

Vocabulary development in Italian children: a longitudinal evaluation of quantitative and qualitative aspects*

LAURA D'ODORICO, STEFANIA CARUBBI,
NICOLETTA SALERNI AND VINCENZO CALVO
University of Padova, Italy

(Received 17 March 1999. Revised 28 June 2000)

ABSTRACT

In this study vocabulary development of a sample of 42 Italian children was evaluated through monthly administration of the Italian version of the CDI. Data collection started at 1;0–1;1 for 32 children and at 1;3–1;4 for the remaining subjects and continued until children's vocabulary reached 200 words. At fixed stages of vocabulary size (50, 100 and 200 words), individual differences in percentile scores and vocabulary composition were examined. Individual growth curves were analysed in order to verify the presence of a vocabulary spurt and the type of lexical items which contributed most to rapid acceleration in vocabulary growth.

Stylistic differences in vocabulary composition were examined regarding the 'referential–expressive' distinction, controlling vocabulary size. Data have shown that general trends in vocabulary development are quite similar to those obtained for other languages using CDI adaptations. Moreover, all children in this sample eventually exhibited a vocabulary spurt, even if some can be defined as 'late spurters'. The type of lexical items which are learned during the spurt depend on both infant vocabulary size and referential score. About 28% of infants in this sample were defined 'referential' when their vocabulary size was about 50 words, but the stylistic differences disappeared at the 100- and 200-word stages.

Composition of vocabulary did not differ in relation to precocity in reaching different stages of vocabulary development. The only exception

[*] This research was supported by grant CNR96.05216.PSo8 awarded to the first author. Address for correspondence: Prof. Laura D'Odorico, Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università di Padova, Via Venezia 8, 35131 Padova, Italy.

was that infants who reached the 50-word stage first also had a vocabulary with a lower proportion of function words.

INTRODUCTION

Vocabulary development in the second year of life is certainly one of the most impressive phenomena of the language acquisition process for both parents and language students. Therefore the number of studies devoted to identifying the characteristics of this development from both a quantitative and a qualitative point of view is not surprising. Most use parental reports which have proven to be a reliable tool for assessing vocabulary development (Bates, Bretherton & Snyder, 1988; Dale, Bates, Reznick & Morisset, 1989; Reznick & Goldfield, 1994; Bornstein & Haynes, 1998). One widely used instrument is the MacArthur Communicative Development Inventory (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) designed for English-speaking children. Recently, however, the CDI has been adapted for different languages allowing a cross-linguistic comparison in early lexical development, observed with the same method of assessment.

Regarding general quantitative aspects of vocabulary development, Italian-speaking children (Caselli, Bates, Casadio, Fenson, Fenson, Sanderl & Weir, 1995) appear to learn new words at a slower rate than English-speaking children. The same result has been found for Icelandic-speaking children (Thordardottir & Weismer, 1996), while no significant difference has been detected for Hebrew, Spanish and Finnish-speaking children (Jackson-Maldonado, Thal, Marchman, Bates & Gutierrez-Clellen, 1993; Lyytinen, Lari, Lausvaara & Poikkeus, 1994; Maital, Dromi, Sagi & Bornstein, 1998). In any language, however, large individual differences are observed in both development rate and total vocabulary size at a given age. Socio-demographic factors can only partially explain the great variability in vocabulary development. A female advantage has been observed in English (Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Fenson *et al.*, 1994; Bornstein & Haynes, 1998), and Hebrew (Maital *et al.*, 1998), but not in Swedish (Berglund & Eriksson, 1998). Mothers' educational level has often been associated with faster vocabulary development, while data about the influence of birth order, which up to now has been collected for English-speaking children only, are not concordant (Jones & Adamson, 1987; Oshima-Takane & Derevensky, 1990; Lieven, Pine & Dresner-Barnes, 1992).

The great variability observed in vocabulary growth rate may also depend on individual differences in the same process of vocabulary acquisition. A good candidate for investigation in this area is the 'VOCABULARY SPURT' phenomenon. A 'vocabulary spurt' has been defined as a marked increase in the rate at which new words are added to a child's productive vocabulary. It represents the change from what is, initially, a slow acquisition of first words,

to a more rapid rate of word learning. A threshold of 50 words is usually considered necessary in order to observe a vocabulary spurt and it is common between 1;2–2;0 (Bloom, 1973; Nelson, 1973; Benedict, 1979; Dromi, 1987; Lucariello, 1987). Although this phenomenon has been observed in many longitudinal studies of English-speaking children, many aspects are still questionable, first of all its universality and relevance to the process of language acquisition. Nelson (1973) suggested that not all children demonstrate a period of rapid acceleration. More recently Goldfield & Reznick (1990) reported that 5 out of 18 children in their sample, who were followed from 1;2 to 1;10, learned their first 75 to 100 words at a gradual rate. However, Mervis & Bertrand (1995) showed that some children, who can be labelled 'late spurters', may demonstrate a vocabulary spurt even after their productive vocabulary has reached 100 words, concluding that 'all normally developing children will eventually evidence a vocabulary spurt'.

Moreover, Goldfield & Reznick (1990) and Mervis & Bertrand (1995) observed a marked increase in the average rate of new word acquisition after the first recorded spurt interval, while Dromi (1987) and Gopnik & Meltzoff (1987) reported cases of level-off and even decline.

Assessment of the generality of the vocabulary spurt, both intra-language and across languages, and the course of development afterwards, is relevant because hypotheses relating this phenomenon to the development of memory (Huttenlocher, 1974), language perception (Plunkett, 1993), conceptual development (Bloom, Lifter & Broughton, 1985; Nelson & Lucariello, 1985; Gopnik & Meltzoff, 1987) or constraint on word learning (Markman, 1991) are admissible only if the vocabulary spurt, as epiphenomenon of one of these capacities, is observable in all children, across different native languages, even if at different ages.

Cross-linguistic evidence of the vocabulary spurt in languages other than English is scarce. For example, Berglund & Eriksson (1998) in their study based on the Swedish Early Communicative Inventory, reported that when the vocabulary score was about 50 to 100 words, the growth curve for all percentiles became steeper, indicating a vocabulary spurt, but their cross-sectional sample does not provide information about the generality of the phenomenon.

