

Does the native language influence lexical composition in very preterm children at the age of two years? A cross-linguistic comparison study of Italian and Finnish children

First Language

2017, Vol. 37(4) 368–390

© The Author(s) 2017

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0142723717698006

journals.sagepub.com/home/fla



Suvi Stolt

University of Helsinki, Finland; University of Oulu, Finland

Silvia Savini and Annalisa Guarini

University of Bologna, Italy

Maria Cristina Caselli

Institute of Cognitive Sciences and Technologies, National Research Council, Italy

Jaakko Matomäki

Turku University Hospital, Finland

Helena Lapinleimu

Turku University Hospital, Finland; University of Turku, Finland

Leena Haataja

Helsinki University Hospital, Finland; University of Helsinki, Finland

Liisa Lehtonen

Turku University Hospital, Finland; University of Turku, Finland

**Rosina Alessandrini, Giacomo Faldella
and Alessandra Sansavini**

University of Bologna, Italy

Corresponding author:

Suvi Stolt, Department of Psychology and Logopedics (Speech and Language Pathology), Faculty of Medicine, University of Helsinki, P.O. Box 9, Helsinki FI-00014, Finland.

Email: suvi.stolt@helsinki.fi

Abstract

This cross-linguistic study investigated whether the native language has any influence on lexical composition among Italian ($N = 125$) and Finnish ($N = 116$) very preterm (born at <32 gestational weeks) children at 24 months (controls: 125 Italian and 146 Finnish full-term children). The investigation also covered the effect of maternal education (ME) on lexical composition. The Italian/Finnish MacArthur Communicative Development Inventory was used for gathering the data. Although the lexicons of the preterm children were smaller than those of the controls, the native language had no major effect on their lexical composition. The ME had a significant effect on preterm children's lexical composition, especially in the Finnish children. The findings indicate that lexical composition is not strongly affected by preterm birth. They also imply that lexical composition is a robust phenomenon that is connected to lexicon size and is not language-specific when analysed in broad terms, although some language-specific features were also detected.

Keywords

Cross-linguistic study, language development, lexical composition, lexicon, preterm children

Introduction

Children at the early stage of language acquisition universally acquire words that refer to things related to their daily context (i.e. to people, toys, food; Clark, 2003). They have also been shown to acquire nouns before predicate terms in many languages (Bornstein et al., 2004; Gentner, 1982). However, the role of the native language in the early acquisition of lexical categories is not fully understood. For example, it is not clear whether the structure of the native language influences the acquisition order of early word categories, and if it does, how. Thus, there is a need for cross-linguistic comparative studies that will provide information on the possible general, or even universal, features of language development. Moreover, the effect of the native language on early lexical composition may differ in clinical populations such as preterm children. Although there has been growing interest in early language development among preterm children in recent years (e.g. Foster-Cohen, Edgin, Champion, Woodward, 2007; Kern & Gayraud, 2007; Perez-Pereira, Fernandez, Gomez-Taibo, & Resches, 2014; Sansavini et al., 2006; Stolt et al., 2007), there have been no cross-linguistic comparisons concerning this group. Such studies could enhance understanding of which features in early development result from very preterm birth, and which ones are influenced by the children's native language or other background factors. This is the first cross-linguistic comparative study to focus on the role of the native language in the acquisition of lexical categories in two-year-old very preterm children (born <32 gestational weeks). The children in the present study were acquiring one of two different languages: a Romance language (Italian) and a Finno-Ugric language (Finnish).

Reorganization during early lexical development

Reorganization occurs in the way the lexicon is composed during the second year of life (Bates et al., 1994). Children primarily acquire social terms (names of people, words connected to routines, onomatopoeic expressions; e.g. Bates et al., 1994) at the very beginning of their lexical development, then after this early phase a shift takes place from referencing to predication and grammar (e.g. Caselli, Casadio, & Bates, 1999). Children actively begin to acquire words with a clear naming function (i.e. nouns) when their lexicon consists of between about 50 to 200 words (i.e. shift to referencing; e.g. Bates et al., 1994; Caselli et al., 1999). Verbs and adjectives are rare in small lexicons, but as children incorporate more words into their vocabulary they begin to focus more actively on them (i.e. shift to predication). When children have acquired reasonably large lexicons (i.e. > 400 words; Bates et al., 1994) they undergo a shift to grammar, when the active acquisition of closed-class words such as prepositions, pronouns, articles, question words, quantifiers and conjunctives begins. This shift from early social terms through referencing and predication to grammar has been observed in various languages (Bates et al., 1994; Caselli et al., 1995; Caselli et al., 1999; Jackson-Maldonado, Thal, Marchman, Bates, & Gutierrez-Clellen, 1993; Maital, Dromi, Sagi, & Bornstein, 2000; Stolt et al., 2007; Stolt, Haataja, Lapinleimu, & Lehtonen, 2008). Furthermore, findings in cross-sectional samples collected at one age point only (Conboy & Thal, 2006; Stolt et al., 2007) further indicate that this developmental shift in the early acquisition of vocabulary may indeed be strongly tied to lexicon size, and not only to the age of the children.

One main claim linked to the early development of lexical categories is that children universally acquire nouns before predicate terms (*the natural partitions hypothesis*; Gentner, 1982). This hypothesis is grounded on the notion that the linguistic distinction between nouns and verbs is based on the perceptual-conceptual distinction between concrete concepts and predicative concepts (i.e. activity, change-of-state, causal relation; Gentner, 1982). The category that corresponds to nouns is conceptually more basic than categories corresponding to predicate terms, which explains why nouns are acquired before predicate terms. This hypothesis has proved to be valid in many languages (Bornstein et al., 2004; Gentner, 1982; Gentner & Boroditsky, 2001), although findings among Korean and Mandarin Chinese children have challenged it (Choi & Gopnik, 1995; Tardiff, 1996). Gentner and Boroditsky (2001) went on to expand the natural partitions hypothesis further with their *division of dominance* hypothesis, according to which open-class words, especially nouns, have a strong *cognitive dominance* meaning that the reference of these words can be picked up by cognitive means (perceptual experience). Closed-class words, on the other hand, exhibit strong *linguistic dominance* in that their meaning does not exist independently of language. Verbs lie on the continuum between cognitive and linguistic dominance: unlike closed-class words they have a denotational function, and yet the target reference they denote is expressed using language structures. Words with strong cognitive dominance are acquired earlier than those with strong linguistic dominance (Gentner & Boroditsky, 2001). The division of dominance may explain, at least to some extent, the developmental shift in the early lexicon described by Bates et al. (1994).

Cross-linguistic comparisons could shed a light on whether or not the developmental shift from early social terms to referencing, predication and grammar happens in parallel among children with different native languages. However, very few cross-linguistic studies have been conducted on this topic. Caselli et al. (1995) analysed the lexical compositions of English-speaking American and Italian-speaking children aged between 8 and 16 months and found a parallel shift in both languages. A later study (Caselli et al., 1999) focused on lexical development in English-speaking American and Italian-speaking children aged between 18 and 30 months, and again a comparable developmental shift in vocabulary was found in both languages, as well as some differences. The percentage of social terms was higher in the Italian children's lexicons, which was attributed to the cultural differences between the two countries. In addition, the Italian children acquired closed-class words gradually as their vocabulary expanded, whereas the American children acquired very few such words before having acquired roughly 400 words (Caselli et al., 1999). This was attributed to differences in grammar between the two languages.

The development of lexical composition in preterm children

Very few studies have investigated early lexical composition among very preterm children, and none of them included cross-linguistic comparison. Kern and Gayraud (2007) analysed lexical development in French preterm children at 24 months, and reported that most of them had acquired fewer nouns, predicate terms and closed-class words than the controls. However, lexical composition was not analysed in relation to lexicon size in this study, and the question remains open as to whether there is a parallel shift from referencing to predication and grammar in the lexicons of French preterm and full-term children.

