CDI included more than 80% of the words in the diary counts. At later ages, however, the proportion of diary words found on the CDI decreased to a minimum of 43% at 2;0. The drop in proportion of diary words captured by the CDI is most dramatic for nouns. At 2;0, the CDI included 79% of closed class words and 55% of predicates, but only 36% of nouns. The CDI captured 40% of the words in the smaller 'other' category. Figure 3 presents the proportion of diary words captured by the CDI for each lexial class.

Given the larger vocabulary estimated by the diary relative to the CDI method, the linear growth rates would be expected to differ. Analysis of the diary study data indicates that Ari learned 90·36 words per month, while learning rate calculated from the CDI data is only 38·79 words per month. Interestingly, however, the SHAPES of the growth curves generated by the diary and CDI methods are almost identical. Previous modelling of the diary data aggregated at the weekly level revealed the logistic as the best fitting curve for lexical growth up to 1;10 (Robinson & Mervis, 1998). The logistic

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

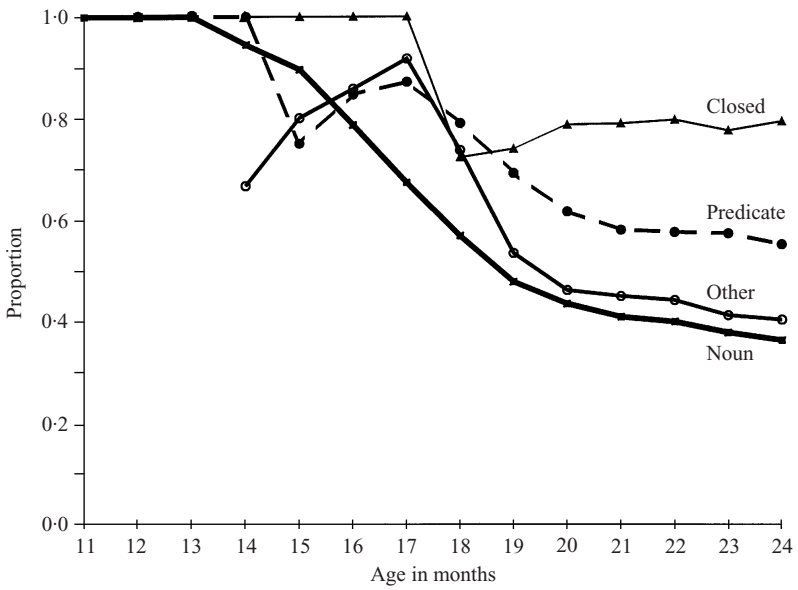


Fig. 3. Proportion of diary study words appearing on the CDI checklist by lexical class.

form proved also to be the best fitting curve for the diary data word counts for the same period when fit to the monthly estimates ($R^2 = 0{\cdot}995$):

$$Y_{diary} = \left(\frac{1}{1100} + 1{\cdot}76 * 0{\cdot}3896^{t}\right) - 1$$

with 1100 set as the upper bound based on a 1;10 vocabulary size of 1080 words and where t is the time point being estimated. As with the weekly and monthly diary data, the logistic curve was the best fitting model for the CDI word counts ($R^2 = 0{\cdot}994$):

$$Y_{diary} = \left(\frac{1}{500} + 1{\cdot}5 * 0{\cdot}4094^{t}\right) - 1,$$

where the upper bound is set to 500 based on a CDI 1;10 vocabulary size of 491 words. Figure 4 presents the fitted and actual curves for the diary and CDI data.

Another similarity between the diary and CDI data can be found by examining the lexical curve 1;10. As can be seen in Figure 1, there is a deviation from the tail plateau of the logistic that marks an increase in word learning rate from the 22nd month to the end of the study at 2;0. The jump in growth rate at this time point in development has been attributed to interaction between lexical growth and the acquisition of plural morphology

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

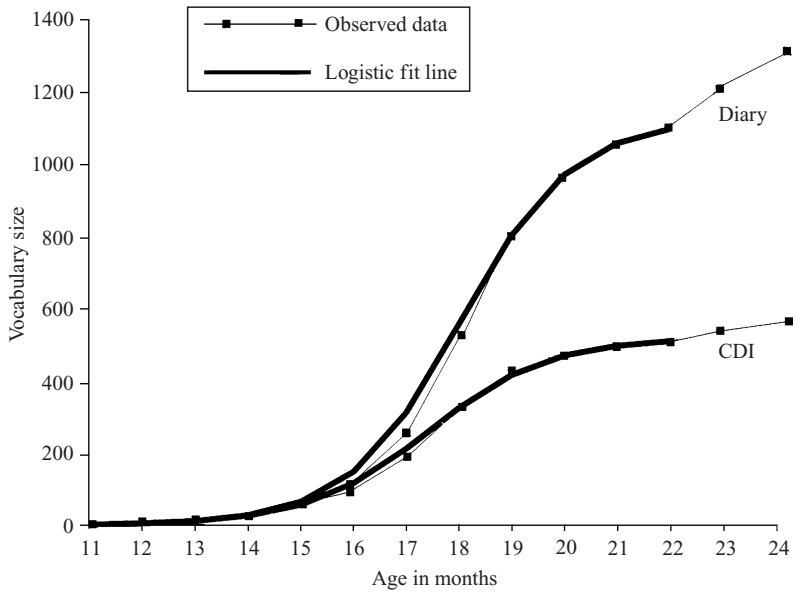


Fig. 4. Fitted logistic growth curves for the diary study and CDI vocabulary growth data.