To obtain information about mechanisms operating in the vocabulary development process, analyses of vocabulary development rate have to be integrated with a more qualitative approach, focusing on the types of words infants actually learn to produce in the first stage of language acquisition, i.e. the composition of vocabulary. In this research area, cross-linguistic studies are particularly important to evaluate the influence of cognitive processes on lexical acquisition. For English-speaking children, nouns form the largest category in the 50–500 word interval, while predicates emerge later than nouns, followed by closed class items (Bates, Marchman, Thal, Fenson,

Dale, Reznick, Reilly & Hartung, 1994). Results of many cross-linguistic studies, all based on adaptations of CDI, (for Italian, Caselli *et al.*, 1995; for Hebrew, Maital *et al.*, 1998; for Spanish, Jackson-Maldonado *et al.*, 1993; for Finnish, Lyytinen *et al.*, 1994; for French, Poulin-Dubois, Graham & Sippola, 1995; for Japanese, Ogura, 1998; for German, Grimm, 1999) presented data confirming the noun–verb sequences. The universality of the predominance of nouns over verbs in early vocabulary development has been challenged on the basis of data regarding the acquisition of Korean (Choi & Gopnik, 1995) and Chinese (Tardif, 1996; Tardif, Gelman & Xu, 1999). The Italian language appears to be a good candidate for testing the hypothesis of the influence of input language on early vocabulary composition in so far as it has structural characteristics (it allows extensive word order variation, the subject is often omitted and verbs are often located in sentence-initial or sentence-final positions; cf. Caselli, *et al.* 1995) which make it quite different from English and more similar to Korean and Chinese, in some aspects. Results of studies examining the actual input of Italian mothers to their children are controversial. On the basis of an analysis of the Calambrone corpus in MacWhinney (2000) (six caregivers' language was transcribed during free interaction with their children who ranged in age from 1;9.1 to 2;2.1), Tardif, Shatz & Naigles (1997) claimed a similarity between English and Italian input (i.e. noun emphasis) which could explain a similarity in the composition of Italian and English children's early vocabulary in spite of structural differences between the two languages. On the other hand, Camaioni & Longobardi (2000) found verb rather than noun emphasis in Italian mother-to-child speech. Although there are methodological differences between the two studies, which could explain the differences in results (the sample size is larger – 15 mothers – and both the mean age and mean MLU are lower than in Tardif *et al.*'s study), the question of the universality of noun bias in early lexical development is still controversial and could benefit from other data such as those which will be analysed in this paper.

Data from cross-linguistic studies could be relevant to the question of the possibility of applying to other languages individual differences in vocabulary acquisition 'style', which have mostly been observed in English-speaking children.

The most widely used dimension for studying differences in style of vocabulary development is the referential-expressive distinction proposed by Nelson (1973). In the original study 'referential children' were defined as those for whom over 50% of their first 50 words were object names, while personal-social routines and formulas made up a large proportion of the first 50 words in children labelled 'expressive'. In her study, Nelson did not find any differences between the referential and expressive groups regarding the age at which they reached 50 words. In the following years, however, other

studies applied this distinction to describe stylistic differences in vocabulary development. The significant advantage for children who adopt a referential approach to language, evident when the comparison between children is based on age (Bates *et al.* 1988), disappears when the comparison is based on fixed vocabulary size (Pine & Lieven, 1990; Bates *et al.* 1994). More recently, Lieven *et al.* (1992) have redefined the referential/expressive distinction suggesting that the proportion of frozen phrases acquired by a child in the first 50 and 100 words is a good candidate for defining nonreferential style in a positive way.

There is not much evidence on the referential-expressive distinction across different languages. Zurer Pearson & Fernández (1994) in a study of lexical development of 20 bilingual (English-Spanish) infants, found that some children had referential vocabularies in both English and Spanish, some had expressive vocabularies in both, and others had referential vocabulary in English and expressive in Spanish or vice versa. These last two possibilities are more frequent when the vocabulary size in the two languages is very different (i.e. there is a dominant language). The only study which has attempted to verify the applicability of the referential-expressive distinction to Italian is Camaioni & Longobardi's (1995) which found only 1 child who could be categorized as 'referential' in a sample of 12 Italian infants.

Distinct from the problem of noun bias in first words, is the question of noun bias in the vocabulary spurt, i.e. whether the vocabulary spurt can be attributed to an acceleration in acquisition of nouns or whether other lexical items may be responsible for this sharp increment in vocabulary size. The types of items (nouns or verbs) which are learned during the spurt are important for verifying the different hypotheses described above and for relating individual variation in vocabulary growth to a faster or slower development of the capacity underlying acquisition of new lexical items. Research in which the vocabulary spurt phenomenon has been examined by simultaneously controlling lexical composition has obtained different results. For example, Gopnik & Meltzoff (1987) and Goldfield & Reznick (1990) maintained that the vocabulary spurt may be more aptly called the 'naming explosion', because nouns constitute the majority of the new vocabulary items. On the other hand, Choi & Gopnick (1995), reported that the majority of their Korean-speaking children had a period of rapid vocabulary growth that could be best classified as a 'verbing spurt'. Moreover, for those children who showed both a verb and a noun spurt, the rapid increase in verbs actually preceded the rapid increase in nouns.

In this study we will investigate some aspects of early lexical development discussed above, observing longitudinally the vocabulary development of a sample of Italian children. More specifically we will investigate individual differences in rate of vocabulary development and their stability in the acquisition of the first 200 words. The longitudinal approach adopted in this

study also allows verification of the generality of phenomenon of the vocabulary spurt in Italian children, the type of lexical items which contribute more to the rapid acceleration of vocabulary and the characteristics of vocabulary development before and after the spurt. The composition of vocabulary at different stages of vocabulary development was examined to verify noun predominance in the early phase of acquisition of Italian and the presence of stylistic variations in the acquisition of lexical items.

Quantitative and qualitative aspects of vocabulary development were also compared to verify if children more precocious in development rate also differed from slower children in vocabulary composition. From a theoretical point of view, many possibilities could have plausible foundations. We can hypothesize that (1) if the order of entry of the different lexical items is strongly related to vocabulary size, as data about English suggest, it does not matter at what age a fixed stage (e.g. 50 words) is reached, no differences should emerge between the vocabularies of early and late talkers; (2) early talkers reach the different stages of vocabulary size first because they have learned more of the simplest items (i.e. routines, onomatopoeic words, names for people); (3) precocity in vocabulary growth is related to a more advanced cognitive capacity which also produces the precocious entry of the more advanced lexical items (verbs and other predicates and closed class items). Data from literature are scarce regarding these hypotheses and difficult to compare because of substantial differences in methodology. Lieven *et al.* (1992) did not find any significant relations between the ages at which 50 and 100 words were reached by their longitudinal sample and the proportion of different lexical items. Camaioni & Longobardi (1995) found, on the other hand, that at 16 months of age earlier talkers produced more common nouns, verbs and articles in their actual speech than later talkers. At 20 months the difference is in favour of early talkers regarding verbs and adjectives, while late talkers produced many more onomatopoeic words. These data seem to confirm hypothesis 3, even if they must be considered with caution because the comparison between the two groups is based on age, not vocabulary size.

METHOD

Participants

Subjects for the present study were 42 Italian children (22 males and 20 females) whose families agreed to participate with their children in a longitudinal study of language development. The families were recruited from local birth records. The total number of families contacted by phone was 60; 46 families gave their consent, but 4 families left the project after a few months. The children and their families were all residents of cities in the northern regions of Italy. The education of mothers and fathers whose

children participated in the study was distributed fairly evenly across three different levels: level 1 – 5–8 years of education, corresponding to elementary and junior high school; level 2 – 13 years of education, corresponding to high school; level 3 – 18 years of education, corresponding to graduate school.

