The Norwegian Communicative Development Inventories – reliability, main developmental trends and gender differences

Hanne Gram Simonsen, University of Oslo

Kristian E. Kristoffersen, University of Oslo

Dorthe Bleses, University of Southern Denmark

Sonja Wehberg, Odense University Hospital

Rune N. Jørgensen, University of Southern Denmark

Correspondence concerning this article should be addressed to Hanne Gram Simonsen, MultiLing, Department of Linguistics and Scandinavian studies, University of Oslo, P.O.Box 1102, Blindern, N-0315 Oslo, Norway. Email: h.g.simonsen@iln.uio.no

Abstract

This paper presents results from a large population-based study of early

communicative development in Norwegian children using an adaptation of the

MacArthur-Bates Communicative Development Inventories, comprising 6,574

children between 8 and 36 months. Data were collected via the Internet. In

accordance with similar studies from other languages, we found that vocabulary

comprehension preceded vocabulary production, and that both use of gestures,

comprehension and production of vocabulary, and grammatical complexity

increased with age. Moreover, boys lagged behind girls in vocabulary production

and comprehension, in grammatical complexity, and in certain types of imitation

– this gender difference seems to come out more clearly in our data than in data

from other languages.

Keywords: MacArthur-Bates CDI, Norwegian, gender differences

The Norwegian Communicative Development Inventories – reliability, main developmental trends and gender differences

Development of language skills in small children is characterized both by an enormous complexity – increasing with age – and by extensive variation from one typically developing child to the next. It is well known that language skills are an important prerequisite for later academic achievements, reading being prominent among these. It is therefore important that we have knowledge about early linguistic development, as well as effective assessment instruments to help us identify those children who are at risk for language problems as early as possible. By identifying these at-risk children we will be able to support their linguistic development by effective interventions, thereby minimizing risk for later problems, for example with acquiring reading skills (see e.g. Scarborough, 2009, for a review of the relationship between early language skills and reading, and on the importance of early identification). Although large-scale assessment instruments are not strong in their predictiveness at case level, they are useful as a first step in identifying possible at-risk children. This paper reports findings from a large-scale population based study of communicative skills in children aged 8 to 36 months learning Norwegian, focusing in particular on gestures, vocabulary and grammar. Data for the study were collected by means of the adaptation into Norwegian of the MacArthur-Bates Communicative Development Inventories.

The target language for the children participating in the study is Norwegian. This is a Germanic language spoken by approximately 5 million citizens in Norway. Norwegian is closely related to the other Scandinavian languages, Swedish and Danish – those languages have in general been considered mutually intelligible, but

seem now to be drifting apart. More distant relatives are Icelandic and Faroese. The Norwegian lexicon is predominantly Germanic, but contains loan words from other languages, Germanic as well as non-Germanic. Morphologically, Norwegian is slightly more complex than English. Like English (and other Germanic languages) Norwegian verbs are divided into two main classes, weak ('regular') and strong ('irregular') – but there are two weak classes in Norwegian (and in the other Scandinavian languages) as opposed to one in English. Verbs are inflected for tense, mood and voice. Nouns have two or three gender classes, depending on dialect, and are inflected for number and definiteness. Adjectives are inflected for gender, number and definiteness.

Until now, research on the development of lexical and grammatical skills in infants and toddlers learning Norwegian has been sparse, and has for the most part consisted of single or multiple case studies (with a few exceptions, see below). In a survey of the acquisition of Scandinavian languages, Plunkett and Strömqvist (1992) analysed data from Swedish, Danish and Norwegian children between one and four years of age, providing an overview of the development across a range of grammatical phenomena, among them number and definiteness, gender and verb inflections.

As for Norwegian, Vanvik (1971) was to our knowledge the first to investigate first language acquisition in this context, with a longitudinal diary study of the speech and language development of his own daughter from birth to eight years of age. The primary focus of Vanvik's research was phonetic and phonological development, but he also made a few notes concerning vocabulary and grammar. Simonsen (1986, 1990) investigated the phonological development of three children aged between 2;0 and 4;1 acquiring Norwegian. Simonsen's focus was also on phonology, but she

included notes on the interaction between phonology and morphology, as well as lists of all the words produced by her participants, thus giving valuable information about early lexical and grammatical skills. More recently, Anderssen (2005, 2007) examined acquisition of prenominal definiteness markers, suffixal definiteness markers and pronouns in three children aged 1;8 - 3;3. Furthermore, Aukrust (1992) and Westergaard (2005) reported MLU values for two small groups of two-year-old children, which is relevant for grammatical complexity measures. In a more experimental vein Torkildsen and colleagues examined aspects of lexical processing in children aged 1;8 and 2;0 using ERP (event-related potentials) (Torkildsen, 2008; Torkildsen et al., 2006, 2007, 2008, 2009).