Three earlier studies (Sansavini et al., 2006; Sansavini, Guarini, & Savini, 2011; Sansavini, Guarini, et al., 2011) focused on lexical composition in Italian preterm children. Using the Italian short form version of the Communicative Development Inventory (CDI), Sansavini, Guarini, et al. (2011) found that very preterm (born at ≤ 32 gestational weeks) children had acquired fewer words than the controls at 12, 18 and 24 months, and had fewer social terms in their lexicons at 18 months. However, interpretations of the results should take into consideration the fact that the short form version of the CDI provides information only on a highly selected set of lexical items. When the long form of the CDI was used, the Italian preterm children exhibited a smaller lexicon size and fewer social terms, nouns, predicates and grammatical function words than the controls at 24 months (Sansavini, Guarini, & Savini, 2011), but at 30 months their and the full-term children's lexicons consisted of comparable percentages of words from different lexical categories (Sansavini et al., 2006). In addition, a recent study (Sansavini et al., 2015), which employed a direct evaluation of nouns and predicates at 24 months in Italian children, showed that nouns are acquired earlier than predicates in both extremely preterm and full-term Italian children, but with a significantly lower number of nouns and predicates in the extremely preterm group. Neither of these studies included a detailed analysis of lexical composition in relation to lexicon size, however. Thus, the question remains open as to whether a comparable shift from referencing to grammar occurs in the lexicons of Italian preterm children as has been found in Italian full-term children (Caselli et al., 1999).

The lexical composition of Finnish preterm (very-low-birth-weight [VLBW], birth weight ≤ 1500 g; born between 23 and 35 gestational weeks) children has been the focus of two previous studies (Stolt et al., 2007; Stolt, Haataja, Lapinleimu, & Lehtonen, 2009). According to the findings of the longitudinal study (Stolt et al., 2009), healthy VLBW children had acquired a comparable number of words from different lexical categories in their expressive lexicons to the controls between the ages of 9 and 18 months, but they had fewer social terms, nouns, verbs and adjectives in their lexicons than the controls at 24 months. When the development of lexical categories was analysed in relation to lexicon size, the findings showed a comparable shift from referencing to predication and to grammar in both groups. Furthermore, the findings of a cross-sectional sample at 24 months (Stolt et al., 2007) also revealed a comparable developmental shift in the lexicons of VLBW children and their controls, and only minor differences between the groups.

The effect of maternal education on the lexicons of preterm children

Results regarding the influence of maternal education (ME) level, or family socio-economic status, on language development among preterm children are not consistent: some studies report a significant positive effect in the preterm group (e.g. Ortiz-Mantilla, Choudhury, Leever, & Benasich, 2008; Sansavini, Guarini, et al., 2011; Stolt et al., 2007), whereas others could establish no influence (Kern & Gayraud, 2007). There has been no reported research thus far on the effect of ME on the composition of the early lexicon in preterm children, although such information could be useful in planning early interventions, for example.

The aims of the present study

One of the main aims of this cross-linguistic study was to find out whether the native language has any effect on the lexical composition (i.e. the shift from social terms to referencing, predication and to grammar; Bates et al., 1994) in very preterm children acquiring Italian and Finnish. These languages differ in several aspects (e.g. on the phonological and morphological levels), which may also influence lexical development. Appendix A sets out the basic characteristics of the two languages. A further aim was to analyse the possible effect of ME on the lexicons of Italian/Finnish very preterm children. The research questions were: (1) Does the native language have any effect on lexical composition (i.e. the shift from social terms to referencing, predication and to grammar; Bates et al., 1994) in very preterm children acquiring Italian and Finnish? and (2) Does ME affect the acquisition of lexical categories in Italian and/or in Finnish very preterm children?

The predictions were as follows. First, with regard to lexical composition it was expected that the lexicons of Italian and Finnish very preterm children would exhibit a comparable robust trend from social words to nouns and predicates, and eventually to closed-class words as found in the full-term control children. This prediction was based on earlier findings among Italian full-term children (Caselli et al., 1999) and Finnish full-term and preterm children, although in smaller samples than the present ones (Stolt et al., 2007). Nevertheless, some differences (e.g. fewer closed-class words in the lexicons of preterm children; cf. Stolt et al., 2007) due to a very preterm birth were also expected to

be found. Second, it was also predicted that linguistic differences in the native languages would affect the rhythm in the acquisition of lexical categories. With regard to social-pragmatic words, given the differences in the percentage of social terms between American English-speaking and Italian children (Caselli et al., 1999), it was expected that the percentage of social terms may differ in the lexicons of Italian and Finnish children. With regard to nouns and predicates it was hypothesized that because the noun and verb inflectional system is more complex in Finnish than in Italian (see Appendix A), Finnish children may more readily acquire nouns and predicates at an early stage of lexical acquisition than Italian children. This prediction was based on the findings that vocabulary and grammar are significantly associated at the end of the second year (e.g. Bates & Goodman, 1999). There is evidence that children use grammatical information to learn the meaning of a novel word, especially verbs (syntactic bootstrapping; e.g. Naigles, 1990; see also Bates & Goodman, 1999). In addition, the type of association between lexicon and grammar may differ in different languages (Thordardottir, Weismer, & Evans, 2002). Thus, given the differences in morphology between Italian and Finnish, it was expected that these differences would also be reflected in lexical acquisition. In addition, it was anticipated that Italian children would acquire closed-class words more actively early on than Finnish children, given that Italian uses articles and prepositions actively (see Appendix A) whereas there are no articles in Finnish, and many meanings that are expressed using prepositions in Italian are expressed using case endings in Finnish. Moreover, given the association between ME and lexicon size at 24 months of age among preterm children reported in a previous study (Stolt et al., 2007), it was expected that ME would affect the development of lexical categories among very preterm children.

The data investigated in the present study have been partially analysed in earlier studies (Caselli, Pasqualetti, & Stefanini, 2007; Sansavini, Savini, et al., 2011; Stolt et al., 2007; Stolt et al., 2013). However, the present samples of preterm children differ slightly from those in the earlier studies in that the analysis incorporates the findings on healthy very preterm children (born <32 gestational weeks; see below for more information on the sample characteristics). The present study comprises a detailed cross-linguistic comparison of (1) the acquisition of lexical categories in very preterm children and (2) the early lexicons of Italian and Finnish children. The findings also shed light on the effect of ME on the acquisition of lexical categories in very preterm children, which has not previously been investigated from a cross-linguistic perspective.

Method

Participants

The participants were Italian ($N = 125$) and Finnish ($N = 116$) very preterm children (Table 1). The two groups were modified from earlier study samples, an Italian (see Sansavini, Guarini, & Savini, 2011) and a Finnish (PIPARI study; Lapinleimu, Haataja, & Lehtonen and the PIPARI study group; see Stolt et al., 2013) sample. The original inclusion criteria for the Italian sample were as follows: preterm children (≤ 32 gestational weeks) who were born within a six-year period (2003–2008) in the neonatal intensive care unit of Bologna University Hospital. Children with major cerebral damage

Table 1. Background characteristics of the Italian (It) and Finnish (Fi) very preterm children (VP; born <32 gestational weeks). The values presented are the numbers and percentages, and cases in which the mean value (M) and standard deviation (SD) are presented are indicated.

	ItVP children N = 125	FiVP children N = 116
Gestational age in weeks, Mean (SD), min.–max.	29 (2), 23–31	28 (2), 23–31
Birth weight in grams, Mean (SD), min.–max.	1150 (381), 475–2275	1128 (316), 400–1855
Spontaneous delivery	7 (6)	48 (41)
SGA –2SD ^a	49 (39)	37 (32)
Male	68 (54)	71 (61)
2nd- or later-born child	28 (22)	24 (21)
Twins or multiple children	35 (28)	45 (39)
BPD	26 (21)	15 (13)
Sepsis	21 (17)	21 (18)
>4 otitis / year	4 (3)	4 (4)
Nursery at 24 months	41 (33)	65 (56)
<i>Cognitive development^b</i>		
Normal/Mild (–1SD) delay	115 (92)	110 (95)
Moderate/Severe (–2SD) delay	10 (8)	4 (4)
<i>Maternal education</i>		
Compulsory	20 (16)	12 (10)
Secondary	61 (49)	27 (23)
Further education	44 (35)	73 (63)

SGA = small for gestational age, BPD = bronchopulmonary dysplasia.