In this sample, only 5 children attended a daycare centre for some hours during the period of the study. However, another 25 children spent part of the day with a babysitter or grandparents because their mothers were working away from home. Complete demographic information of the children and their parents is given in the Appendix.

Design

After having obtained the consent of parents to participate in the project, a trained interviewer visited the families at home and administered the Italian version of the MacArthur Communicative Development Inventory (Caselli & Casadio, 1995) to mothers. The first questionnaire was administered at an age which ranged from 12 to 16 months (see Appendix). Each month the administration was repeated. An effort was made to visit participants within 1 week of the child's monthly birthday. Due to family holidays we missed one monthly questionnaire for 3 children, and two monthly questionnaires for another child.

Data collection was planned to continue up to the months when children's vocabulary size reached 200 words. However, for 5 children (see Appendix) vocabulary development was only monitored up to 24 months of age, when they had not yet reached that number of words.

The Italian version is modelled as closely as possible on the English version in terms of overall format, number and type of lexical categories and number of items. The CDI Infant form (Words and Gestures) was administered to infants up to 1;5, while the Toddler form (Words and Sentences) was administered from 1;6 to the end of the study.

The Infant form consists of a vocabulary list, for which both comprehension and production is assessed, and a section (Action and Gesture) in which nonverbal communicative and representational skills are assessed. The Italian Toddler form contains a vocabulary production checklist of 670 words and another two sections designed to assess morphological and syntactical production. In this work only data regarding vocabulary production are analysed.

Data reduction

The total number of words produced by each child, according to the mother's report was calculated for each monthly assessment. The criterion of the first increase of more than 20 new words in any monthly interval was

adopted in this study as indicating the phenomenon of vocabulary spurt. The number of 20 words was selected considering that (a) in previous longitudinal studies with an interval of two or three weeks, the vocabulary spurt has been considered as the first session in which more than 10 or 12 new words were acquired (Bloom & Capatides, 1987; Reznick & Goldfield, 1992) and (b) in the study by Poulin-Dubois *et al.* (1995), which used a longer interval, i.e. four weeks, the criterion of 15 new general nominals was adopted.

At fixed stages of vocabulary size, i.e. 50, 100 and 200 words, we calculated the percentile ranks for each subject, utilising normative data on the Italian population collected by Caselli & Casadio (1995). In order to analyse the composition of vocabulary, by considering the effect of vocabulary size, we used the nearest questionnaire to the fixed size for each stage. For the 50-word stage the actual range of vocabulary size was 40–64 (mean number of words = 52); for the 100-word stage it was 88–148 (mean number of words = 109); for the 200-word stage it was 186–303 (mean number of words = 223). We missed the 100-word stage analysis for 5 subjects (all males): one of them jumped from 46 to 233 words in one month, while the other four left the project at 2;0, when they had not yet reached the 100-word stage. We missed the 200-word stage analysis for these last three subjects and for another subject (male) who left the project at 2;0 months with a vocabulary size of about 100 words. Therefore we have complete data (50, 100 and 200-word stages) for 36 subjects (16 males and 20 females), while for one subject only data to compare the 50- and 100-word stages are available and for another only a comparison between the 50- and 200-word stages is possible.

The composition of children's vocabularies at the three points of development was analysed following the procedures outlined in Caselli *et al.* (1995). More specifically we calculated:

1. Percentage of common nouns, which include only words from adult language that stand for concrete objects (animal names, vehicles, toys, food and drink, clothing, body parts, furniture and rooms, small household objects). Names of people, sound effects and places to go were excluded.¹
2. Percentage of predicates (verbs and adjectives combined).
3. Percentage of grammatical function words, which include pronouns, question words, prepositions and quantifiers.²
4. Cumulative percentage of onomatopoeic words, routines and names of people.

[1] Caselli *et al.* (1995) suggested not to include the category 'places to go' in the count of common nouns because many of its items function more like adverbials in the adult language.

[2] We used the expression 'function words' in order to maintain the same terminology adopted in the cross-linguistic study by Caselli *et al.* (1995) which is also focused on Italian. This label however refers to a category which includes the same lexical items which in the literature are usually referred to as 'closed class words'.

Moreover, for the 200-word questionnaire only, we counted the number of personal pronouns that mothers credited to their children.

RESULTS

Vocabulary growth

The thirty-seven children who completed the project reached the 200-word stage at 2;6. In this respect, our sample is comparable with normative data which reports that at this age only children below the tenth percentile do not yet have a vocabulary of 200 words (Caselli & Casadio, 1995).

There is, however, great variability in rate of vocabulary development. At 19 months of age, for example, about 40% of children have not yet reached the 50-word stage, but about 30% have a vocabulary size equal to, or greater than, 100 words (see Fig. 1).

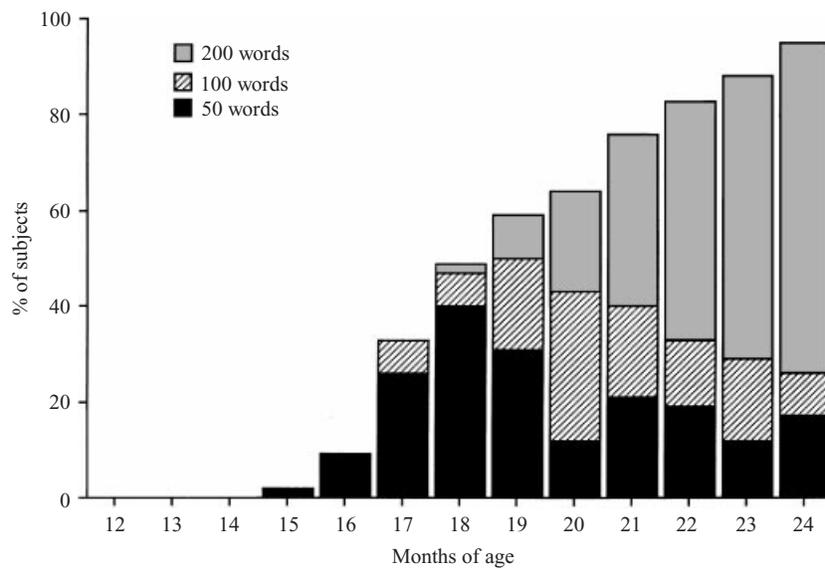


Fig. 1. Percentage of infants reaching the 50-, 100- and 200-word stages from 12 to 24 months of age.