There are also a few studies based on parental questionnaires. Lorenzen and Pedersen (2001) and Ingvaldsen (2001) used a previous Norwegian adaptation of the MacArthur-Bates CDI,[1] and reported vocabulary measures for children at three different ages: at 1;2 (n=24), 1;7 (n=20), and 2;0 (n=22) years. Janson and colleagues (Janson, 2003; Janson & Smith, 2003; Janson & Squires, 2004) reported on a limited set of communicative skills in 1172 children aged 0;4 – 5;0 based on the Ages and Stages Questionnaires. Finally, Schjølberg et al. (2009) produced a preliminary report as part of the larger *The Norwegian Mother and Child Cohort Study,* which included questions concerning early communicative skills. A problem with this report, however, is that language skills were examined from a very general perspective.

This brief summary of earlier studies clearly identifies the need for more detailed large-scale studies within the area of early lexical and grammatical skills of children learning Norwegian. One way of meeting this need is by way of parental

---

[1] This previous adaptation was made by Lars Smith, professor of psychology at the University of Oslo. However, it was not published, nor normed for Norwegian.

reports. Parents are best-placed to observe their children's communicative skills, and it has therefore turned out that they are particularly reliable sources of knowledge about these skills. In addition, parental reports give information about linguistic skills across different situations, thus providing more representative data than can be obtained through structured tests or laboratory samples. Parental reports are also a cost-efficient means for assessing linguistic skills in children, in particular for the early phases of development. Therefore they are an invaluable tool for collecting the large amounts of data that are necessary for establishing population-based norms.

Parental reports also have their limitations. Neither phonological development nor frequency of different vocabulary items is assessed, nor is spontaneous speech distinguished from imitations (Bates, Dale, & Thal, 1995). Furthermore, there is always a risk that parents both over-report and under-report on their children's skills, although other areas of development seem to be more vulnerable to this than language (Fenson et al., 2007). In addition, at the lowest age-range (until 12 months) the predictive power of the CDI is limited, probably due to the large variation in children at this age. Thus caution must be taken in the use of the instrument for screening before 16-24 months (Fenson et al., 2000). Finally, the biased sampling generally found in surveys of this kind, in that parents with lower education are underrepresented, is important in evaluating children's performance, in particular when applying and interpreting scores for children from lower income families and at the youngest ages (Fenson, et al, 2000).

The *MacArthur-Bates Communicative Development Inventories (CDI)* (Fenson et al, 1994; 2007) is the most widely used parent report instrument today. It focuses on the development of gestures, vocabulary and grammar in infants and

toddlers up to the age of 2;6 to 3 years. The instrument was originally developed for children learning American English, but has been adapted into more than 50 different languages, spoken as well as signed (Dale & Penfold, 2011). For both linguistic and cultural reasons, these adaptations differ from the American original to a varying degree. Still, the fact that they have been adapted from the same original makes them useful for cross-linguistic comparison (Bleses et al, 2008b; Caselli et al, 1995; Caselli, Casadio, & Bates, 1999; Caselli, Monaco, Trasciani, & Vicari, 2008; Devescovi et al, 2005; Maital, Dromi, Sagi, & Bornstein, 2000; McBride-Chang et al, 2008; Tardif, Gelman, & Xu, 1999). In addition to typically developing children, the CDI instrument has also been used for studying language skills in children from atypical populations (see e.g. Caselli et al., 1998; Berglund, Eriksson & Johansson, 2001; Chilosi, Cipriani, Bertuccelli, Pfanner, & Cioni, 2001; Thal, Reilly, Seibert, Jeffries & Fenson, 2004).

