

Journal of Child Language

<http://journals.cambridge.org/JCL>

Additional services for *Journal of Child Language*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Observational and checklist measures of vocabulary composition: what do they mean?

Julian M. Pine, Elena V. M. Lieven and Caroline Rowland

Journal of Child Language / Volume 23 / Issue 03 / October 1996, pp 573 - 590
DOI: 10.1017/S0305000900008953, Published online: 26 September 2008

Link to this article: http://journals.cambridge.org/abstract_S0305000900008953

How to cite this article:

Julian M. Pine, Elena V. M. Lieven and Caroline Rowland (1996). Observational and checklist measures of vocabulary composition: what do they mean?. Journal of Child Language, 23, pp 573-590 doi:10.1017/S0305000900008953

Request Permissions : [Click here](#)

Observational and checklist measures of vocabulary composition: what do they mean?*

JULIAN M. PINE

University of Nottingham,

ELENA V. M. LIEVEN

University of Manchester

AND

CAROLINE ROWLAND

University of Nottingham

(Received 21 June 1995. Revised 28 November 1995)

ABSTRACT

Observational and checklist measures of vocabulary composition have both recently been used to look at the absolute proportion of nouns in children's early vocabularies. However, they have tended to generate rather different results. The present study is an attempt to investigate the relationship between such measures in a sample of 26 children between 1;1 and 2;1 at approximately 50 and 100 words. The results show that although observational and checklist measures are significantly correlated, there are also systematic quantitative differences between them which seem to reflect a combination of checklist, maternal-report and observational sampling biases. This suggests that, although both kinds of measure may represent good indices of differences in vocabulary size and composition across children and hence be useful as dependent variables in correlational research, neither may be ideal for estimating the absolute proportion of nouns in children's vocabularies. The implication is that questions which rely on information about the absolute proportion of particular kinds of words in children's vocabularies can only be properly addressed by detailed longitudinal studies

[*] We would like to thank all the families that took part in the research reported here. We would also like to thank two anonymous reviewers for their constructive comments on an earlier draft of this paper. This research was funded by the Economic and Social Research Council, Grant Number: R000234221. Address for correspondence: Julian M. Pine, Department of Psychology, University of Nottingham, NG7 2RD, UK. E-mail: jp@nott.psyc.ac.uk

in which an attempt is made to collect more comprehensive vocabulary records for individual children.

INTRODUCTION

In the twenty or so years since the publication of Nelson's (1973) monograph there has been a proliferation of studies on children's early lexical development which have been based on a variety of different kinds of data and have involved the use of a variety of different kinds of measures (see Lieven, Pine & Dresner Barnes, 1992 for a review). Against this background, the development of standardized vocabulary checklists, the most notable of which is the MacArthur Communicative Development Inventory (MCDI) (Bates, Bretherton & Snyder, 1988; Dale, Bates, Reznick & Morisset, 1989; Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994), has represented an important step forward for a number of reasons.

Firstly, it has allowed the construction of a relatively detailed picture of the course of early language development based on a very much larger number of children than would ever have been possible using maternal diary data alone (Bates *et al.*, 1994); secondly, it has provided researchers with a standardized measuring instrument which allows them to locate the children in their own, often relatively small samples with respect to the rest of the population (Bates *et al.*, 1994); and, thirdly, it has allowed the investigation of proposals about the relationship between stylistic variation and background variables such as SES, Sex and Birth order in samples which are large enough to allow meaningful conclusions to be drawn (Bates *et al.*, 1994).

However, although the development of vocabulary checklists such as the MCDI has undoubtedly opened up new possibilities in language development research, the ease with which such instruments allow researchers to collect data on children's early lexical development also carries with it some potential dangers, particularly if such data are used as substitutes for more fine-grained measures of children's early vocabulary and usage based on more labour-intensive longitudinal studies. It is thus important to be clear, firstly, about the precise nature of the relationship between checklist measures and other kinds of measures of vocabulary composition used in the literature, and, secondly, about the kind of information that checklist and other kinds of measures do and do not provide, and hence about the kind of questions they can reasonably be used to address.

Different kinds of measures of vocabulary composition

At least four different kinds of measures of vocabulary composition have been used in the literature on early lexical development. These are: measures derived from maternal diaries (e.g. Nelson, 1973; Pine & Lieven 1990); measures derived from maternal checklists (e.g. Bates *et al.*, 1988, 1994); observational measures (e.g. Tardif, 1996); and measures based on some

combination of maternal report and observational measures (e.g. Lieven *et al.*, 1992; Bloom, Tinker & Margulis, 1993; Choi & Gopnik, 1995).