To investigate which factors are related to a faster or slower rate of development, we first analysed the influence of demographic variables (sex, position in the family, mother's education, father's education) on the children's percentile scores at 50, 100 and 200 words (Table 1). These data were evaluated with a series of one-way ANOVAs. Results showed a significant female advantage ($F_{1,40} = 5.32; p < 0.05$) and an effect of mother's education ($F_{2,39} = 5.34; p < 0.01$), but only for the first stage of vocabulary

TABLE 1. Mean group percentile scores at the three stages of vocabulary sizes

| | 50 words | | 100 words | | 200 words | |
|------------------------|---------------------------|----------------|------------------|---------------|------------------|---------------|
| | Percentile score | Age (months) | Percentile score | Age (months) | Percentile score | Age (months) |
| Sex | | | | | | |
| Males | 30.63* (21.99)† | 20.3 (2.8) | 32.64 (17.87) | 21.7 (2.3) | 46.17 (27.06) | 23.1 (2.8) |
| Females | 47.45 (25.24) | 19.2 (2.3) | 46.35 (25.33) | 20.2 (2.7) | 58.85 (24.47) | 22.5 (3.0) |
| Position in the family | | | | | | |
| First-born | 39.44 (25.33) | 19.3 (3.0) | 43.14 (17.12) | 20.4 (2.4) | 58.50 (22.74) | 21.9 (2.6) |
| Other | 37.20 (24.63) | 19.1 (2.50) | 35.53 (29.70) | 21.7 (2.9) | 45.00 (29.37) | 23.7 (2.9) |
| Mother's education | | | | | | |
| 5–8 years | 36.00 (27.86) | 19.6 (3.5) | 39.22 (28.69) | 21.1 (3.1) | 47.67 (26.99) | 22.8 (2.8) |
| 13 years | 25.14 (20.0) | 20.8 (2.6) | 29.17 (25.49) | 22.4 (2.9) | 41.09 (27.14) | 24.4 (3.3) |
| 18 years | 51.47 (38.64) | 17.6 (1.4) | 48.69 (13.35) | 19.7 (1.5) | 63.59 (21.87) | 21.4 (2.0) |
| Father's education | | | | | | |
| 5–8 years | 25.11 (23.47) | 21.2 (3.3) | 37.55 (32.38) | 21.7 (3.1) | 43.67 (25.25) | 23.5 (3.1) |
| 13 years | 41.33 (24.44) | 18.9 (2.7) | 40.14 (23.68) | 20.7 (2.1) | 52.35 (26.31) | 22.6 (2.5) |
| 18 years | 43.17 (24.66) | 18.5 (2.3) | 40.88 (20.06) | 20.8 (2.8) | 56.88 (26.92) | 22.4 (3.2) |

* Bold characters highlight significant differences.

† Values in parentheses are standard deviations.

development, i.e. the 50-word stage. A *post hoc* test (Tukey–HSD) revealed that children whose mothers have the highest level of education were credited with a significantly higher vocabulary than children whose mothers have a medium level of education. By contrast, children whose mothers have a lower level of education did not differ significantly from the other two groups. In our data gender accounts for 12 % of the variance, while mother's education is responsible for about 20 % of variance.

Father's education and position in the family do not show any significant effect at any stage of vocabulary development.

A faster or slower rate of vocabulary growth is a stable individual characteristic of our subjects. Children who reached the 50-word stage first,

VOCABULARY DEVELOPMENT IN ITALIAN CHILDREN

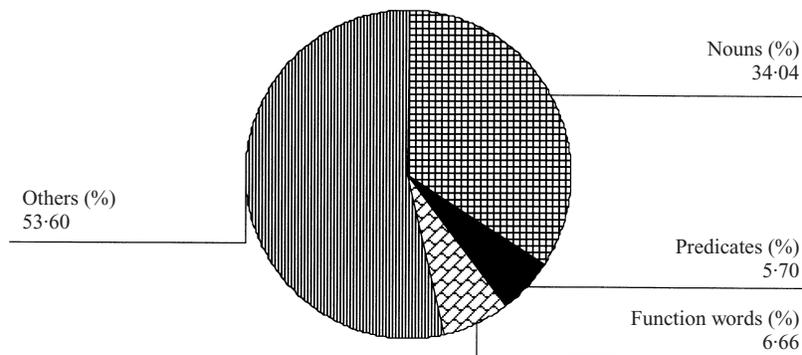


Fig. 2. Percentage of different lexical items at 50-word stage.

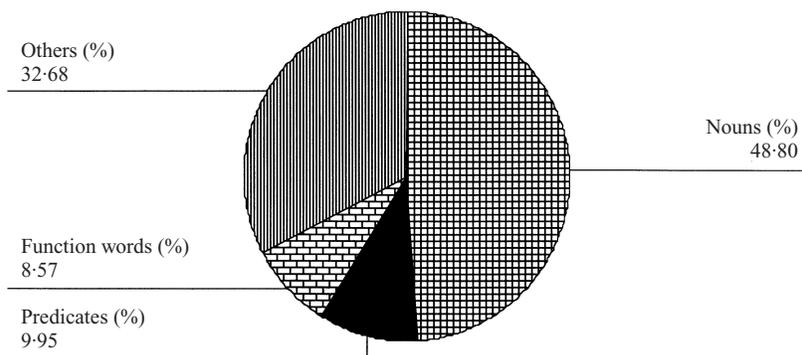


Fig. 3. Percentage of different lexical items at 100-word stage.

also reached the 100-word ($r = 0.74, p < 0.01$) and 200-word stage ($r = 0.73, p < 0.01$) first.

Vocabulary composition

Figures 2, 3 and 4 show vocabulary composition at the different stages of vocabulary size examined in this study (50, 100 and 200 words). About half the vocabulary items in the earliest stage of vocabulary development are therefore onomatopoeic words, routines or names of people, a datum which has cross-cultural stability (cf. Caselli *et al.* 1995, Berglund & Eriksson, 1998). At the next two stages, however, the most frequent lexical items are nouns. We also examined the rate of change from one vocabulary level to another, adopting the procedure described by Bates *et al.* (1994), but calculating rate-of-change statistics for individual children allowed by the longitudinal design adopted in this study.

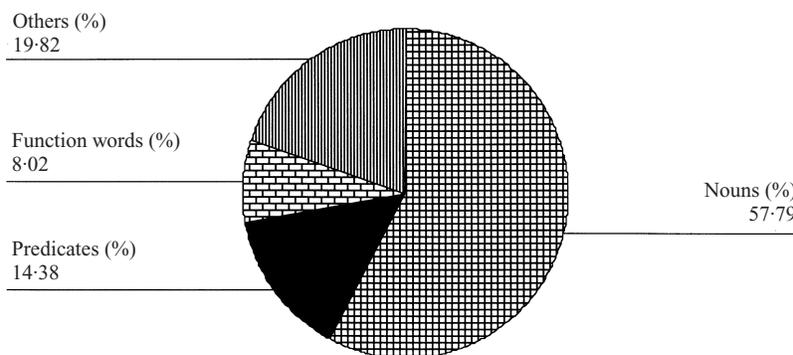


Fig. 4. Percentage of different lexical items at 200-word stage.