^aSGA status was defined based on the same reference values for all children (= –2SD from the age- and gender-specific growth charts for Finnish children; Dietz, Callaghan, Smith, & Sharma, 2009) to allow comparison between the groups in the present study. Following this procedure, there was no difference between the groups in the number of SGA children ($\chi^2(1, N = 240) = 1.51, p = .22$).

^bInformation on cognitive development was missing for two of the Finnish VP children, and information on maternal education level was missing for four of the Finnish VP children.

confirmed in an ultrasound scan (e.g. periventricular leukomalacia, intraventricular haemorrhage >II grade, hydrocephalus), retinopathy of prematurity (>II grade), congenital malformations, or visual or hearing impairment were excluded. The original inclusion criteria for the Finnish sample were as follows: preterm children (<37 gestational weeks) who were born within a six-year period (2001–2006) in Turku University Hospital, and whose birth weight was <1500 g. The criteria were expanded from the beginning of 2004 to cover children born at <31+6 gestational weeks, including those with a higher birth weight than <1500 g if a child was born at <31+6 gestational weeks. Furthermore, the family had to live in the Turku University Hospital catchment area and had to understand Finnish or Swedish in order to be able to complete the follow-up forms. The Italian and Finnish samples of preterm children were modified for the present study to make the groups comparable: we report findings on healthy very preterm children (born

<32 gestational weeks). The exclusion criteria for the present study were a major brain abnormality at term verified using serial brain ultrasound (see Sansavini, Savini, et al., 2011, and Rademaker et al., 2005 and Reiman et al., 2008 for the definition used in the Italian/Finnish sample), a cerebral palsy diagnosis, and/or bilateral hearing loss >40 dB. All the preterm children in the present study were being brought up in monolingual Italian- or Finnish-speaking families.

The Italian and Finnish groups of preterm children were comparable (see Table 1) in terms of birth weight, GA, gender, order of birth (i.e. firstborn vs later born), being small for gestational age (SGA; see the note in Table 1), being a singleton vs twin/triplets, etc., and having bronchopulmonary dysplasia (i.e. a need for supplemental oxygen at 36 weeks of gestational age), sepsis and recurrent otitis media (>4 otitis media in one year). Those with severe intraventricular haemorrhaging (grades III and IV) were excluded from both samples.

The cognitive development of the preterm children was examined at 24 months (corrected age, i.e. the age calculated from the expected date of delivery). The general development quotient (DQ) of the Revised Griffiths Mental Development Scales was used in Italy, and the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development II was used in Finland. The Italian and Finnish groups of very preterm children did not differ in terms of how many children with moderate or severe cognitive delay ($-2SD$: $DQ < 77$ or $MDI < 70$ standard scores) there were in each group, $\chi^2(1, N = 239) = 2.18, p = .14$.

ME level was categorized according to the educational system in Italy and in Finland as follows: compulsory (Italian sample: ≤ 10 years; Finnish sample: ≤ 9 years), secondary (Italian: 11–13 years; Finnish: 10–12 years) and further education (Italian: > 13 years; Finnish: > 12 years; see Table 1). The Finnish mothers of very preterm children had a higher educational level than the respective Italian mothers, $\chi^2(1, N = 237) = 17.48, p < .001$.

Groups of Italian ($N = 125$) and Finnish ($N = 146$) full-term (> 37 GA) children served as controls in the present study. The Italian and Finnish groups included 56 males (45%) and 71 males (49%), respectively, $\chi^2(1, N = 271) = 0.40, p = .53$. Almost all the controls were singletons (1 twin in the Finnish sample). All of the controls were brought up in monolingual families. The Italian control group was formed from two earlier study groups ($N = 70$, Caselli et al., 2007; $N = 55$, Sansavini, Guarini, & Savini, 2011). Unfortunately, exact information on cognitive development was not available for all the Italian controls (i.e. all the children from Caselli et al.'s study and 29 from Sansavini, Guarini, & Savini's study). However, their typical development was verified by a paediatrician, their parents and, when applicable, teachers. Exact information on cognitive development was available for 26 children ($DQ \geq 87$ standard scores) from the Sansavini, Guarini, and Savini (2011) study. The following information on ME was available for the children from Caselli et al.'s study (2007): 57% had at least one parent with secondary education, and 35% had at least one parent with further education. The ME levels of the controls from Sansavini, Guarini, and Savini's (2011) study were as follows: 7 (14%) with compulsory education, 19 (39%) with secondary education and 23 (47%) with further education (ME information was missing for 6 children; the percentages were

calculated from the available data). The Finnish full-term children were the controls from the PIPARI study (see Stolt et al., 2013), and they were all developing cognitively according to their age ($MDI \geq 84$ standard scores). The ME of the Finnish controls was as follows: 8 (6%) with compulsory education, 48 (33%) with secondary education and 81 (56%) with further education (ME data were missing for 9 controls).

Tools and analysis

The data were collected at 24 months using the standardized Italian and Finnish versions of the MacArthur–Bates Communicative Development Inventory (MB-CDI; Words and Sentences, WS form; Fenson et al., 1994; ItCDI: Caselli et al., 2007; FinCDI: Lyytinen, 1999). The percentages of social terms, nouns, predicates and closed-class words are comparable in the ItCDI/FinCDI (see Appendix B), although the total number of items is slightly higher in the ItCDI than in the FinCDI. The mean corrected age (i.e. calculated from the expected date of delivery, not the actual birth date) at which the CDI inventory was filled in was two years and two days (SD 13) for the Italian and two years and six days (SD 15) for the Finnish preterm children. The mean ages at which the CDI inventory was filled in were two years and a day (SD 5) for the Italian controls and two years and 11 days (SD 11) for the Finnish controls.

The following lexical categories were analysed (see Appendix B; Bates et al., 1994; Caselli et al., 1995, 1999): *social terms* including the names of people, early sound effects and words associated with games and routines; *common nouns* including words with a clear naming function; *predicate terms* including words from the CDI categories: ‘action words’ and ‘descriptive words’ (a common predicate-terms category was used because both verbs and adjectives indicate a state, an action or a relationship with a ‘primary’ word, i.e. noun); *closed-classed words* including those marking linguistic relations between content words in linguistic structures.

The percentages of words in different categories (social terms, nouns, predicates and closed-class words) were used in the analysis, calculated from the total number of words in the lexicon measured for each child using the CDI, and then analysed in relation to lexicon size. The children were therefore divided into seven sub-groups based on their lexicon size (cf. Caselli et al., 1999; see Table 2 for the numbers and percentages of children in each sub-group). An analysis of variance (ANOVA) was carried out to see if the children were distributed evenly among the seven sub-groups. In this analysis, lexicon size was used as a dependent variable, and the four study groups (Italian preterm, Italian full-term, Finnish preterm, Finnish full-term children) and seven lexicon-size sub-groups were used as between-group factors (cf. Caselli et al., 1999). There was a significant effect for lexicon size, $F(6, 484) = 2119, p < .001$, confirming that these groups did indeed differ by lexicon size. However, there was no significant main effect of group, $F(3, 484) = 1.23, p = .30$, and nor was the interaction between group and lexicon size significant, $F(18, 484) = 1.49, p = .09$. Thus, the four groups were evenly balanced in terms of vocabulary level.

First, an ANOVA was used to investigate the possible differences in lexicon size in Italian and Finnish very preterm and full-term children. In this ANOVA analysis, birth

Table 2. The lexicon-size sub-groups and the numbers of very preterm and full-term children in each sub-group (ItVP = Italian very preterm children; ItFT = Italian full-term children; FiVP = Finnish very preterm children; FiFT = Finnish full-term children).