Both validity and reliability have been found to be good for the CDI instrument (see Fenson et al., 2007 and Law & Roy, 2008 for reviews). Two recent studies, O'Toole and Fletcher (2010) and Trudeau and Sutton (2011), present thorough validations of the Irish and the Quebec French CDIs respectively, comparing the CDI results (vocabulary and grammar) with spontaneous speech from large samples of children. They both found high correlations both for concurrent and predictive validity.

This continually growing body of CDI-based research has focused both on developmental trends, on variation, on the composition of early vocabulary, and on the relationship between gestures and vocabulary, between receptive and productive vocabulary and between vocabulary and grammar. Generally, the following results

have emerged from this research: Similar developmental trends and range of variation across languages; communicative gestures and receptive vocabulary are strongly correlated; productive vocabulary and grammatical complexity are strongly correlated; receptive and productive vocabulary are moderately correlated; and two developmental spurts, one in productive vocabulary somewhere between ages 1;4 and 1;8, and one in morphosyntactic development closely after, somewhere between ages 2;0 and 2;6 (for summaries, see e.g. Bates & Goodman, 1999; Bleses et al., 2008b; Dale & Goodman, 2005; Law & Roy, 2008; Hall, Holler, Rumney & Kidd, 2013).

There is discussion in the literature whether such accelerations in productive vocabulary and morphosyntactic development should actually be characterized as 'spurts', and on an individual basis not all children exhibit such spurts (Bates et al., 1995; Tomasello, 2003). More steady increases in vocabulary acquisition are also found (Trudeau & Sutton, 2011), and in other cases both accelerations and plateaus are noted (Stolt, Haataja, Lapinleimu, & Lehtonen, 2008). However, the temporal asynchrony and close correlation between acquisition of productive vocabulary and grammar seem well documented (Bates & Goodman, 1997, 1999).

The question of gender differences in early communicative development has also been addressed in several studies based on parental reports. Generally, a small advantage for girls has been found for many languages, but not necessarily across all language skills. Better performance by girls has been reported most consistently for word production, more rarely for word comprehension, and in some cases also for communicative gestures. On the other hand, several studies have found no gender differences at all (see Eriksson et al., 2012, for a review of findings from earlier studies). Eriksson et al. (2012) studied combined CDI data from more than 13, 000

children from 10 different European languages (Austrian German, Basque, Croatian, Danish, Estonian, French, Galician, Slovene, Spanish, and Swedish) and found a small, but consistent girl advantage in early gestures, in productive vocabulary, and in combining words, a difference that increased with age. This difference was found to be robust across languages and cultures.

For the present study we used an adaptation into Norwegian of the CDI (Kristoffersen et al., 2008) to investigate early communicative skills in typically developing monolingual children between 0;8 and 3 years of age. In an earlier paper (Kristoffersen et al., 2013), the methodology of the investigation, in particular the use of a web-based data collection, as well as the validity are discussed. The validity of the Norwegian adaptation was found to be comparable to that of CDI-studies on other languages. In the current paper we (1) analyse the reliability of the Norwegian CDI, before addressing two general questions: (2) What characterizes the development and variation in linguistic skills in infants and toddlers aged 0;8 – 3;0 learning Norwegian as measured by the Norwegian CDI?; and (3) Are there gender differences concerning these skills?

Method

The study is based on an adaptation into Norwegian of the *MacArthur-Bates Communicative Development Inventories* (CDI). It consists of two different forms, an infant form (*Words and Gestures*) covering development between 0;8 years and 1;4-1;8, years and a toddler form (*Words and Sentences*) covering the period from 1;4 to 2;6–3;0 years. The infant form assesses the first signs of understanding, productive skills like labelling and imitation, and the size of receptive and productive vocabulary, in addition to communicative actions and gestures. The toddler form has an extensive

vocabulary checklist assessing productive vocabulary, a section focusing on children's ability to talk about actions, objects and persons that are not present, as well as a section covering inflections, word combinations and grammatical complexity.

The Norwegian adaptation of the CDI

In 2006, a first version was constructed on the basis of the American original, and evaluated by a group of experts on early communicative development in Norwegian children within the fields of linguistics and psychology.[2] This version was tested in a pilot study in 2007, collecting parental report data from 17 children, six with the *Words and Gestures* form, and 11 with the *Words and Sentences* form, as well as a form on background information. Parents were asked to evaluate the items selected and suggest possible missing items. They were also asked to report the time spent on completing the forms, and evaluate the instructions given.