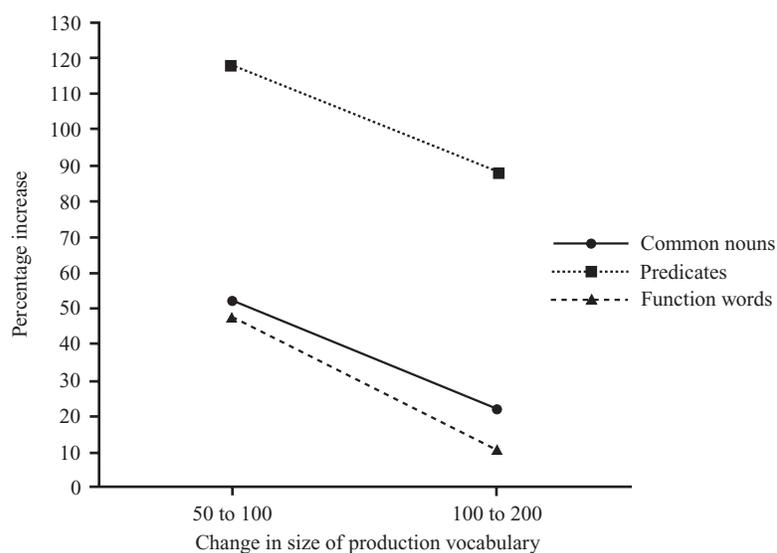


Fig. 5. Percentage increase in number of common nouns, predicates and function words from one vocabulary level to another.

In Figure 5 we can see that for all classes of lexical items considered in this analysis there is a deceleration in growth rate as vocabulary size increases, a pattern similar to that observed by Bates *et al.* (1994). However the rate of change is always smaller than that observed in the cross-sectional sample of English-speaking children and predicates appear to have a much higher rate of change than nouns and closed class items in both comparisons.

In order to examine the course of development in greater detail, we first compared the percentages of different lexical items calculated for each

VOCABULARY DEVELOPMENT IN ITALIAN CHILDREN

TABLE 2. *Descriptive statistics for the vocabulary composition measures in the first and the second 50 words*

| Measures | M | s.D. | Range |
|--------------------|-------|-------|-------|
| 1st 50 word | | | |
| Common nouns (%) | 34.04 | 10.29 | 10-53 |
| Predicates (%) | 5.70 | 4.40 | 0-16 |
| Function words (%) | 6.66 | 4.33 | 0-20 |
| 2nd 50 word | | | |
| Common nouns (%) | 60.35 | 11.65 | 33-86 |
| Predicates (%) | 13.54 | 6.96 | 0-31 |
| Function words (%) | 10.62 | 6.22 | 2-29 |

TABLE 3. *Descriptive statistics for the vocabulary composition measures in the first and the second 100 words*

| Measures | M | s.D. | Range |
|--------------------|-------|-------|-------|
| 1st 100 word | | | |
| Common nouns (%) | 48.80 | 10.18 | 27-69 |
| Predicates (%) | 9.95 | 4.11 | 1-20 |
| Function words (%) | 8.57 | 4.38 | 2-18 |
| 2nd 100 word | | | |
| Common nouns (%) | 55.74 | 6.14 | 46-70 |
| Predicates (%) | 14.37 | 3.58 | 8-23 |
| Function words (%) | 8.01 | 2.79 | 2-14 |

TABLE 4. *Concurrent and longitudinal correlations between the percentage of common nouns, predicates and function words in the first and the second 50 words*

| 1st 50 words | 1st 50 words | | | 2nd 50 words | | |
|-------------------|--------------|-------|---------|--------------|---------|--------|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| 1. Nouns | | -0.25 | -0.53** | +0.40* | -0.57** | -0.30 |
| 2. Predicates | | | +0.23 | +0.02 | -0.15 | +0.06 |
| 3. Function words | | | | -0.43* | +0.22 | +0.46* |

individual child in the first and second 50 words (table 2) and then the percentages of the different lexical items calculated in the first and second 100 words (table 3). In the passage from the first to second 50 words there is a significant increase in the percentage of nouns ($t = 12.93$; $p < 0.001$), predicates ($t = 5.33$; $p < 0.001$) and function words ($t = 4.54$; $p < 0.001$). Otherwise, in the comparison between the first and second 100 words, only the percentage of predicates increases significantly ($t = 6.88$; $p < 0.001$).

Stability of individual differences in vocabulary composition is moderate comparing the first and second 50 words (table 4) and becomes stronger comparing the first and second 100 words (table 5).

TABLE 5. *Concurrent and longitudinal correlations between the percentage of common nouns, predicates and function words in the first and in the second 100 words*

| 1st 100 words | 1st 100 words | | | 2nd 100 words | | |
|-------------------|---------------|---------|---------|---------------|---------|---------|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| 1. Nouns | | -0.71** | -0.79** | +0.73** | -0.40* | -0.69** |
| 2. Predicates | | | +0.38* | -0.62** | +0.49** | +0.43* |
| 3. Function words | | | | -0.59* | +0.38 | +0.79* |

Moreover, two strategies appear to differentiate children who have a preference for learning nouns from those who concentrate on learning predicates and function words. The relative proportion of nouns in the first 50 words is negatively correlated with the proportion of predicates in the following 50 words and the proportion of nouns in the first 100 words is negatively correlated with the proportion of both predicates and function words in the second 100 words.

A birth order effect on children's tendency to acquire personal pronouns earlier was investigated by examining differences in the raw number of items pertaining to this category in the questionnaires in which children reached about 200 words. For Italian-speaking children second-borns were credited by mothers with a significantly higher number of personal pronouns ($F_{1,35} = 4,26$; $p < 0.05$) as observed by Oshima-Takane & Derevensky (1990) and Pine (1995) for English-speaking children.

Our data also demonstrated that there are individual variations in the 'referential' dimension as observed for Italian children in Caselli *et al.*'s cross-linguistic study (1995). They are in contrast with data reported by Camaioni & Longobardi (1995) who found only one 'referential' child out of 12 subjects, using a criterion of a noun-to-total-word ratio of 40% or more as an index of referential style. Applying the same criterion to our subjects at the 50-word stage, 12 out of 42 subjects can be classified as referential, even if our 'noun category' is even more restricted than the one used by Camaioni & Longobardi (1995).

With increased vocabulary size, however, variability in the referentiality dimension diminishes. At the 100-word stage 'referential' children are 78% of the sample and at the 200 word stage they constitute 100%. These data confirm that stylistic differences in vocabulary development have to be verified with a strict control of vocabulary size, as suggested by Pine &

Lieven (1990), and that referential vocabulary proportions increase with age, irrespective of style differences, also for Italian-speaking children. In Camaioni & Longobardi's sample of 12 children the range of vocabulary was 30–100 words and this can probably explain the differences with respect to our data.

The majority of our referential subjects individuated at the 50-word stage are females (58%), first born (75%) with at least one parent with a high level of education (75%), but the only variable which significantly differentiates referential children from nonreferential children is mother's educational level (Mantel–Haenszel test for linear association $p < 0.03$).