Lexicon size	ItVP	ItFT	FiVP	FiFT
	<i>N</i> = 125	<i>N</i> = 125	<i>N</i> = 116	<i>N</i> = 146
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
<50	27 (22)	5 (4)	19 (16)	21 (14)
50–100	17 (14)	8 (6)	9 (8)	11 (8)
101–200	23 (18)	27 (22)	16 (14)	18 (12)
201–300	18 (14)	36 (29)	28 (24)	22 (15)
301–400	18 (14)	15 (12)	21 (18)	33 (23)
401–500	9 (7)	15 (12)	18 (16)	32 (22)
>500	13 (10)	19 (15)	5 (4)	9 (6)

status (preterm vs full-term) and native language (Italian vs Finnish) were used as between-group factors. Subsequently, separate ANOVAs were run on the percentage of each of social terms, nouns, predicates and closed-class words. The design for these analyses was 2 (birth status) \times 2 (native language) \times 7 (lexicon-size sub-groups). Post hoc comparisons for multiclass independent variables were conducted using Tukey's method.

Lastly, the effect of ME (i.e. compulsory, secondary, further education) on lexicon size and on the percentage of social terms, nouns, predicates and closed-class words in the vocabulary was analysed using a one-way ANOVA. This analysis was conducted separately for the Italian and Finnish children, and it was run only for the preterm children given that exact information on ME was not available for all full-term children. The significance level in all the tests was $p < .05$.

Results

Lexicon size

The mean lexicon size was 219 (SD 177) words in the Italian very preterm children and 293 (SD 160) words in the Italian controls. The respective values for the Finnish children were 244 (SD 155) and 276 (SD 164) words. The ANOVA analysis indicated a significant main effect of the birth status, $F(1, 508) = 13.36, p < .001$, the full-term children having roughly 53 more words in their lexicons than the very preterm children (95% confidence interval: 25–82). However, the main effect of the native language was not significant, $F(1, 508) = 0.09, p = .77$, and neither was the interaction between birth status and native language, $F(1, 508) = 2.14, p = .14$.

Lexical composition

Table 3 and Figures 1–4 give the mean percentages of the words in different lexical categories categorized by lexicon size for all the groups. Table 3 also gives the respective standard deviations.

Social terms. According to the ANOVA, the birth status had no significant effect on the percentage of social terms. However, the native-language effect was significant, $F(1, 484) = 9.42, p = .02$: the percentage of social terms was significantly higher among the Italian than among the Finnish children (difference: 2%; 95% CI: 0.01–0.04). The effect of lexicon size on the percentage of social terms was also significant, $F(6, 484) = 235.46, p < .001$: the percentage was higher in the lexicon-size sub-group with <50 words than in the sub-group with 50–100 words (difference: 24%; 95% CI: 0.19–0.29, $p < .001$); in the sub-group with 50–100 words than in the one with 101–200 words (difference: 12%; 95% CI: 0.07–0.16, $p < .001$); and in the sub-group with 101–200 words than in the one with 201–300 words (difference: 6%; 95% CI: 0.03–0.10, $p < .001$). The birth status by lexicon size interaction was significant, $F(6, 484) = 2.49, p = .02$: the percentage of social terms was significantly higher (9%) in the small lexicons (<50 words) of the preterm children than in those of the controls (95% CI: 0.05–0.14, $p < .001$). The interaction between native language and lexicon size was also significant, $F(6, 484) = 4.64, p < .001$: the percentage of social terms was higher in the lexicons of Italian children than in those of the Finnish children in the sub-group with 50–100 words (difference: 14%; CI: 0.09–0.18, $p < .001$).

Nouns. The effect of birth status on the percentage of nouns was not significant, but the effect of native language was, $F(1, 483) = 6.20, p = .013$: the percentage of nouns was significantly higher in the lexicons of the Finnish children than in those of the Italian children (difference: 2%; 95% CI: 0.00–0.04). The effect of lexicon size on the percentage of nouns was also significant, $F(6, 483) = 46.28, p < .001$, the percentages being significantly lower in the lexicon-size sub-group with <50 words than in the sub-group with 50–100 words (difference: 16%; 95% CI: 0.11–0.22, $p < .001$), and in the sub-group with 50–100 words than in the one with 101–200 words (difference: 6%; 95% CI: 0.01–0.10, $p = .013$). The interaction between birth status and lexicon size was not significant, but that between native language and lexicon size was, $F(6, 483) = 5.67, p < .001$: the percentage of nouns was higher in the sub-group with 50–100 words in the Finnish children compared to the Italian children (difference: 15%; 95% CI: 0.10–0.20; $p < .001$).

Predicate terms. The effect of birth status on the percentage of predicate terms was not significant, but the effect of native language was, $F(1, 484) = 55.92, p < .001$: the percentage of predicate terms was significantly higher in the lexicons of the Finnish children than in those of the Italian children, the difference being 3% (95% CI: 0.02–0.04). The effect of lexicon size on the percentage of predicate terms was also significant, $F(6, 484) = 164.79, p < .001$: the percentage of predicate terms was significantly lower in the lexicon-size sub-group with <50 words than in the one with 50–100 words (difference: 4%; 95% CI: 0.01–0.07, $p = .001$); in the sub-group with 50–100 words than in the one with 101–200 words (difference: 6%; 95% CI: 0.03–0.09, $p < .001$); in the sub-group with 101–200 words than in the one with 201–300 words (difference: 4%; 95% CI: 0.2–0.6, $p < .001$); in the sub-group with 201–300 words than in the one with 301–400 words (difference: 3%; 95% CI: 0.01–0.05); and in the sub-group with 301–400 words than in the one with 401–500 words (difference: 4%; 95% CI: 0.02–0.06, $p < .001$). The

Table 3. The percentages of words in different lexical categories calculated from the total number of words in the lexicon. The values presented are the mean percentages and the standard deviations of each lexicon-size sub-group, shown separately for the Italian (It) and Finnish (Fi) very preterm children (VP) and for the controls (FT).

VP	Language	Language			
		<50	50–100	101–200	201–300
Social terms	It	59 (19)	37 (9)	23 (5)	15 (2)
	Fi	63 (22)	26 (6)	18 (5)	13 (1)
Nouns	It	29 (17)	40 (12)	52 (8)	52 (6)
	Fi	26 (18)	55 (8)	54 (7)	54 (5)
Predicates	It	3 (4)	8 (5)	12 (6)	15 (5)
	Fi	7 (6)	9 (3)	16 (5)	20 (4)
Closed-class words	It	8 (11)	12 (8)	7 (4)	9 (4)
	Fi	1 (3)	5 (4)	6 (1)	7 (7)
		301–400	401–500	>500	Total
Social terms	It	13 (2)	10 (1)	10 (1)	28
	Fi	12 (1)	10 (1)	10 (0)	22
Nouns	It	53 (6)	48 (5)	44 (4)	44
	Fi	50 (3)	47 (2)	46 (1)	47
Predicates	It	18 (5)	24 (3)	25 (3)	13
	Fi	23 (3)	27 (2)	27 (1)	18
Closed-class words	It	8 (4)	8 (3)	11 (3)	9
	Fi	8 (1)	8 (2)	9 (1)	6
FT		<50	50–100	101–200	201–300
Social terms	It	52 (14)	41 (7)	21 (5)	15 (2)
	Fi	52 (17)	25 (8)	20 (4)	13 (2)
Nouns	It	34 (11)	39 (10)	50 (8)	53 (7)
	Fi	33 (15)	54 (8)	53 (7)	53 (5)
Predicates	It	2 (2)	8 (7)	15 (5)	16 (6)
	Fi	7 (5)	10 (6)	15 (6)	21 (3)
Closed-class words	It	11 (7)	8 (4)	8 (5)	8 (3)
	Fi	4 (6)	6 (3)	6 (3)	6 (2)
		301–400	401–500	>500	Total
Social terms	It	12 (1)	10 (1)	10 (1)	18
	Fi	11 (1)	11 (1)	10 (0)	19
Nouns	It	52 (5)	48 (5)	44 (3)	48
	Fi	51 (4)	46 (3)	45 (2)	48
Predicates	It	19 (5)	23 (3)	25 (2)	17
	Fi	24 (3)	27 (2)	27 (1)	20
Closed-class words	It	8 (3)	9 (4)	12 (3)	9
	Fi	7 (2)	9 (2)	10 (2)	4