(which Ari acquired before the onset of syntax). It is hypothesized that an initial slowing in vocabulary growth – the tapering at the end of the logistic curve – is caused by a competition for cognitive resources required for learning morphology. Once plural morphology is mastered, resources are freed and vocabulary growth rate increases once again – the point at which the curve deviates from the logistic form. For Ari, the increase in rate occurs at 1;10 (Robinson & Mervis, 1998). The deviation from the logistic curve, can be identified in both the diary study growth curve and, although more subtly, in the CDI curve.

DISCUSSION

As predicted by Fenson *et al*. (1993), the CDI did not capture all of the lexical items found in a comprehensive diary listing, especially of a child such as Ari, with a somewhat precocious vocabulary. The discrepancy between the larger diary study estimates and the CDI grew as vocabulary size increased with age. The CDI performs best for words in lexical classes that are more limited in diversity (e.g. the CDI captures a larger number of closed class words, which represent a smaller proportion of the English lexicon). Finally, the CDI method of data collection proved to be sufficient for growth curve modelling endeavours.

When comparing children to the CDI norms or other children's CDI scores, the discrepancy between CDI vocabulary size and actual vocabulary

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

size should not pose a problem. However, there is a caveat. The lack of uniformity in the proportion of words captured by the CDI across lexical classes may lead to the underestimation of some children's vocabulary knowledge. The uneven distribution across lexical class in some sense penalizes children who learn a higher proportion of nouns. For example, consider the case where child A has a larger vocabulary than child B, but most of Child A's vocabulary is in the form of nouns while child B's vocabulary is distributed across lexical classes in a more typical fashion. In this instance, it is possible for child A to be scored on the CDI as having the same, or even a smaller, vocabulary size than child B. This problem is likely more pronounced for older children whose vocabulary is large. In Ari's case, there is a dramatic drop in the proportion of words captured by the CDI at approximately 1;7 (see Figure 1), which is likely caused both by Ari's rapidly increasing vocabulary and the high proportion of nouns found in his lexicon. Although the problem with uneven distribution of missed words across word type is noted in the single case presented in this study, the actual impact of this problem for research remains to be quantified.

Despite the CDI's underestimation of diary vocabulary at any given time point, the growth curve of the CDI data mirrored the logistic form described by the diary curve. Given that the CDI and diary data were tallied at the monthly level, it is interesting to note that the curve generated from diary data aggregated at the weekly level also fell within the logistic curve family. Therefore, lexical acquisition data gathered via the CDI chacklists at monthly intervals should yield valid curve forms even though some parameters will differ (e.g. upper bound and growth rate) from true levels. Also, it is possible that deviations from the fitted growth curves will be captured by monthly CDI data, allowing the identification of some important transition points in language development.

Clearly the present comparison of the CDI method to data gathered using a systematic diary study does not represent a complete validation of the CDI for all longitudinal purposes. Not all children will necessarily follow the same lexical acquisition patterns as Ari. Other children with lower growth rates may present non-logistic forms of growth. The impact such variation in lexical learning has on the validity of the CDI relative to diary study methods should be an issue of further investigation. Additionally, the method of scoring the CDI used in this study assumed perfect memory on the part of the 'parent' filling out the checklist. That is, in the diary study, each word was recorded online as it was produced by the participant, therefore requiring no memory load on the observer. In an actual study using the CDI, it would be much more likely that a parent would be given the checklist and then asked to think back over the past month and remember any words that their child produced. A greater number of omission errors would be expected under these conditions.

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press

The present comparison of the CDI and the diary method does, however, increase confidence that the CDI administered at monthly intervals should provide a good index of a child's expressive vocabulary in the longitudinal context. Based on the results contained in this study one could expect word counts derived from monthly CDI administrations to generate adequate data for growth curve modelling. It should be noted, however, that not all applications of the CDI for longitudinal purposes are supported by this study. For instance, the CDI may not provide estimates valid for answering questions concerning relative proportions of vocabulary growth by lexical class. (See Pine, Lieven & Rowland, 1996 for additional discussion of this issue.) The use of the CDI as an instrument to track vocabulary makes large N longitudinal studies feasible, which should yield valuable information for the examination of individual differences in paths to early language development.

REFERENCES

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Sciences* **34**, 565–71.

Dale, P. S., Bates, E., Reznick, S. & Morisett, C. (1989). The validity of a parent report instrument at 20 months. *Journal of Child Language* **16**, 239–50.

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

Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (ed.), *Concepts and conceptual development* (201–33). Cambridge: C.U.P.

Mervis, C. B., Mervis, C. A., Johnson, K. E. & Bertrand, J. (1992). Studying early lexical development: the value of the systematic diary method. In C. Rovee-Collier & L. P. Lipsitt (eds), *Advances in infancy research*. Vol. 7 (291–378). Norwood, NJ: Ablex.

Pine, J. M., Lieven, E. V. M. & Rowland, C. (1996). Observational and checklist measures of vocabulary composition: what do they mean? *Journal of Child Language* **23**, 573–89.

Robinson, B. F. & Mervis, C. B. (1998). Disentangling early language development: modeling lexical and grammatical acquisition using an extension of case-study methodology. *Developmental Psychology* **34**, 363–75.

Tomasello, M. & Mervis, C. B. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development* **59** (5, Serial No. 242) 174–79.

https://doi.org/10.1017/S0305000998003663 Published online by Cambridge University Press