The vocabulary spurt

According to our definition of 'vocabulary explosion', for each child we calculated the age at which it occurred and the vocabulary size from which the 'jump' was made. Moreover, we calculated how common nouns, predicates and function words contribute to the increase in vocabulary size which constitutes the spurt. Table 6 summarizes descriptive statistics for

TABLE 6. *Age, vocabulary size and composition of vocabulary spurt*

| | Mean | s.d. | Minimum | Maximum |
|--------------------|-------|-------|---------|---------|
| Age | 19.87 | 2.90 | 15 | 27 |
| Words | 48.84 | 18.17 | 11 | 90 |
| Increase (words) | 33.76 | 28.03 | 20 | 191 |
| Common nouns (%) | 56.88 | 14.73 | 16.67 | 79.41 |
| Predicates (%) | 11.85 | 8.87 | 0 | 38.46 |
| Function words (%) | 9.31 | 9.34 | 0 | 42.31 |

these variables. It appears clear that this phenomenon can be found in a large age range and in a vocabulary size much smaller and much larger than 50 words.

Even if there is a significant positive correlation between Age at spurt and Vocabulary size at spurt ($r = 0.55$; $p < 0.001$), there are children who exhibited the spurt starting from a small vocabulary but at an older age than the others (see figure 6 for examples of different growth curves). In any case, however, vocabulary growth in the months following the 'explosion' is always very rapid. Before the age at which the spurt was detected, the average monthly increment was about 7 words, while after the spurt it was about 52 (range 22–114).

From these data we hypothesized an advantage in vocabulary growth for children who showed the vocabulary explosion at a younger age which was tested computing Pearson's correlations between age at vocabulary explosion and percentile score at the three stages of vocabulary development. The

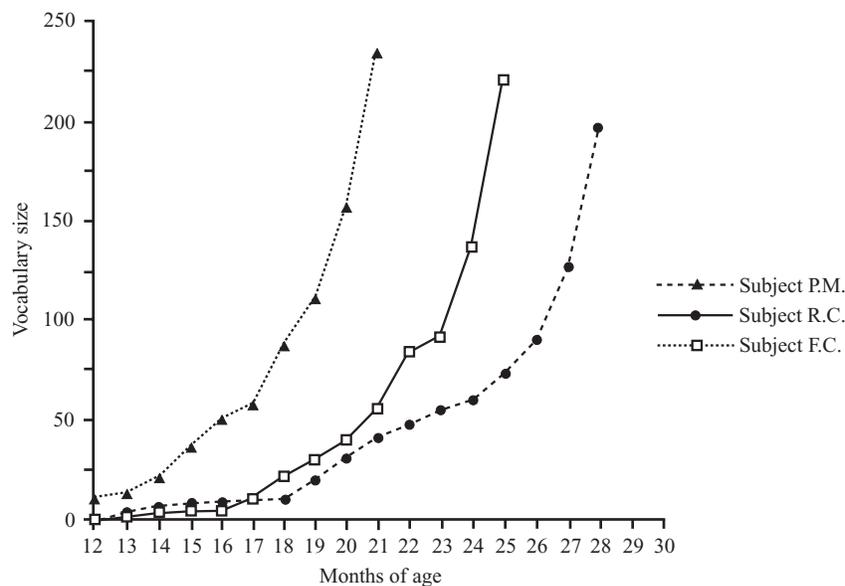


Fig. 6. Example of individual vocabulary growth curves.

results showed highly significant negative correlations between age at vocabulary explosion and percentile scores at 50 ($r = -0.77$; $p < 0.001$), 100 ($r = -0.84$; $p < 0.001$) and 200 words ($r = -0.80$; $p < 0.001$).

Even if in our data common nouns always constitute the largest relative percentage of vocabulary increment (with the only exception being a subject for whom function words constituted 40% of the increment) there is, however, a great variability in the relative contributions of predicates and function words. We hypothesized that (a) this variability is related to vocabulary size at the moment of the spurt, in so far as according to normative data for Italian (as well as English) verbs, adjectives and function words do not develop until common nouns are a well-established component of the vocabulary; (b) the style of vocabulary development, i.e. the referentiality dimension is negatively related to the appearance of predicates and function words in the vocabulary spurt. Results of the Pearson's correlation showed a significant positive relationship between the percentage of predicates in the increment and vocabulary size at the spurt ($r = 0.37$; $p < 0.03$) and a significant negative relationship between both the percentage of predicates and function words in the increment and the percentage of common nouns at the 50-word stage (our index of referentiality) ($r = -0.56$; $p < 0.001$; $r = -0.39$; $p < 0.02$ respectively). Moreover, these relationships are still significant even if for each variable the correlation is computed controlling the contribution of the other variable ($r = 0.27$; $p < 0.05$ re-

TABLE 7. *Concurrent and longitudinal relationships between the percentile score at 50, 100 and 200 words and vocabulary composition*

| | Percentile score | | |
|---|------------------|----------|----------|
| | 50-word | 100-word | 200-word |
| Nouns (%) | | | |
| 50-word | 0.17 | 0.16 | 0.27 |
| 100-word | 0.06 | 0.14 | 0.09 |
| 200-word | 0.07 | -0.01 | 0.09 |
| Predicates (%) | | | |
| 50-word | 0.03 | 0.03 | -0.24 |
| 100-word | 0.11 | 0.16 | 0.06 |
| 200-word | -0.07 | 0.18 | -0.02 |
| Function words (%) | | | |
| 50-word | -0.34* | -0.25 | -0.23 |
| 100-word | -0.21 | -0.17 | -0.08 |
| 200-word | -0.16 | -0.08 | -0.04 |
| Routines + onomatopoeic words + names of people (%) | | | |
| 50-word | 0.15 | 0.21 | 0.25 |
| 100-word | 0.01 | -0.26 | -0.20 |
| 200-word | 0.21 | 0.05 | -0.03 |

* $p < 0.03$.

spectively for the relationships between vocabulary size and percentage of predicates, controlling for referentiality; $r = -0.52$; $p < 0.001$ and $r = -0.36$; $p < 0.02$ for the relationships between referentiality and percentage of predicates or function words, controlling for vocabulary size).

Relationship between vocabulary composition and percentile scores

Table 7 shows the results of correlations made between percentile scores at the 50-, 100- and 200-word stages and vocabulary composition to investigate the possibility that a different vocabulary growth rate is also related to different vocabulary composition. From the table we can see that the only lexical variables which are significantly less frequent in the vocabulary of more precocious children are function words at the 50-word level. Percentage of common nouns is not related to a faster rate of vocabulary composition when vocabulary size is controlled, also confirming the absence of a relation between referential style and a faster vocabulary growth rate in Italian children.

DISCUSSION

The principal aim of this work was to investigate which aspects of Italian children's vocabulary development show similarities with studies focused on the acquisition of other languages and, on the other hand, which aspects

appear to be linked to characteristics of the input language. We chose to use mothers' reports of children's vocabulary instead of recordings of free speech in order to maintain the same assessment method used in most cross-language studies.