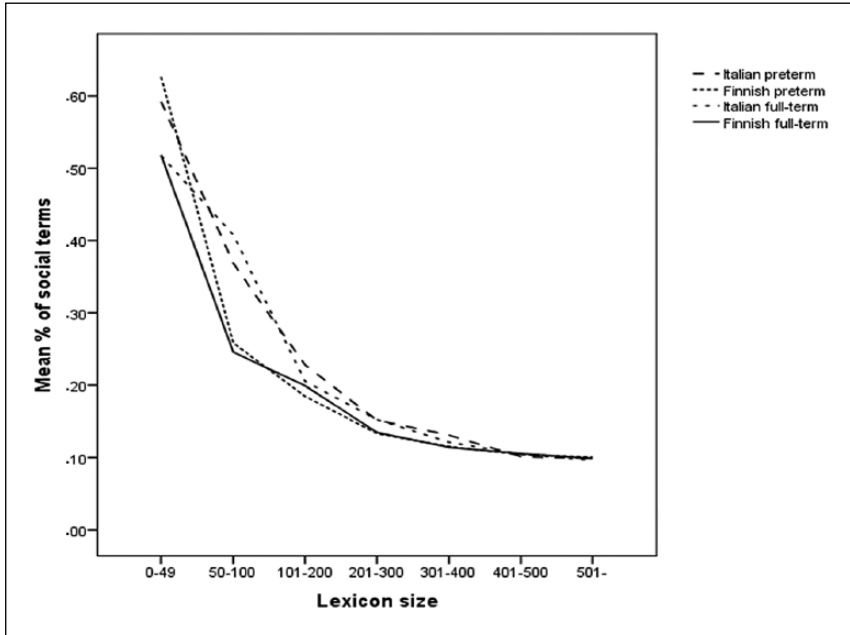


Figure 1. The percentage of social terms in the lexicons of the Italian and Finnish very preterm and full-term children. Mean percentages of the lexicon-size sub-groups are presented.

interaction between birth status and lexicon size was not significant, and neither was that between native language and lexicon size.

Closed-class words. No significant effect of birth status on the percentage of closed-class words was found, but there was a significant native-language effect, $F(1, 484) = 29.44$, $p < .001$: the percentage of closed-class words was significantly higher in the lexicons of the Italian children than in those of the Finnish children (difference: 2%; 95% CI: 0.02–0.03). The effect of lexicon size on the percentage of closed-class words was also significant, $F(6, 484) = 4.80$, $p < .001$, although no significant differences were found between contiguous lexicon-size sub-groups. The interaction between birth status and lexicon size was not significant, but that between native language and lexicon size was, $F(6, 484) = 4.31$, $p < .001$: the percentage of closed-class words was higher in the Italian than in the Finnish children in the following lexicon-size sub-groups: <50, 50–100 and 201–300 words, the respective differences being 7% (95% CI: 0.05–0.09, $p < .001$), 5% (95% CI: 0.02–0.07, $p < .001$) and 2% (95% CI: 0.00–0.04, $p = .04$).

The effect of ME on the lexicon in Italian and Finnish very preterm children

The ANOVA revealed a non-significant main effect of ME on lexicon size among the Italian very preterm children, but a significant effect of ME on the percentage of nouns,

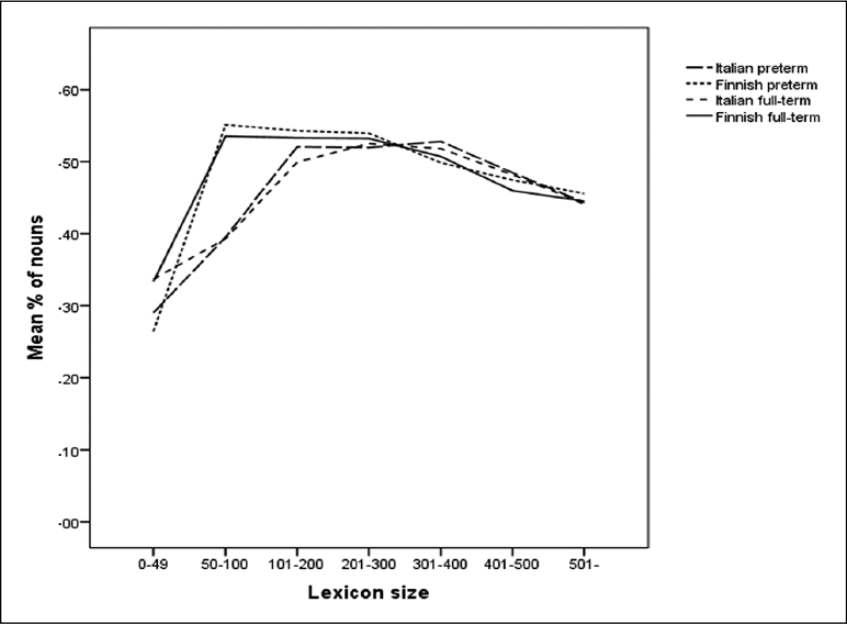


Figure 2. The percentage of nouns in the lexicons of the Italian and Finnish very preterm and full-term children. Mean percentages of the lexicon-size sub-groups are presented.

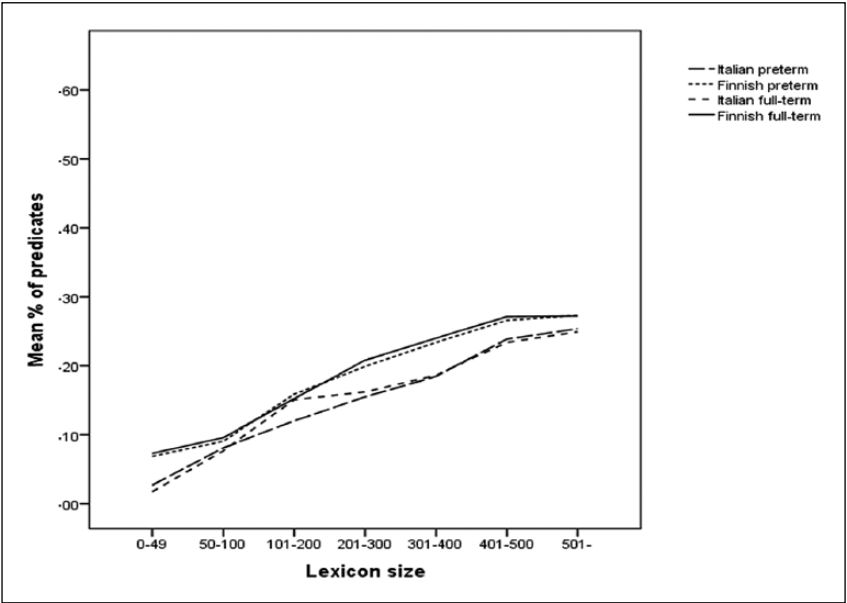


Figure 3. The percentage of predicate terms in the lexicons of the Italian and Finnish very preterm and full-term children. Mean percentages of the lexicon-size sub-groups are presented.

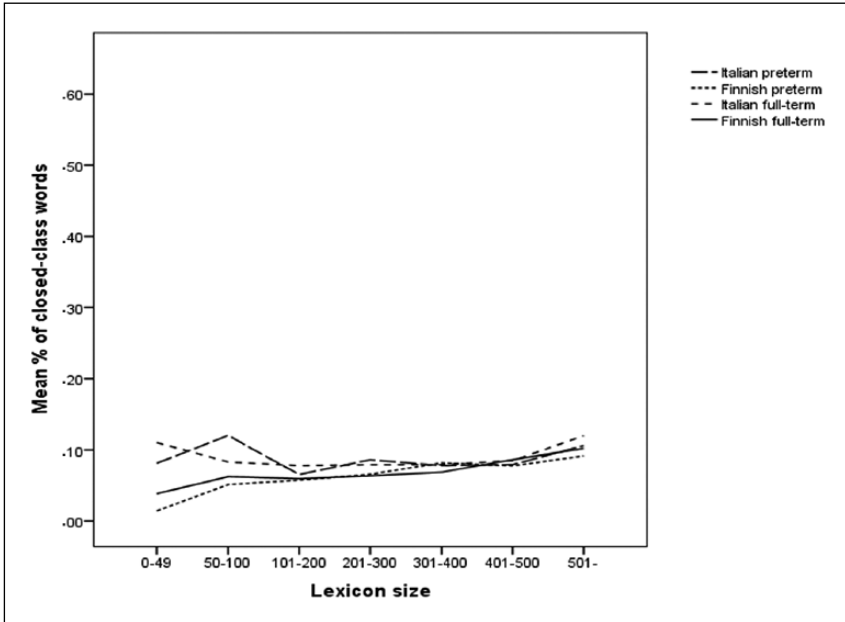


Figure 4. The percentage of closed-class words in the lexicons of the Italian and Finnish very preterm and full-term children. Mean percentages of the lexicon-size sub-groups are presented.