It has tended to be assumed in the literature that these different kinds of measures are more or less equivalent. Moreover, if one is only interested in using them as measures of stylistic variation between children, this assumption is probably justified. Thus Pine (1992) has reported correlations of 0.66, $p < 0.05$ and 0.89, $p < 0.005$ between maternal diary and observational measures of referential style at 50 and 100 words, and Hampson (1989) has reported a correlation of 0.47, $p < 0.03$, between maternal checklist and observational measures of referential style at thirteen months, rising to 0.65, $p < 0.007$, when vocabulary size was controlled by restricting the analysis to the eighteen most advanced children in her sample.

However, different measures of vocabulary composition have also recently been used to support claims and counterclaims about the numerical importance of object words in children's early vocabularies both within and across languages (Gentner, 1982; Bloom *et al.*, 1993; Au, Dapretto & Song, 1994; Caselli *et al.*, 1995; Choi & Gopnik, 1995; Gopnik & Choi, 1995; Tardif, 1996). Such claims depend on information not about the relative proportion of particular kinds of words in different children's vocabularies, but about the absolute proportion of such words in the vocabularies of particular children or groups of children. However, the assumption that different measures of vocabulary composition are equivalent in this absolute sense is very different from the assumption that they are reliable in relative terms. That is to say, the fact that children who score high on checklist measures of referentiality also tend to score high on observational measures of referentiality is no guarantee that the absolute values of the checklist scores can be taken as valid indices of the absolute proportion of nouns in children's vocabularies.

In fact, there are a number of reasons for believing that there may be systematic quantitative differences between observational, diary and checklist measures of vocabulary composition, and that these may be responsible for some of the conflicting findings reported in the literature. Thus, as Bloom *et al.* (1993) point out, where studies have reported observational and maternal-report measures of the proportion of common nouns in children's vocabularies (whether from diaries or checklists), there has been a tendency for the observational measures to be somewhat lower than the maternal report measures. These differences have rarely been analysed directly. However, in one study where such an analysis was performed, Pine (1992) found a significant difference between maternal diary and observational measures of 'referentiality' at 50 and 100 words which seemed to be due partly to a noun bias in maternal reporting and partly to differences in the likelihood of nouns and other kinds of words occurring during the observational session.

This suggests that there are systematic quantitative differences between

maternal diary and observational measures of vocabulary composition which need to be taken into account when making claims about the absolute proportion of particular kinds of words in children's early vocabularies. However, it is unclear how checklist measures fit into this broader picture. For example, checklists such as the MCDI may eliminate maternal reporting biases by controlling the sample from which mothers are able to report vocabulary items. Alternatively, they may themselves incorporate a noun bias by sampling more exhaustively across the range of nouns in children's underlying vocabularies than they do across the range of other kinds of words (Nelson, Hampson & Kessler Shaw, 1993). This means that while it is possible that checklists may generate reasonably accurate measures of the absolute proportion of nouns in particular children's vocabularies, this can certainly not be taken for granted. Indeed, an alternative possibility is that they may result in measures which overestimate this figure even more seriously than do maternal diary measures.

The present study represents an attempt to address this question by examining not only the extent to which checklist measures of vocabulary composition correlate with observational measures of vocabulary composition, but also the extent to which these different measures result in quantitatively different vocabulary composition scores. The aim is to discover whether there are any systematic differences between observational estimates, checklist estimates and more comprehensive measures of vocabulary composition, particularly with respect to the important common noun category, and to investigate the reasons for any differences which do emerge.

METHOD

The study

The present study formed part of a more extensive longitudinal study of the early language development of 26 children whose parents had responded to an advertisement in the local newspaper. Children were recruited to the study on the basis that they had productive vocabularies of fewer than 10 words. Their progress was then monitored by telephone until the points at which their mothers reported vocabularies of approximately 10, 50 and 100 words. At each of these points the children were visited at home. During these visits mothers were asked to fill in a vocabulary checklist (see below), and a 40 minute audio-recording was made of their spontaneous language use, where the first 20 minutes was a recording of the child at lunch and the second 20 minutes a recording of the child at play. This procedure was intended, firstly, as a means of collecting vocabulary composition measures which were controlled as carefully as possible for vocabulary size at or around 50 and 100 words (Lieven & Pine, 1990; Pine & Lieven, 1990), and secondly as a means

of collecting observational vocabulary composition measures which were reasonably representative of the child's normal language use.