The parents all found the instructions easy to understand. They took between 10 and 80 minutes to complete the forms. Following the parents' suggestions, a few words (for example *pc*) were added and some were removed; but, in all, this version of the CDI was only slightly revised. Before constructing the web-based forms used in the present study, however, we revised the forms again, this time aiming to bring the Norwegian adaptation as close as possible to the Danish adaptation, in order to facilitate cross-linguistic comparison between two closely related languages with comparable grammatical systems but with quite different phonologies (see Kristoffersen et al. (2013) for details).

---

[2] Master of linguistics Kristin Wium made the draft for the first version of the Norwegian CDI. The expert group consisted of Lars Smith and Stephen von Tetzchner, professors of psychology at the University of Oslo, PhD (linguistics) Janne von Koss Torkildsen, and Hanne Gram Simonsen and Kristian E. Kristoffersen, professors of linguistics at the University of Oslo.

Participants

The children, who were all Norwegian citizens with the exact age of 0;8, 0;9, 0;10 ... or 3;0 years a few weeks after they received the invitation to participate in the study, were randomly selected from the official Norwegian birth register by Statistics Norway. In all, 20,400 families with children aged between eight months and three years were contacted.[3] Those who decided to take part in the study were directed to the website with the CDI forms, and by means of a username (the date of birth of the child) and a password they got access to the actual forms. The children's ages are calculated according to the date of completion of the forms by the parents.

Data were collected in two rounds, all from the same sample. Parents who had not responded in the first round after having received the invitation, were contacted by mail again after approximately two months. Those who answered then were reassigned to the month corresponding to the date of birth of their child. After the second round, a total of 7,555 forms had been completed, with 2,699 for *Words and Gestures* and 4,856 for *Words and Sentences*, yielding a response rate of 37%. The response rate varied between each monthly stage, with 22% as an extreme at the lowest end for the 8- month-olds, and 54% as an extreme at the highest end for the 21- month-olds.

For the child to be included in the final dataset, the following four criteria had to be met: (1) no frequent contact with other languages than Norwegian; (2) birth at full term (after week 36); (3) combined hospital stay should not exceed 4 weeks; (4) no serious, well-founded parental concern for the language development of the child.

---

[3] Since information about individuals would be handled in the study, the Norwegian Social Science Data Service (NSD) reviewed the methods for collecting and storing data and approved all procedures. The procedures were also evaluated and found appropriate by Statistics Norway.

Children with limited hearing because of frequent ear infections were not excluded, but profoundly deaf children were, as well as children who had physical or mental disadvantages in acquiring language. Also, children about whom day care personnel had raised concern were excluded from the final dataset. No more than 40 children were excluded by this fourth criterion.

The age of the child had to be between 0;8 and 1;8 months for the infants, and between 1;4 and 3;0 for the toddlers. On the website the parents were automatically directed to the appropriate form according to the child's date of birth. All children between 0;8 and 1;4 were directed to the *Words and gestures* form, while the children between 1;8 and 3;0 were directed to the *Words and sentences* form. The children between 1;4 and 1;8 were randomly directed to either the *Words and gestures* or the *Words and sentences* form, resulting in two groups of equal size.

Finally, at least one question in the form had to be answered. After applying these criteria, 981 of the original 7,555 children were excluded, so the final dataset consisted of 6,574 parental reports, 2,359 for *Words and gestures*, and 4,251 for *Words and sentences*.

The final sample of children was balanced with respect to gender (49% boys and 51% girls), and the sibling status of the children matched that of the child family population relatively well. (An overview of the participants, boys and girls, at each monthly stage is listed in the Appendix.) Furthermore, the participating families came from all main dialect areas of Norway, in a proportion that corresponded very closely (> 99%) to the general population in these areas. The sample was biased with respect to education, in that parents with higher education were overrepresented and parents with lower education underrepresented compared to the population of families with

children in Norway – see Table 1. This is an observation made in other CDI studies,

including the US and the Danish ones (Fenson et al., 2007, Bleses et al., 2008a).