$F(2, 122) = 3.30, p = .04$. There were 8% more nouns in the lexicons of the children whose mothers had further as opposed to compulsory education (95% CI: 0.00–0.17, $p < .05$), as well in the lexicons of children whose mothers had secondary as opposed to compulsory education (95% CI: 0.00–0.17, $p < .05$). There was no significant difference in the percentage of nouns between the children whose mothers had a secondary as opposed to further education.

According to the ANOVA, ME had a significant effect on lexicon size among the Finnish very preterm children, $F(2, 109) = 4.42, p = .01$. The lexicon sizes of those children whose mothers had further education were 134 words larger than those with a mother educated to the compulsory level (95% CI: 22–246, $p = .01$). There were no significant differences in lexicon size between children whose mothers had a secondary education as opposed to compulsory education, or a secondary education as opposed to further education. Furthermore, ME significantly affected the percentages of the following lexical categories: social terms, $F(2, 109) = 7.89, p < .001$, nouns, $F(2, 108) = 6.88, p = .002$, and closed-class words, $F(2, 109) = 3.85, p = .02$. The lexicons of the Finnish very preterm children whose mothers had further education contained 24% fewer social terms (95% CI: 0.09–0.39, $p < .001$), 14% more nouns (95% CI: 0.05–0.23, $p = .001$), and 3% more closed-class words (95% CI: 0.00–0.05, $p = .02$) than those of their counterparts whose mothers had a compulsory education. In addition, the lexicons of children whose mothers had a secondary education contained 22% fewer social terms (95% CI: 0.06–0.38, $p = .00$) and 12% more nouns (95% CI: 0.02–0.22, $p = .01$) than those of their counterparts whose mothers had a compulsory education. There was no

significant difference in the percentage of social terms, nouns or closed-class words between the lexicons of children whose mothers had further education as opposed to a secondary education.

Discussion

The primary aim of this cross-linguistic study was to investigate the possible effect of the native language on lexical composition among very preterm children acquiring Italian or Finnish. Even though the very preterm children in both language groups had smaller lexicons than the full-term children, a comparable shift from early social terms to referencing, predication and early grammar was detected among the Italian and Finnish preterm children and their full-term peers. ME had a significant effect on lexical skills at 24 months, especially among the Finnish very preterm children.

The vocabulary of very preterm children

The vocabulary of the very preterm children was less extensive than that of the controls regardless of the native language. This confirms the results of previous analyses run on single language samples (Foster-Cohen et al., 2007; Kern & Gayraud, 2007; Sansavini, Savini, et al., 2011; Stolt et al., 2013; see however Perez-Pereira et al., 2014). Furthermore, given that the data on preterm children were collected at two years of corrected age and that there was still a significant difference in lexicon size between the preterm and full-term children, the finding supports the view that age correction is needed in the early clinical follow-up of very preterm children (cf. Cattani et al., 2010).

A significant novel finding of the present study was that the native language had no major effect on lexical composition specifically among Italian/Finnish very preterm children. This was the first cross-linguistic study to focus on the possible effect of the native language on the lexical composition of preterm children. However, the result is in line with earlier findings run on single language samples indicating that lexical composition of preterm children is comparable to that of full-term controls when lexicon size is taken into consideration in the analysis (Stolt et al., 2007; Stolt et al., 2009; see also Kern & Gayraud, 2007; Schults, Tulviste, & Haan, 2013). Another novel finding was that the lexical composition of the Italian very preterm children was comparable to that of the Italian controls at 24 months. With regard to Finnish very preterm children, the present study verifies earlier results (Stolt et al., 2007; Stolt et al., 2009), but now in a larger group of children. In conclusion, the results of the present cross-linguistic study show that the native language does not significantly influence the lexical composition of very preterm children. In other words, the results indicate that early lexical composition is a robust phenomenon that is not strongly affected by very preterm birth.

Minor differences in the lexical composition were detected between the preterm and full-term children: the percentage of social terms in the small lexicons (≤ 50 words) of preterm children in both language groups was higher than that of their controls. This might be due to the high variation in lexical composition at this developmental stage. Alternatively, it may be that, as preterm children have been shown to have weaker

language skills at 24 months than their full-term controls (e.g. Foster-Cohen et al., 2007), they might maintain a small lexicon composed primarily of social terms linked to daily routines for a longer period. Another possibility is that the parents of preterm children tend to focus on routines and baby-talk for a longer time than the parents of full-term children (Suttora & Salerni, 2011).

The effect of the native language on the lexical composition of Italian and Finnish children

The lexical composition was generally highly comparable in all four groups under study, regardless of the native language. The implication of this is that the composition of the early lexicon is not language-specific in broad terms. The few earlier cross-linguistic studies have reported comparable findings (Caselli et al., 1995; Caselli et al., 1999; see also Bornstein et al., 2004; Conboy & Thal, 2006), although they did not include a Finno-Ugric language. More research is needed to gain further information on the possible universality of lexicon composition. The present findings also give further support for the natural partition hypothesis (Gentner & Boroditsky, 2001): the percentage of nouns was higher than the percentage of predicate terms in all the lexicon-size sub-groups and in both language groups studied (cf. Bornstein et al., 2004).

Some language-specific features in the lexical compositions were detected among the Italian and Finnish children. The fact that these features were found in two groups (i.e. in very preterm and full-term children) of Italian/Finnish children verified the findings. The percentage of social terms was higher in the small lexicons of the Italian children (50–100 words) than in those of the Finnish children. A comparable result was reported in a previous study (Caselli et al., 1999): the percentages of social terms in lexicons of 50–100 words were roughly 39% and 27%, respectively, among Italian children and children acquiring American English. The respective percentages in the present study were 41% and 25% among Italian and Finnish controls. It is possible that Italian children may acquire early social terms more actively because Italian parents are likely to use such words more actively than Finnish parents (cf. Caselli et al., 1999). Finnish people tend to tolerate silence, whereas Italians are often perceived as lively and social. It was shown in a recent study (De Geer, Tulviste, Mizera, & Tryggvason, 2002), for example, that Finnish parents were less talkative than Swedish parents while socializing with their children at dinnertime. This acknowledged difference between Italian and Finnish people may be reflected in early parent–child interaction, and hence in early lexical acquisition. On the other hand, there were higher percentages of nouns in the small lexicons (50–100 words) of the Finnish children than in those of the Italian children, and the Finnish children were also more active in acquiring predicate terms. An earlier comparative study on the lexical composition of American English and Italian children (Caselli et al., 1999) reported no significant differences in the percentage of predicate terms in lexicons containing more than 100 words, whereas in the present study the differences persisted in large lexicons. There is clear evidence of a strong association between lexical and grammatical development at the end of the second year (e.g. Bates & Goodman, 1999), although the causal direction is not clear (i.e. if the lexical development influences grammatical development or vice versa; Bates & Goodman, 1999). Children have

been shown to use grammatical information to learn the meaning of a novel word, especially verbs (syntactic bootstrapping; e.g. Naigles, 1990; see also Bates & Goodman, 1999). On the other hand, it was found in a recent study on Finnish children (Stolt et al., 2013) that nominal/verb lexicon sizes were significant predictors of the number of nominal/verb inflections acquired at 24 months. Previous findings from cross-linguistic studies also indicate that the association between lexicon and grammar may not be identical in different languages (Caselli et al., 1999; Thordardottir et al., 2002), and variations have been attributed to grammatical differences in the languages concerned. In sum, the differences between Italian and Finnish children in the percentages of nouns and predicate terms reported in the present study may well be attributable to differences in the structure (i.e. morphology) of the two languages. It may be that the more complex noun and verb morphology of Finnish encourages Finnish children to pay attention to the category of nouns and predicate terms from early on. Finally, the higher percentage of closed-class words in the small lexicons (<100 words) of the Italian children compared to the Finnish children could also be attributable to structural differences between the languages: articles and prepositions are used actively in Italian whereas there are no articles in Finnish, and many meanings expressed through the use of prepositions in Italian are conveyed by morphological inflections in Finnish.