The checklist

The checklist used in this study was a version of the vocabulary checklist section of the MCDI (Infants) which had been modified in the following ways. Firstly, certain North American English vocabulary items (e.g. *diaper*) had been 'translated' into their British English equivalents (e.g. *nappy*). Secondly, since comprehension was not the focus of the present study, mothers were requested only to check off items which their child actually produced rather than items which she produced and/or comprehended. Thirdly, since we were interested in collecting direct measures of the proportion of unanalysed phrases as well as common nouns in children's early vocabularies, an additional multi-word utterance checklist was appended to the main single-word checklist, and mothers were requested only to check off items on the single-word checklist if they had heard their child produce them as one-word utterances. The multi-word checklist was made up of phrases which had often been reported by mothers in diaries in the authors' previous research on early vocabulary development (see Appendix A). Finally a section was added to the end of both the single-word and multi-word checklists in which mothers were asked to write down any vocabulary items which they had heard their child produce but which were not listed on the checklist. This was intended as a means of collecting more exhaustive maternal-report vocabularies which could be combined with additional vocabulary items from the observational data to provide vocabulary records which were as comprehensive as possible in the context of the present study.

Vocabulary composition measures

Data from the 50 and 100 word checklists and from transcriptions of the 50 and 100 word tapes were combined to produce a variety of different 50 and 100 word vocabulary records as follows:

Observed vocabularies

These vocabularies were based on separate inventories of types taken from the 50 and 100 word recordings, and were used to derive observational estimates of vocabulary composition.

Checklist vocabularies

These vocabulary records were based on the items checked by mothers on the single- and multi-word checklists, and were used to derive checklist estimates of vocabulary composition based entirely on maternal recognition.

Total maternal report vocabularies

These vocabulary records were based on a combination of all the vocabulary items checked by the mothers on the single- and multi-word checklists and all the items reported by the mothers in the two free recall sections of the checklist, and were used to derive maternal-report measures of vocabulary composition that were more comprehensive than the checklist measures.

Total vocabularies

These vocabularies were based on a combination of all the vocabulary items checked by the mothers on the single- and multi-word checklists; all the items reported by the mothers in the two free recall sections of the checklist; and any additional items found on the 50 and 100 word recordings which had not already been reported by the mothers. They were used to derive vocabulary composition measures which were as comprehensive as possible, within the limits of the present study.

Coding

The coding scheme used to classify the vocabulary data is present below (see Lieven *et al.*, 1992; Pine and Lieven, 1993 for more detailed discussions of this system).

The coding scheme

I. SINGLE-WORD UTTERANCES: utterances consisting of one word as defined by the adult language, including reduplications (e.g. *bye bye*, *night night*) and word-compounds (e.g. *Christmas tree*, *teddy bear*). These are subdivided as follows:

- A. COMMON NOUNS: words which refer to a class of possible referents in the adult language (e.g. *milk*, *dog*, *book*);
- B. ONOMATOPOEIC WORDS: words which represent the sound that specific objects make in the adult language (e.g. *brum*, *moo*, *meow*, *tick tock*);
- C. PROPER NOUNS: words which refer to a unique referent in the adult language (e.g. *Mummy*, *Grandma*, *Postman Pat*, *Father Christmas*);
- D. INTERACTIVE WORDS: words with little or no referential value in the adult language whose meaning is derived from the interactive situation in which they occur, such as greetings, or conversational devices, etc. (e.g. *bye bye*, *hooray*, *peepo*, *ta*);
- E. OTHER WORDS: all other single-word utterances, including verbs (e.g. *want*, *play*, *go*), adjectives (e.g. *more*, *pretty*, *blue*), adverbs/prepositions (e.g. *up*, *down*, *there*), pronouns (e.g. *me*, *mine*, *this*, *that*) and ambiguous words (e.g. *poo*, *wee wee*).

OBSERVATIONAL AND CHECKLIST MEASURES

2. MULTI-WORD UTTERANCES: utterances which consist of more than one word (as defined in the adult language) including forms which are frozen in adult speech or in the child's, but excluding word compounds and reduplications. These are subdivided as follows:

- A. FROZEN PHRASES: utterances which contain two or more words which have not previously occurred alone in the child's vocabulary or which contain one such word, provided it does not occur in the same position in a previous multi-word utterance;
- B. INTERMEDIATE: utterances in which:
 - (i) all the words or phrases have previously occurred independently in the child's vocabulary, provided none of them has occurred in the same position in two previous multi-word utterances,
and/or
 - (ii) one or more of the words or phrases has occurred in the same position in one but only one other previous multi-word utterance, provided that the word or phrase which makes up the remainder of the utterance has already occurred independently in the child's vocabulary.
- C. CONSTRUCTED: utterances which:
 - (i) contain one or more words or phrases which have occurred independently in the child's vocabulary, together with a word or phrase which has occurred in the same position in at least two other previous multi-word utterances.
or
 - (ii) conform to a positional pattern as in (Ci) above which is already established in the child's vocabulary, regardless of whether the variable word has occurred previously.