Italian infants demonstrated a great variability in vocabulary development rate which can be partially explained by socio-demographic factors, as has been shown for English-speaking children. More specifically, females reached the 50-word stage first, but the effect of gender was no longer significant in the following stages. This result is in agreement with data reported by Huttenlocher *et al.* (1991), who found a greater acceleration for females in early vocabulary growth, while gender effects declined in the 1;8–2;0 age period. The influence of mother's education on children's vocabulary growth has also been reported by Fenson *et al.* (1994), although in their large sample this variable explains only about 2% of the variance. More recently, Bornstein & Haynes (1998) have confirmed the advantage of females in language development assessed at 1;8 and the positive influence of mothers' education level.

Differently from Fenson *et al.* (1994), we did not find a positive linear relation between mother's education and vocabulary size. In our sample mothers with a low level of education reported a higher vocabulary size for their children than mothers with a medium level of education. We checked the possibility that these results may be due to confounding variables characterizing the mothers with a medium level of education (e.g. sex of children, fathers' education, children's position in the family, external work of mothers), but all appeared to be evenly distributed across the different education sub-groups. Fenson *et al.* (1994) reported a reliable tendency for lower-SES parents to check more items on the receptive vocabulary scale than higher-SES parents. We hypothesized that this explanation also fitted our data on expressive scale, but other investigations would be necessary to verify it.

Particular attention was devoted in this study to analysing the vocabulary spurt phenomenon. Our data gave evidence of the universality of the vocabulary spurt, in so far as all children in our sample showed it, even if at different ages. Moreover our data showed that the number of new lexical items which are learned each month after the spurt is about five times the number of words entering the vocabulary monthly before the spurt. Consequently, infants who showed evidence of this phenomenon at an earlier age, also had a clear advantage in rate of vocabulary growth.

The analysis of the vocabulary spurt we adopted in this study, which takes into consideration not only the age at which the spurt occurs, but also the vocabulary size before the spurt, also allowed us to analyse which type of lexical items were acquired at that point. We individuated two variables which influence the type of word: vocabulary size and style of acquisition.

Children who exhibited the vocabulary spurt after having reached quite a large vocabulary size also showed a high percentage of predicates as new words while 'referential' children showed a lower percentage of predicates and function words. These two variables (vocabulary size and style of acquisition), however, had an independent (and perhaps additive) influence on the nature of new words learned during the spurt. The contrasting results reported in the literature about this topic, therefore, could be due to the confounding effect of these variables.

Data on vocabulary composition are more controversial. General trends in the order in which the different lexical items are acquired confirm the theoretical arguments advanced in favour of a universal noun-verb sequence. In our data, however, it seems that predicates showed a steeper growth rate than has been observed for English. In fact, nouns are the most frequent class of lexical items at both the 100- and 200-word levels, but predicates showed a significantly higher rate of growth than nouns in the passage from 100 to 200 words. Moreover an analysis of individual differences showed a differentiation between children who tend to acquire more nouns and those who tend to acquire more predicates and function words, which was stable across development of the first 200 words. This result is compatible with the high salience of verbs in Italian mothers' input (Camaioni & Longobardi, 2000) and suggests the possibility that vocabulary development after the first 100 words could be more linked to specificity of target language than hypothesized by Caselli *et al.* (1995) when analysing the very early stages of development.

The stylistic differences in vocabulary acquisition do not seem to be strictly related to a faster or slower rate of vocabulary growth. The only lexical characteristics which distinguish slower children is a higher percentage of function words at the 50-word level. In this respect our data confirm Bates *et al.*'s (1988) suggestion that the early use of closed class words reflects a 'holistic' approach to language and may be associated with a slower rate of development.

The last cross-linguistic similarity we observed regards the possibility of applying the referential/expressive distinction to Italian children who have acquired their first 50 words, and the faster acquisition of personal pronouns by second-born children.

In conclusion, vocabulary development in Italian children appears to follow very general patterns which have cross-linguistic validity. The main difference we observed, i.e. the different growth rate of predicates in the second level of vocabulary development (from 100 to 200 words), can be explained by both the structural properties of Italian and the actual nature of the input received by Italian children. Even if the validity of mothers' report has been largely documented, our data concerning the influence of mothers' level of education on the expressive scale raises some doubts about the

reliability of data from questionnaires administered to low-educated mothers. It suggests that it may be necessary to collect more cross-linguistic data on children's actual speech in order to obtain a more complete picture of vocabulary development across languages.

REFERENCES

- Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms*. New York: 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.
- Benedict, H. (1979). Early lexical development: comprehension and production. *Journal of Child Language* **6**, 183–200.
- Berglund, E. & Eriksson, M. (1998). *Communicative development in Swedish children 16–28 months old. The Swedish Early Communicative Development Inventory – Words and Sentences* (rep. No. 845). Stockholm, Sweden: Stockholm University, Department of Psychology.
- Bloom, L. (1973). *One word at a time: the use of single-word utterances before syntax*. The Hague: Mouton.
- Bloom, L., Lifter, K. & Broughton, J. (1985). The convergence of early cognition and language in the second year of life: problems in conceptualization and measurement. In M. Barrett (ed.), *Single-word speech*. New York: Wiley.
- Bloom, L. & Capatides, J. B. (1987). Expression of affect and the emergence of language. *Child Development* **58**, 1513–22.
- Bornstein, M. H. & Haynes, M. O. (1998). Vocabulary competence in early childhood: measurement, latent construct and predictive validity. *Child Development* **69**, 654–71.
- Camaioni, L. & Longobardi, E. (1995). Nature and stability of individual differences in early lexical development of Italian-speaking children. *First Language* **15**, 203–18.
- Camaioni, L. & Longobardi, E. (2000). Noun versus verb emphasis in Italian mother-to-child speech. Manuscript submitted for publication.
- 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.
- Caselli, M. C. & Casadio, P. (1995). *Il primo vocabolario del bambino*. Milano: Franco Angeli.
- Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: a cross-linguistic study. *Journal of Child Language* **22**, 497–529.
- Dale, P. S., Bates, E., Reznick, S. & Morisset, C. (1989). The validity of a parent report instrument of child language at 20 months. *Journal of Child Language* **16**, 239–50.
- Dromi, E. (1987). *Early Lexical Development*. Cambridge: C.U.P.
- 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.
- Goldfield, B. A. & Reznick, J. S. (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. *Journal of Child Language* **17**, 171–83.
- Gopnick, A. & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development* **58**, 1523–31.
- Grimm, H. (1999). *Störungen der Sprachentwicklung: Grundlagen – Ursachen*. Göttingen: Hogrefe.
- Huttenlocher, J. (1974). The origin of language comprehension. In R. L. Solso (ed.), *Theories in cognitive psychology: the Loyola Symposium*. Potomac, MD: Erlbaum.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology* **27**, 236–48.
- Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E. & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language* **20**, 523–50.