The present cross-sectional data which were collected at 24 months further support the view that the development of lexical composition may be strongly connected to lexicon size and not necessarily to the age of the children (cf. Bates et al., 1994; Caselli et al., 1999; Conboy & Thal, 2006; Stolt et al., 2007). Indeed, lexicon size may be an even more significant factor in compositional development than age. It may be that as children acquire their first lexicon they build on what they have already learnt, despite their age. Onomatopoeic expressions and words connected to routine situations (i.e. social terms) are acquired first (Bates et al., 1994; Caselli et al., 1999). These very early words do not necessarily have an exact adult-like meaning. Nevertheless, children do practise the symbolic function of language by picking up an expression and connecting it to the reference while using early social terms. Nouns, which are conceptually basic and easily referenced, are acquired next (Gentner & Boroditsky, 2001). As Bates et al. (1994) suggest, children practise referencing during the period of active noun acquisition. In other words, children learn that words have a meaning and are connected to the surroundings in a specific manner. On the other hand, predicate terms may not be acquired actively before children have learned enough words to which they relate (Bates & Goodman, 1999). Finally, closed-class words have a strong linguistic meaning which cannot be understood independently of language (Caselli et al., 1999; Gentner & Boroditsky, 2001). It may be that children cannot acquire these words before they have acquired enough content words and before they can understand enough language structures sufficiently well (cf. Bates & Goodman, 1999; Bates et al., 1994; Caselli et al., 1999; Gentner & Boroditsky, 2001).

The effect of ME on the vocabulary in very preterm children

ME had more effect on the lexical skills at 24 months among the Finnish very preterm children than in the Italian group. The Finnish mothers of preterm children also had a higher level of education in general. The effect of parental education is not always evident

(e.g. Kern & Gayraud, 2007; Ortiz-Mantilla et al., 2008). Nevertheless, the finding of a clear effect of ME on lexical skills in a group with mothers with a higher educational level than in the other group implies that early mother–child interaction could support early lexical acquisition among preterm children (Sansavini, Guarini, & Caselli, 2011; Stolt et al., 2014). However, interpretations of this finding should take into account the fact that in the present study the effect of ME was analysed only in the preterm children. A previous study (Stolt et al., 2007) reported a significant effect of ME on lexicon size in a preterm group, but not in a group of full-term children. Hence, it is possible that the effect of ME on both lexicon size and the percentages of the lexical categories differed between the groups of preterm and full-term children in the present study as well. Because ME information was not available for all the full-term children, its influence on lexicon size and lexical categories in this group could not be analysed in the present study.

Conclusion

Even though the lexicons of the very preterm children were smaller than those of the controls at 24 months of age in both language groups, the native language had no major effect on the early development of lexical categories in Italian and Finnish very preterm children, except for those with very small vocabularies. ME was associated with lexical composition especially in the Finnish very preterm children. The present findings indicate that lexical composition is not strongly affected by preterm birth. The results also imply that lexical composition is a robust phenomenon, which is connected to lexicon size and is not language-specific when analysed in broad terms, although some language-specific features were also detected.

Acknowledgements

We are grateful to the parents and their infants for their participation in the studies. We thank Silvia Vandini for her help with the medical examination of the Italian preterm infants, Cristina Fabbri for her help coding the data on the Italian preterm children, Patrizio Pasqualetti for his help managing the database on the Italian full-term controls; Petriina Munck, Riikka Korja, Anniina Peltola, Annika Lind and Timo Tuovinen for assessing the cognitive development of the Finnish children. This study is part of the PIPARI study. The members of the PIPARI study group are Satu Ekblad, Eeva Ekholm, Leena Haataja, Mira Huhtala, Pentti Kero, Riikka Korja, Harry Kujari, Helena Lapinleimu, Liisa Lehtonen, Marika Leppänen, Hanna Manninen, Jaakko Matomäki, Jonna Maunu, Petriina Munck, Pekka Niemi, Pertti Palo, Riitta Parkkola, Jorma Piha, Annika Lind, Liisi Rautava, Päivi Rautava, Milla Ylijoki, Hellevi Rikalainen, Katriina Saarinen, Matti Sillanpää, Suvi Stolt, Anniina Väliaho, Päivi Tuomikoski-Koiranen, Timo Tuovinen and Tuula Äärimaa.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: **Italian sample:** Italian Ministry of Education, University and

Research. Grant number: year 2008 - prot. 2008J2WEEK. Grant title: ‘Gestures and language in children with atypical and at-risk developmental profiles: relationships among competences, mother-child interaction modalities and proposals of intervention’; Coordinator: Alessandra Sansavini. Grant number: year 2007- STRAT07SAN. Grant title: ‘Early communicative-linguistic and cognitive abilities: risks linked to preterm birth’; Coordinator: Alessandra Sansavini.

Finnish sample: This article was written during a research period funded by the Academy of Finland (grant awarded to Suvi Stolt; grant number: 139124. Project title: “The language development of very low birth weight children - early development, background variables and predictive value”).

The main funders of the PIPARI study are: the Emil Aaltonen Foundation, the Academy of Finland, C.G Sundells stiftelse and the Fund for Neonatal Research in Southwestern Finland Fund.

References

- Bates, E., & Goodman, J. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–80). London, England: Lawrence Erlbaum.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, S., ... Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, *21*, 85–123.
- Bornstein, M., Cote, L., Maital, S., Painter, K., Park, S. Y., Pascual, L., ... Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean and American English. *Child Development*, *75*, 1115–1139.
- Caselli, 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*, 159–199.
- Caselli, C., Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, *26*, 69–111.
- Caselli, C., Pasqualetti, P., & Stefanini, S. (2007). *Parole e frasi nel ‘Primo vocabolario del bambino’*. *Nuovi dati normativi fra 18 e 36 mesi e forma breve del questionario* [The Italian MB-CDI – Words and Sentences complete and short forms: Normative data on children from 18 to 36 months]. Milano, Italy: FrancoAngeli.
- Cattani, A., Bonifacio, S., Fertz, M., Iverson, J., Zocconi, E., & Caselli, C. (2010). Communicative and linguistic development in preterm children: A longitudinal study from 12 to 24 months. *International Journal of Language and Communication Disorders*, *45*, 162–173.
- Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, *22*, 497–529.
- Clark, E. (2003). *First language acquisition*. Cambridge, UK: Cambridge University Press.
- Conboy, B., & Thal, D. (2006). Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development*, *77*, 712–735.
- De Geer, B., Tulviste, T., Mizera, L., & Tryggvason, M.-T. (2002). Socialization in communication: Pragmatic socialization during dinnertime in Estonian, Finnish and Swedish families. *Journal of Pragmatics*, *34*, 1757–1786.
- Dietz, P., Callaghan, W., Smith, R., & Sharma, A. (2009). Low pregnancy weight gain and small for gestational age: A comparison of the association using 3 different measures of small for gestational age. *American Journal of Obstetrics & Gynecology*, *199*, e1–e7.
- Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5).
- Foster-Cohen, S., Edgin, J., Champion, P., & Woodward, L. (2007). Early delayed language development in very preterm infants: Evidence from the MacArthur–Bates CDI. *Journal of Child Language*, *34*, 655–675.

- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, thought and culture* (pp. 301–333). London, England: Lawrence Erlbaum.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and word learning. In M. Bowerman, & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge, UK: Cambridge University Press.
- 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, 523–549.
- Karlsson, F. (1999). *Finnish: An essential grammar*. London, England: Routledge.
- Kern, S., & Gayraud, F. (2007). Influence of preterm birth on early lexical and grammatical development. *First Language*, 27, 159–173.
- Lyytinen, P. (1999). *Varhaisen kommunikaation ja kielen kehityksen arviointimenetelmä* [The Finnish, long form version of the MacArthur–Bates Communicative Development Inventories]. Jyväskylä, Finland: Niilo Mäki Instituutti.
- Maital, S., Dromi, E., Sagi, A., & Bornstein, M. (2000). The Hebrew Communicative Development Inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language*, 27, 43–67.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374.
- Nespor, M., & Bafile, L. (2008). *I suoni del linguaggio* [Speech sounds]. Bologna, Italy: Il Mulino.
- Ortiz-Mantilla, S., Choudhury, N., Leevers, H., & Benasich, A. (2008). Understanding language and cognitive deficits in very low birth weight children. *Developmental Psychobiology*, 50, 107–126.
- Perez-Pereira, M., Fernandez, P., Gomez-Taibo, M., & Resches, M. (2014). Language development of low risk preterm infants up to the age of 30 months. *Early Human Development*, 90, 649–656.
- Rademaker, K., Uiterwaal, C., Beek, F., van Haastert, I., Liefink, A., Groenendaal, F., ... de Vries, L. S. (2005). Neonatal cranial ultrasound versus MRI and neurodevelopmental outcome at school age in children born preterm. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 90, F489–F493.
- Reiman, M., Kujari, H., Maunu, J., Parkkola, R., Rikalainen, H., Lapinleimu, H., ... PIPARI Study Group. (2008). Does placental inflammation relate to brain lesions and volume in preterm infants? *Journal of Pediatrics*, 152, 642–647.
- Sansavini, A., Bello, A., Guarini, A., Savini, S., Alessandroni, R., Faldella, G., & Caselli, M. C. (2015). Noun and predicate comprehension/production and gestures in extremely preterm children at two years of age: Are they delayed? *Journal of Communication Disorders*, 58, 126–142.
- Sansavini, A., Guarini, A., Alessandroni, R., Faldella, G., Giovanelli, G., & Salvioli, G. (2006). Early relations between lexical and grammatical development in very immature Italian preterm. *Journal of Child Language*, 33, 199–216.
- Sansavini, A., Guarini, A., & Caselli, M. (2011). Preterm birth: Neuropsychological profiles and atypical developmental pathways. *Developmental Disabilities Research Reviews*, 17, 102–113.
- Sansavini, A., Guarini, A., & Savini, S. (2011). Linguistic and cognitive delays in very preterm infants at 2 years: General or specific delays? *Revista de Logopedia, Foniatria y Audiologia*, 31, 133–147.
- Sansavini, A., Guarini, A., Savini, S., Broccoli, S., Justice, L., Alessandroni, R., & Faldella, G. (2011). Longitudinal trajectories of gestural and linguistic abilities in very preterm infants in the second year of life. *Neuropsychologia*, 49, 3677–3688.
- Sansavini, A., Savini, S., Guarini, A., Broccoli, S., Alessandroni, R., & Faldella, G. (2011). The effect of gestational age on developmental outcomes: A longitudinal study in the first two years of life. *Child: Care, Health & Development*, 37, 26–36.

- Schults, A., Tulviste, T., & Haan, E. (2013). Early vocabulary in full term and preterm Estonian children. *Early Human Development, 89*, 721–726.
- Stolt, S., Haataja, L., Lapinleimu, H., & Lehtonen, L. (2008). Early lexical development of Finnish children – A longitudinal study. *First Language, 28*, 259–279.
- Stolt, S., Haataja, L., Lapinleimu, H., & Lehtonen, L. (2009). Early lexical development of prematurely born very-low-birth-weight children, and its relations to language skills at 2;0. *Journal of Communication Disorders, 42*, 107–123.
- Stolt, S., Klippi, A., Launonen, K., Munck, P., Lehtonen, L., Lapinleimu, H., ... PIPARI Study Group. (2007). Size and composition of lexicon in prematurely born very low birth weight and full-term Finnish children at two years of age. *Journal of Child Language, 34*, 283–310.
- Stolt, S., Korja, R., Matomäki, J., Lapinleimu, H., Haataja, L., & Lehtonen, L. (2014). Early relations between language development and the quality of mother–child interaction in very-low-birth-weight children. *Early Human Development, 90*, 219–225.
- Stolt, S., Matomäki, J., Haataja, L., Lapinleimu, H., Lehtonen, L., & PIPARI Study Group. (2013). The emergence of grammar in very-low-birth-weight Finnish children at two years of age. *Journal of Child Language, 40*, 336–357.
- Suttora, C., & Salerni, N. (2011). Maternal speech to preterm infants during the first 2 years of life: Stability and change. *International Journal of Language and Communication Disorders, 46*, 464–472.
- Tardiff, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*, 492–504.
- Thordardottir, E., Weismer, S., & Evans, J. (2002). Continuity in lexical and morphological development in Icelandic and English-speaking 2-year-olds. *First Language, 22*, 2–28.

Appendix A

Basic characteristics of Italian and Finnish

Italian. Italian is a Romance language (Bornstein et al., 2004; Caselli et al., 1995) with 30 phonemes (23 consonants and 7 vowels; Nespor & Bafille, 2008). Both bound and free morphemes are used. Bound morphemes are used to mark nouns, adjectives and verbs for number and gender, and verbs for tense, aspect and mood. There are several tenses for verbs for each mood (for indicative and subjunctive: present and past, simple and compound; for conditional: gerund and infinitive; for participle: present and past; for imperative: present and future). Free morphemes are used to mark the basic syntactic relations between the words in the sentences. In particular, both definite and indefinite articles are used to mark nouns, adjectives and pronouns for number and gender. The basic word order is subject-verb-object (SVO) but this order may change for pragmatic reasons (Bornstein et al., 2004; Caselli et al., 1995).

Finnish. Finnish is a Finno-Ugric language which has 21 phonemes (13 consonants and 8 vowels), fewer than in most European languages (Karlsson, 1999). The basic characteristics of the morpho-syntax is the addition of bound morphemes and suffixes to word stems. Both nominal words (nouns, adjectives, pronouns and numerals) and verbs are inflected morphologically. Nominal words are inflected by means of 15 case endings. Verb forms are built up by adding endings to stems and there is a subject-verb agreement.

There are three grammatical persons, each of which can be expressed in the singular and the plural. In addition, there are four tense forms (two simple: present and past; two compound: perfect, pluperfect). No articles are used. The word order is relatively free, although typically SVO.

Appendix B

Number of words listed per class and category in the Italian and the Finnish version of the MacArthur–Bates Communicative Development Inventory (ItCDI; FinCDI). The percentages of items calculated from the total number of items on the ItCDI/FinCDI are also presented.

Categories	ItCDI n (%)	FinCDI n (%)
<i>Social terms (total)</i>	66 (10)	59 (10)
Sound effects	12 (2)	13 (2)
People	29 (4)	24 (4)
Games and routines	25 (4)	22 (4)
<i>Nouns (total)</i>	281 (42)	255 (43)
Animal names	43 (6)	38 (6)
Vehicles	14 (2)	12 (2)
Toys	18 (3)	15 (3)
Food and drink	68 (10)	58 (10)
Clothing	28 (4)	28 (5)
Body parts	27 (4)	24 (4)
Rooms and furniture	33 (5)	32 (5)
Small household items	50 (8)	48 (8)
<i>Predicates (total)</i>	166 (25)	160 (27)
Action words	103 (15)	106 (18)
Descriptive words	63 (9)	54 (9)
<i>Closed-class words (total)</i>	92 (14)	71 (12)
Pronouns	23 (3)	24 (4)
Question words	7 (1)	8 (1)
Prepositions	17 (3)	20 (3)
Articles and quantifiers	10 (2)	9 (2) ^a
Particles	6 (1)	10 (2)
Auxiliary verbs	15 (2)	–
Words about space and quantity	14 (2)	–
<i>Others (total)</i>	65 (10)	50 (8)
Outside things + Places to go	31 (5) + 22 (3) = 53 (8)	38 (6) ^b
Words about time	12 (2)	12 (2)
Total number of items	670	595

^aThere are no articles in the FinCDI.

^bThe categories 'Outside things' and 'Places to go' are combined in the FinCDI.