The most important features of this coding scheme, at least in the context of the present study, are, firstly, that it defines common nouns in an adult-centred rather than a child-centred way, and is hence comparable with most recent studies of early vocabulary composition; and, secondly, that it involves an explicit attempt to differentiate between unanalysed and productive multi-word utterances. This allows unanalysed phrases to be included as part of the child's 'single-word' vocabulary, and potentially productive utterances (i.e. intermediate and constructed utterances) to be excluded on the grounds that they result from the generative combination of items already present in the child's 'single-word' vocabulary.

All four of the different 50 and 100 word vocabularies for each child were coded according to this scheme and inter-rater reliability was assessed by having a second coder independently classify all the items in four children's total 100 word vocabularies. Reliability expressed as percentage agreement was a satisfactory 99.2% ($\kappa = 0.99$). Coded vocabulary records were then

used to derive scores for the proportion of common nouns and the proportion of frozen phrases in each of the children's four 50-word and four 100-word vocabularies.

Subjects

Subjects for the study were 26 children of mixed SES, including 14 girls and 12 boys and 12 first-borns and 14 later-borns, 50-word measures are reported for 24 children since two children (one girl and one boy) were found to have total maternal-report vocabularies which were closer to 100 than to 50 words at the 50-word recording session. 100-word measures are reported for 22 children since one child (a girl) had been lost from the study by this stage, and three children (one girl and two boys) were found to have total maternal-report vocabularies of more than 130 words at the 100-word recording session. The children ranged in age from 1;1 to 1;11 at 50 words and from 1;3 to 2;1 at 100 words (see Table 1). Their total maternal report

TABLE 1. *Subject details*

	Mean	S.D.	Range
Age at 50 word recording	17.6	2.8	1;1-1;11
Total maternal report vocabulary at 50 word recording	53.8	10.1	41-74
Total vocabulary at 50 word recording	64.0	13.1	46-92
Age at 100 word recording	20.1	2.9	1;3-2;1
Total maternal report vocabulary at 100 word recording	108.7	11.7	92-129
Total vocabulary at 100 word recording	132.2	17.1	102-166

vocabularies ranged in size from 41 to 74 words at 50 words and from 92 to 129 words at 100 words. These figures increased with the inclusion of additional vocabulary items from the 50 and 100 word recording sessions to give a range of 46 to 92 words at 50 words and 102 to 166 words at 100 words. However, there were no significant correlations between vocabulary size and either the proportion of common nouns or the proportion of frozen phrases at 50 or 100 words for either the total maternal report or the total vocabulary measures.

RESULTS

Descriptive statistics for the checklist and observational measures of the proportion of common nouns and frozen phrases at 50 and 100 words are presented in Table 2. These show substantial variation on all of the variables

OBSERVATIONAL AND CHECKLIST MEASURES

TABLE 2. *Descriptive statistics for checklist and observational measures of vocabulary composition at 50 and 100 words*

	Mean (%)	S.D.	Range
50 word measures			
Common nouns on checklist	51.0	15.0	20.6-85.0
Common nouns in observation	31.4	18.3	0-60.0
Frozen phrases on checklist	6.1	5.6	0-19.6
Frozen phrases in observation	10.8	11.0	0-35.7
100 word measures			
Common nouns on checklist	57.6	8.2	41.4-71.6
Common nouns in observation	36.9	13.4	11.1-66.7
Frozen phrases on checklist	5.1	3.5	0-12.8
Frozen phrases in observation	12.3	9.2	1.2-34.3

TABLE 3. *Correlations between checklist and observational measures of vocabulary composition at 50 and 100 words*

	<i>r</i>
50 word measures	
Common nouns	0.56**
Frozen phrases	0.53**
100 word measures	
Common nouns	0.64**
Frozen phrases	0.45*

* $p < 0.05$, two-tailed.** $p < 0.01$, two-tailed.TABLE 4. *T-test statistics for differences between checklist and observational measures at 50 and 100 words*

	<i>t</i>
50 word measures	
Common nouns	6.01***
Frozen phrases	2.46*
100 word measures	
Common nouns	9.48***
Frozen phrases	4.06**

* $p < 0.05$, two-tailed.** $p < 0.01$, two-tailed.*** $p < 0.001$, two-tailed.

at both measurement points. However, the most striking thing about these figures is probably the difference between the observational and checklist measures of the percentage of common nouns at both 50 and 100 words (around 20% in each case). These differences are even larger than those reported by Pine (1992) for observational and maternal diary measures (i.e. 9.7% at 50 words and 13.7% at 100 words), suggesting that, far from eliminating the noun bias in maternal reporting, checklist measures of vocabulary composition may actually accentuate this bias (see below).