However, the web-based method of data collection did not add to this bias: compared

to the Danish CDI study, which was paper based, both the response rates and the

education biases are comparable (see Kristoffersen et al (2013) for details).

TABLE 1. *Educational level of parents responding in the Norwegian CDI study*

Parent education	Norwegian CDI		Child family population in Norway
	N	%	%
Basic (< 9 years)	453	3	11
Short further education (12 years)	3949	30	32
Medium further education (14 - 16 years)	5524	42	37
Long further education (>16 years)	3134	24	16

Results

Reliability

We measured the reliability in terms of internal consistency and test-retest

reliability. Internal consistency was evaluated by calculating Cronbach's *alpha* for the

following four scores: (1) Number of words in receptive vocabulary (*CDI: Words and

Gestures*, 20 subscores, total number of items 395, $\alpha = 0.98$), (2) Number of words in

productive vocabulary (*CDI: Words & Gestures*, 20 subscores, total number of items

395, $\alpha = 0.98$) and (3) the same for *CDI: Words & Sentences (*22 subscores, total

number of items 731, $\alpha = 0.99$) and (4) Number of gestures (*CDI: Words & Gestures*,

5 subscores, total number of items 63, $\alpha = 0.90$). All results show high internal

consistency and are comparable to other published values (e.g., Fenson et al., 2007;

Bleses et al., 2008a).

To examine test-retest reliability, we made use of a self-elected sample –

parents were invited to complete additional CDI forms at later months. For 364

children included in the final infant dataset, there are two forms. For 35 pairs, the

second observation is more than two months later than the first; these are excluded in

the following (since reliability can only be evaluated if essentially the same is

measured – however, a child's ability might change rapidly). For the remaining 329

pairs, the median lag (in days) is 46, (minimum 31, 25% quantile 42, 75% quantile 52,

maximum 91). Test-retest Spearman correlations (r) are calculated for each age group

and each of the considered scales separately and presented together with the

corresponding number of observations (N) in Table 2.

TABLE 2. *Correlations between two observations of the same infant at different age points (n=329)*

		0;8-	0;9-	0;10-	0;11-	1;0-	1;1-	1;2-	1;3-	1;4-	1;5-	1;6-	1;7-
Comprehension score	N	42	44	37	31	36	38	22	21	15	21	15	7
	r	.74	.81	.79	.82	.84	.73	.85	.95	.64	.62	.76	.68
Production score	N	42	44	37	31	36	38	22	21	15	21	15	7
	r	.56	.47	.59	.71	.58	.75	.87	.93	.78	.82	.97	.86
Gestures score	N	41	44	37	31	35	38	22	21	15	21	15	7
	r	.68	.79	.62	.65	.85	.60	.73	.95	.39	.62	.81	.40

Correlations for the comprehension score lie between .62 and .95, which seems

satisfactory given the possibility of a child's rapid development within 46 days.

Correlations are somewhat lower for the production score (between .47 and .97),

especially in the beginning. Correlations for gestures lie between .60 and .95, with two exceptions of .39 and .40 at ages 1;4 and 1;7, respectively.

For 456 children included in the final toddler dataset, there are two forms available. For 59 pairs, the second observation is more than two months later than the first; these are excluded in the following. For the remaining 397 pairs, the median lag (in days) is 46, (minimum 30, 25% quantile 42, 75% quantile 52, maximum 85). Test-retest Spearman correlations (r) are calculated for each age group separately and presented together with the corresponding number of observations (N) in table 3.

TABLE 3. *Correlations between two observations of the same toddler at different age points (n=397)*

		1;4-	1;5-	1;6-	1;7-	1;8-	1;9-	1;10-	1;11-	2;0-	2;1-	2;2-	2;3-
Production score	N	19	17	24	29	40	44	44	24	22	21	20	17
	r	.88	.88	.84	.94	.86	.92	.95	.93	.90	.96	.87	.93
		2;4-	2;5-	2;6-	2;7-	2;8-	2;9-	2;10-	2;11-	3;0-			
Production score, cont.	N	8	14	10	12	11	10	9	2	8			
	r	.79	.94	.74	.84	.85	.40	.92	1.00	.79			

Correlations are constantly high (between .74 and 1.00) with an exception of .40 at age 2;9 (based on measurements of 10 children).