- Jones, C. & Adamson, L. B. (1987). Language use in mother-child and mother-child-sibling interaction. *Child Development* **58**, 356-66.
- 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.
- Lucariello, J. (1987). Concept formation and its relation to word learning and use in the second year. *Journal of Child Language* **14**, 309-32.
- Lyytinen, P., Lari, N., Lausvaara, A. & Poikkeus, A. M. (1994). Language skills of Finnish-speaking toddlers. *Poster presented at the XIIIth Biennial Meetings of the International Society for the Study of Behavioural Development*, June 28-July 2, Amsterdam, The Netherlands.
- MacWhinney, B. (2000). *The CHILDES Project: tools for analyzing talk* Volume II: the database. Mahwah, NJ: Erlbaum.
- Maital, S. L., Dromi, E., Sagi, A., Bornstein, M. H. (1998). The Hebrew Communicative Development Index: language specific properties and cross-linguistic generalizations. Manuscript submitted for publication.
- Markman, E. M. (1991). The whole object, taxonomic and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. A. Byrnes (eds), *Perspectives on language and thought*. Cambridge: C.U.P.
- Mervis, C. B. & Bertrand, J. (1995). Early lexical acquisition and the vocabulary spurt: a response to Goldfield & Reznick. *Journal of Child Language* **22**, 461-8.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development* **38**. No. 149.
- Nelson, K. & Lucariello, J. (1985). The development of meaning in first words. In M. Barrett (ed.), *Single-word speech*. New York: Wiley.
- Ogura, T. (1998). Lexical development in Japanese children and U.S. children. Paper presented at the Annual Conference of the Japanese Psychological Association, Tokyo.
- Oshima-Takane, Y. & Derevensky, J. (1990). *Do later-born children delay in early language development?* Poster presented at the International Conference on Infant Studies, Montreal, Canada.
- Pine, J. M. (1995). Variation in vocabulary development as a function of birth order. *Child Development* **66**, 272-81.
- Pine, J. M. & Lieven, E. (1990). Referential style at 13 months: why age-defined cross-sectional measures are inappropriate for the study of strategy differences in early language development. *Journal of Child Language* **17**, 170-87.
- Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language* **20**, 43-60.
- Poulin-Dubois, D., Graham, S. & Sippola, L. (1995). Early lexical development: the contribution of parental labelling and infants' categorization abilities. *Journal of Child Language* **22**, 325-43.
- Reznick, J. & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology* **28**, 406-13.
- Reznick, J. & Goldfield, B. A. (1994). Diary vs. representative checklist assessment of productive vocabulary. *Journal of Child Language* **21**, 465-72.
- Tardif, T. (1996). Nouns are not always learned before verbs: evidence from mandarin speakers' early vocabularies. *Developmental Psychology* **32**, 492-504.
- Tardif, T., Gelman, S. A. & Xu, F. (1999). Putting the "noun bias" in context: a comparison of English and Mandarin. *Child Development* **70**, 620-35.
- Tardif, T., Shatz, M. & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian and Mandarin. *Journal of Child Language* **24**, 535-65.
- Thordardottir, E. T. & Weismer, S. E. (1996). Language assessment via parent report: development of a screening instrument for Icelandic children. *First Language* **16**, 265-85.
- Zurer Pearson, B. & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning* **44**, 617-53.

APPENDIX

| Subject | Sex | Mother educ. | Father educ. | Birth order | Age | Vocabulary size |
|---------|-----|--------------|--------------|-------------|---------|-----------------|
| AIT | M | 1 | 1 | 2 | 1;00.07 | 11 |
| AND | M | 3 | 3 | 1 | 1;01.07 | 12 |
| BOR | M | 2 | 2 | 2 | 1;02.23 | 19 |
| CAM | F | 3 | 3 | 1 | 1;00.04 | 4 |
| COL | M | 2 | 2 | 2 | 1;01.12 | 10 |
| DUR | F | 1 | 1 | 1 | 1;00.20 | 11 |
| FAC | F | 3 | 3 | 1 | 0;11.25 | 4 |
| FAL† | M | 2 | 1 | 1 | 1;01.12 | 2 |
| FAN | M | 2 | 3 | 1 | 1;00.07 | 12 |
| FED | F | 2 | 1 | 1 | 1;01.11 | 5 |
| FER | F | 2 | 2 | 1 | 1;04.04 | 19 |
| FOR | F | 2 | 2 | 2 | 0;11.28 | 0 |
| GRA | M | 1 | 2 | 1 | 0;11.23 | 5 |
| GRO | F | 2 | 1 | 2 | 1;00.05 | 21 |
| LAN | M | 3 | 3 | 1 | 1;00.13 | 6 |
| LAO | M | 3 | 2 | 1 | 1;00.12 | 5 |
| LOD† | M | 1 | 2 | 2 | 0;11.23 | 6 |
| LOK† | M | 1 | 2 | 2 | 0;11.23 | 3 |
| LOV | F | 3 | 2 | 1 | 1;03.12 | 9 |
| MAC | M | 2 | 1 | 1 | 1;00.22 | 11 |
| MAR | F | 3 | 2 | 2 | 1;03.27 | 38 |
| MAS | M | 3 | 3 | 1 | 1;04.02 | 38 |
| MAZ | F | 3 | 3 | 1 | 0;11.27 | 8 |
| MAO | M | 3 | 3 | 1 | 1;00.03 | 7 |
| MEN | M | 2 | 2 | 2 | 1;01.24 | 3 |
| MIO | F | 3 | 3 | 1 | 1;00.10 | 7 |
| NOD | F | 2 | 3 | 1 | 0;11.25 | 0 |
| PAG | F | 3 | 3 | 2 | 1;00.00 | 11 |
| PAL | F | 3 | 3 | 1 | 1;04.18 | 25 |
| PAS | F | 2 | 2 | 1 | 1;04.02 | 37 |
| PED | M | 2 | 1 | 1 | 1;01.16 | 21 |
| PET | F | 3 | 3 | 2 | 1;00.12 | 8 |
| POL | F | 2 | 1 | 1 | 1;01.12 | 10 |
| RAN | M | 2 | 2 | 2 | 1;01.06 | 9 |
| RIG | M | 3 | 2 | 2 | 1;00.11 | 0 |
| ROS | F | 1 | 1 | 2 | 1;04.01 | 14 |
| SAG | F | 1 | 2 | 2 | 1;01.01 | 21 |
| SAN† | M | 1 | 1 | 1 | 1;00.26 | 6 |
| SAV | M | 1 | 1 | 2 | 1;01.07 | 3 |
| SCH* | M | 3 | 3 | 1 | 1;00.18 | 9 |
| VEZ | M | 3 | 3 | 1 | 1;00.07 | 6 |
| ZUC | F | 3 | 3 | 2 | 0;11.30 | 7 |

* Subject missing 100-word questionnaire.

† Subjects missing 200-word questionnaires.