Relationships between all the pairs of checklist and observational measures presented in Table 2 were investigated using Pearson product moment correlations and correlated *t*-tests. The results of the first of these analyses are presented in Table 3 and show substantial and significant correlations between the different types of measures for both the percentage of common nouns and the percentage of frozen phrases at both measurement points. This is encouraging for two reasons. Firstly, it provides further confirmation of the validity of using checklist estimates as measures of the relative numerical importance of common nouns in children's early vocabularies; and secondly, it suggests that, with some minor modifications, the checklist approach can be extended to provide more direct measures of the relative proportion of unanalysed phrases in children's early vocabularies. This is potentially quite important since the tendency to acquire unanalysed phrases early in vocabulary development has often been seen as a defining feature of a 'non-referential' style (Peters, 1977, 1983; Bretherton, McNew, Snyder & Bates, 1983; Bates *et al.*, 1988; Lieven *et al.*, 1992; Pine & Lieven, 1993). However, it is currently measured by the MCDI in only a very indirect way, (i.e. by calculating the percentage of closed class items reported by the mother on a single-word checklist).

The results of the second analysis are presented in Table 4. These show statistically significant differences between all the pairs of observational and checklist measures described above. Thus the checklist scores provide significantly higher estimates of the percentage of common nouns in children's vocabularies at both 50 and 100 words, significantly lower estimates of the percentage of frozen phrases at both 50 and 100 words, and significantly lower estimates of the percentage of frozen phrases at both measurement points. These findings underline the problems inherent in assuming that different measures of vocabulary composition will result in proportional scores which are more or less equivalent in absolute terms. They also demonstrate how different kinds of measures can give very different answers to questions about the numerical importance of nouns in children's early vocabularies. Thus, the checklist measures seem to indicate that for the majority of the children (13 out of 24 at 50 words and 19 out of 22 at 100 words), over 50% of the words in their vocabularies are common nouns. On the other hand, the observational measures paint a very different

picture, with only 5 out of 24 children showing referential vocabulary scores of over 50 % at 50 words and only 4 out of 22 children showing referential vocabulary scores of over 50 % at 100 words.

It may be tempting to interpret these differences solely in terms of a noun bias in maternal reporting, and hence as a reason for favouring observational measures over maternal report measures, at least as indices of the absolute proportion of nouns in children's vocabularies. However, there are a number of reasons why this conclusion should not be accepted too readily. Firstly, there is the fact that observational measures are generally much less comprehensive than maternal-report measures and are thus only very indirect measures of the proportion of words in the child's underlying vocabulary. As such, they may be inherently less reliable than maternal report measures which are typically based on a much larger number of vocabulary items (e.g. Bates *et al.*, 1988; Pine, 1992). They may also be overly sensitive to differences in the amount that particular children talk. Secondly, there is the possibility that since observational measures tend to be based on recordings made in a particular limited range of contexts, they may themselves be biased in systematic ways. Thus there may be certain kinds of words which are rarely or never observed simply because the situations in which they are typically used are rarely or never sampled. Thirdly, there is the possibility that differences in observational and maternal-report measures may reflect overall differences in the frequency with which different kinds of vocabulary items tend to be used by the child and hence in the likelihood of such words being observed during any particular recording session. Indeed Pine (1992) provides some evidence that this is in fact the case, at least at 100 words, though his analysis of maternal 'errors of omission' also suggests that, for continuous diary measures, a noun bias in maternal reporting is also a factor.

These considerations, when taken together, suggest that although maternal report measures may systematically overestimate the proportion of nouns in children's early vocabularies, there may also be a tendency for observational measures to systematically underestimate this figure. The implication is that neither kind of measure may be ideal as a means of estimating the absolute proportion of nouns in children's vocabularies, and that the best one may be able to do is to combine information from maternal report and observational measures in order to derive vocabulary measures which are as comprehensive as possible. This issue was investigated further by comparing the checklist and observational measures with two other measures of vocabulary composition, (i.e. the total maternal-report measures which were based on a combination of the recognition and recall sections of the checklist; and the total vocabulary measures which were based on a combination of the recognition and recall sections of the checklist and any additional vocabulary items taken from the 50 and 100 word tapes).

TABLE 5. *Descriptive statistics for maternal report and total vocabulary measures of vocabulary composition at 50 and 100 words*

	Mean (%)	S.D.	Range
50 word measures			
Common nouns in maternal report	46.8	14.1	18.2-82.9
Common nouns in total vocabulary	42.9	12.4	15.8-74.1
Frozen phrases in maternal report	7.0	5.5	0-18.8
Frozen phrases in total vocabulary	9.1	7.3	0-26.3
100 word measures			
Common nouns in maternal report	50.9	7.9	35.6-64.9
Common nouns in total vocabulary	46.1	7.3	32.0-60.1
Frozen phrases in maternal report	7.0	4.0	0-15.2
Frozen phrases in total vocabulary	10.0	5.1	0.7-17.7

TABLE 6. *T-test statistics for differences between checklist, maternal report, total vocabulary and observational measures of vocabulary composition at 50 and 100 words*

	<i>t</i>
50 word measures	
Common nouns (checklist vs. maternal report)	6.12***
Common nouns (maternal report vs. total)	5.65***
Common nouns (total vs. observational)	4.28***
Frozen phrases (checklist vs. maternal report)	2.06*
Frozen phrases (maternal report vs. total)	3.46**
Frozen phrases (total vs. observational)	1.25
100 word measures	
Common nouns (checklist vs. maternal report)	9.72***
Common nouns (maternal report vs. total)	8.92***
Common nouns (total vs. observational)	4.96***
Frozen phrases (checklist vs. maternal report)	3.97**
Frozen phrases (maternal report vs. total)	5.20***
Frozen phrases (total vs. observational)	1.66

* $p < 0.10$, two-tailed.** $p < 0.01$, two-tailed.*** $p < 0.001$, two-tailed.

Descriptive statistics for these measures are presented in Table 5. It can be seen from these figures that both sets of measures result in scores which are intermediate in size between the checklist and observational scores, and that, for both common nouns and frozen phrases, the total maternal report scores are more similar to the checklist scores and the total vocabulary scores more similar to the observational scores. Differences between the ordered means of each of the four types of scores for common nouns and frozen phrases at 50 and 100 words were therefore analysed using correlated *t*-tests. The reasoning

OBSERVATIONAL AND CHECKLIST MEASURES

was that differences between the checklist measures and the total maternal-report measures would be informative about biases in the checklist itself; that differences between the total maternal-report and the total vocabulary measures would be informative about biases in maternal reporting; and that differences between the total vocabulary measures and the observational measures would be informative about biases in observational sampling.

The results of these analyses for both common nouns and frozen phrases at 50 and 100 words are presented in Table 6. The pattern with respect to frozen phrases is fairly straightforward since there are significant differences between the total maternal-report and total vocabulary measures at both 50 and 100 words, a near-significant and a significant difference between the checklist and total maternal report measures at 50 and 100 words respectively; but no significant differences between the observational and total vocabulary measures at either 50 or 100 words. These results suggest that although there is a tendency for maternal-report measures to result in underestimates of the proportion of frozen phrases in children's early vocabularies (probably as a result of a combination of checklist and maternal reporting biases), there may be little to choose between observational and more comprehensive measures of vocabulary composition when it comes to estimating the absolute proportion of frozen phrases in children's early vocabularies.

The pattern with respect to common nouns is slightly more complicated since there are significant differences at all levels at both 50 and 100 words. Thus, the significant differences between the observational and total vocabulary measures suggest that observational measures may underestimate the proportion of nouns in children's underlying vocabularies, probably because of differences in the likelihood of common nouns and other kinds of words occurring during any particular observational session; the significant differences between the total maternal-report and the total vocabulary measures suggest a noun bias in maternal reporting, which presumably results from differences in the likelihood of mothers failing to recall nouns and failing to recall other kinds of vocabulary items; and the significant differences between the checklist and total maternal-report measures suggest a bias in the composition of the checklist itself which has the effect of accentuating the noun bias that already exists in maternal reporting. The implication is that there is no easy solution to the problem of estimating the absolute proportion of nouns in children's early vocabularies and that the only way of coming up with a reasonably accurate estimate may be to collect vocabularies which are as comprehensive as possible for individual children. On the other hand, it is also worth pointing out that the more comprehensive the vocabulary measures become in the present study, the more they seem to shift towards the observational scores. Given that even the total vocabulary scores used here are probably not based on complete records of the children's underlying vocabularies, this suggests that even these measures may

overestimate the proportion of nouns in children's vocabularies to some degree, and that the true value of this figure may be somewhere between the total vocabulary scores and the observational scores. Whatever the case, it is clear that it is significantly lower than that suggested by both kinds of maternal report estimates, a finding which obviously needs to be taken into consideration when making claims about the absolute proportion of nouns in children's early vocabularies.

To summarize, the results reported here suggest that although checklist and observational measures of vocabulary composition are significantly correlated and hence may represent equally good indices of differences in vocabulary size and composition across children, there are also systematic quantitative differences between them which need to be taken into account when making claims about the absolute proportion of particular kinds of words in children's early vocabularies. These differences seem to be attributable to a variety of different factors including biases in the checklist itself, biases in maternal reporting, and biases in observational sampling, though these have different effects on different variables. The implication is that neither kind of measure is ideal as a basis for estimating the absolute proportion of particular kinds of words in children's vocabularies, and that the best one may be able to do is to base such estimates on more comprehensive measures of vocabulary composition derived from vocabulary records in which maternal-report and observational data are combined.

DISCUSSION

The aim of the present study was to investigate the relationship between observational and checklist measures of vocabulary composition by examining not only the extent to which such measures correlate with each other, but also the extent to which they result in quantitatively different vocabulary composition scores. The results show that although observational and checklist measures are significantly correlated, there are also systematic quantitative differences between them which seem to reflect a combination of checklist, maternal-report and observational sampling biases. This suggests that, although both kinds of measure may represent good indices of differences in vocabulary size and composition across children and hence be useful as dependent variables in correlational research, neither may be ideal for estimating the absolute proportion of nouns in children's vocabularies.

These results have a number of implications for research on children's early lexical development. Firstly, they provide further confirmation of the validity of checklist measures of vocabulary composition as indices of stylistic variation between children, and show how, with some minor modifications, checklist instruments can be adapted to provide more direct measures of the proportion of unanalysed phrases in children's early vocabularies. Given that the tendency to acquire unanalysed phrases early in vocabulary development

has often been seen as a defining feature of a 'non-referential' style (Peters, 1977, 1983; Bretherton *et al.*, 1983; Bates *et al.*, 1988; Lieven *et al.*, 1992; Pine & Lieven, 1993), this would seem to represent an important extension of the checklist methodology.

Secondly, they go some way towards explaining recent conflicting findings in the literature with respect to the numerical importance of nouns in children's vocabularies in both English (Lieven *et al.*, 1992; Bloom *et al.*, 1993; Bates *et al.*, 1994) and Korean (Au *et al.*, 1994; Choi & Gopnik, 1995; Gopnik & Choi, 1995). Thus the relatively low referential vocabulary scores reported by Lieven *et al.* (1992) and Bloom *et al.* (1993) could be explained, at least in part, by the fact that they are based on a combination of maternal diary and observational measures, whereas the relatively high referential vocabulary scores reported by Bates *et al.* (1994) are based solely on checklist data. Similarly, Choi & Gopnik's (1995) relatively low referential vocabulary scores could be explained by the fact that they are based on a combination of maternal report and observational data, whereas Au *et al.*'s (1994) relatively high referential vocabulary scores are based solely on checklist data from a Korean translation of the MCDI.

Thirdly, and most importantly, they suggest that if researchers are really interested in arriving at accurate estimates of the absolute proportion of particular kinds of words in children's early vocabularies, they must be prepared to collect more comprehensive records of particular children's vocabularies, rather than relying on less labour-intensive methods such as the use of maternal checklists or 'one-shot' observational measures. What this conclusion reflects is the fact that the real problem with using observational or checklist measures to answer questions about the absolute proportion of particular kinds of words in children's early vocabularies is not the possibility that such measures might be unreliable, nor even the possibility that they might be biased in systematic ways, but rather the fact that since they are not even supposed to be based on complete vocabulary records for individual children, they are simply not appropriate for this kind of task.

In fact, observational measures provide only very indirect measures of the proportion of particular kinds of words in the children's underlying vocabulary and are probably best seen as measures of usage rather than of vocabulary composition *per se*. On the other hand, checklist measures are based on standardized instruments which have been designed to measure individual differences between children by selectively sampling across the words in their underlying vocabularies. The scores which they generate are therefore only really meaningful in an individual differences context (i.e. as indices of differences in vocabulary size and composition across children). This should make one suspicious not only about conclusions based on the absolute value of the scores themselves, but also about conclusions based on absolute differences between the scores obtained for different categories of

words, (e.g. nouns and verbs); or on absolute differences between the scores obtained for the same category in different languages, (e.g. English and Korean, or English and Italian). This is because it is unclear how the scores for different languages or different categories of word map onto the child's underlying vocabulary knowledge. We therefore have no way of knowing whether a score of 20% for nouns is equivalent to a score of 20% for verbs, or that a noun score of 50% in English is equivalent to a noun score of 50% in Korean.

The implication is that questions about the absolute proportion of particular kinds of words in children's vocabularies can only be properly addressed by detailed longitudinal studies in which an attempt is made to collect comprehensive vocabulary records for a reasonably large number of children. Of course, such studies are relatively rare in the literature. However, they are not non-existent. Thus Lieven *et al.*'s (1992) and Bloom *et al.*'s (1993) studies of English-speaking children and Choi & Gopnik's (1995) study of Korean- and English-speaking children all use measures based on a combination of maternal report and observational data collected within a longitudinal framework. The results of the present study suggest that it is likely to be these studies that provide the best currently available estimates of the absolute proportion of words in English-speaking and Korean-speaking children's early vocabularies since it is these studies that are based on the most comprehensive vocabulary records. Moreover, it is worth noting that, in contrast to the results of several recent checklist studies (e.g. Au *et al.*, 1993; Bates *et al.*, 1994; Caselli *et al.*, 1995), the figures reported in these studies not only suggest that there are important cross-linguistic differences in the kind of words which children learn early in development (Choi & Gopnik, 1995), but also that even among English-speaking children, subjects with a majority of common nouns or object names in their first 50 words may be relatively rare (Lieven *et al.*, 1992; Bloom *et al.*, 1993).

REFERENCES

- Au, T. K., Dapretto, M. & Song, Y. K. (1994). Input versus constraints: early word acquisition in Korean and English. *Journal of Memory and Languages* 33, 567-82.
- Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanism*. Cambridge: Cambridge University Press.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J. & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* 21, 85-123.
- Bloom, L., Tinker, E. & Margulis, C. (1993). The words children learn: evidence against a noun bias in early vocabularies. *Cognitive Development* 8, 431-50.
- Bretherton, I., McNew, S., Snyder, L. & Bates, E. (1983). Individual differences at twenty months: analytic and holistic strategies in language acquisition. *Journal of Child Language* 10, 293-320.
- Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L. & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development* 10, 159-99.

OBSERVATIONAL AND CHECKLIST MEASURES

- Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: a cross-linguistic study. *Journal of Child Language* **22**, 1-33.
- Dale, P., Bates, E., Reznick, J. & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language* **16**, 239-49.
- 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).
- Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. A. Kuczaj (ed.), *Language development, II: Language, thought and culture*. Hillsdale, NJ; Erlbaum.
- Gopnik, A. & Choi, S. (1995). Names, relational words, and cognitive development in English and Korean speakers: nouns are not always learned before verbs. In M. Tomasello & W. E. Merriman (eds), *Beyond names for things: young children's acquisition of verbs*. Hillsdale, NJ: Erlbaum.
- Hampson, J. (1989). Elements of style: maternal and child contributions to referential and expressive styles of language acquisition. Unpublished doctoral dissertation. City University of New York.
- Lieven, E. V. M. & Pine, J. M. (1990). Review of 'From first words to grammar: individual differences and dissociable mechanisms' by E. Bates, I. Bretherton and L. Snyder. *Journal of Child Language* **17**, 495-501.
- Lieven, E. V. M., Pine, J. M. & Dresner-Barnes, H. (1992). Individual differences in early vocabulary development: redefining the referential-expressive distinction. *Journal of Child Language* **19**, 287-310.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development* **38** (1-2).
- Nelson, K., Hampson, J. & Kessler Shaw L. (1993). Nouns in early lexicons: evidence, explanations and implications. *Journal of Child Language* **20**, 61-84.
- Peters, A. M. (1977). Language learning strategies: does the whole equal the sum of the parts? *Language* **53**, 560-73.
- (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.
- Pine, J. M. (1992). How referential are 'referential' children? Relationships between maternal-report and observational measures of vocabulary composition and usage. *Journal of Child Language* **19**, 75-86.
- Pine, J. M. & Lieven, E. V. M. (1990). Referential style at thirteen months: why age-defined cross-sectional measures are inappropriate for the study of strategy differences in early language development. *Journal of Child Language* **17**, 625-31.
- & — (1993). Reanalysing rote-learned phrases: individual differences in the transition to multi-word speech. *Journal of Child Language* **20**, 551-71.
- Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology* **32**, 492-504.

APPENDIX A: MULTI-WORD CHECKLIST

The following is a list of phrases which typically appear in young children's vocabularies. Could you please circle any phrase which you have heard your child say. If your child uses a different pronunciation of a phrase (e.g. 'Ee are' for 'Here you are' or 'Der tis' for 'There it is') circle the phrase anyway. Remember, this is a 'catalogue' of phrases that are used by many different children. Don't worry if your child is saying only a few at the moment.

A drink	In a minute	Sit down
All gone	In there	Stop it
Another one	It's a car	That one
Catch it	I want it	Thank you
Come on	Kick it	That's mine
Daddy gone	Light on	That's nice
Do it	Like it	There it is
Don't do that	More juice	This one
Dropped it	My car	There you are
Get down	No more	Up stairs
Go away	Oh dear	Want it
Good boy	One two three	What is it?
Good girl	On there	What's that?
Here it is	Open it	Where is it?
Here you are	See it	Yes please
Mummy do it	See you	
I get it	Shoes on	

Could you now write down in the space provided any other things which your child says which are more than one word long. Please include here all the phrases you can think of, even if they are very similar to each other (e.g. 'Its a cat', 'It's a dog', 'It's a lorry') or are similar to phrases which you have circled above (e.g. 'More juice', 'More milk', 'More cheese', etc.)

1	11	21
2	12	22
3	13	23
4	14	24
5	15	25
6	16	26
7	17	27
8	18	28
9	19	29
10	20	30