Developmental trends, with gender differences

Results for the developmental trends of the different scores are presented graphically by percentiles (10%, 25%, 50%, 75%, 90%) within the single age groups. To illustrate gender differences, the medians (50%-percentiles) are shown separately for boys and girls. In addition, results of a 2-way ANOVA are reported for each score

separately, where (continuous) age, gender and their interaction are included as covariates. Significance level is 5%. To give an impression of average effect sizes, effect sizes for age and gender are reported together with 95% confidence intervals based on a linear regression model without interaction.

For more detailed descriptive statistics, we refer to the database CLEX, found at http://www.cdi-clex.org (see Jørgensen, Dale, Bleses & Fenson, 2009). This is a website for a cross-linguistic database containing lexical data from adaptations of the MacArthur-Bates CDI, providing tools for a range of analyses within and across languages. The Norwegian CDI is included in the CLEX database.

Developmental trends: Infants
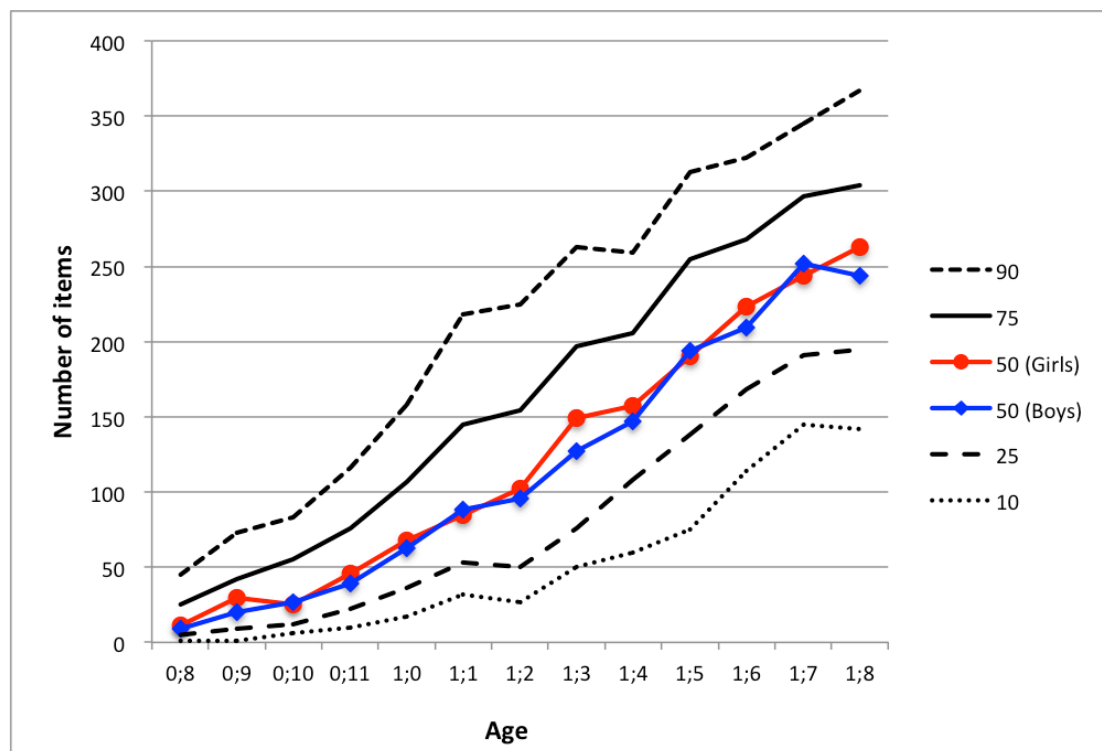
Figure 1. Percentiles of receptive vocabulary size by age and gender (CDI:WG; 395 items).

Figure 1 presents selected percentiles of receptive vocabulary size in infants by age and gender, comprising 395 items. Not surprisingly, the receptive vocabulary of Norwegian children increased noticeably with age, $F(12,2333) = 239.1, p<0.0001$. At age 0;8 half of the children did not understand more than 10 words (items), whereas 50% of the children one year older were reported to understand at least 257 words. Gender also had a significant effect, $F(1,2333) = 7.0, p=0.008$. The median number of words understood by girls was larger than the boys' median for almost all age groups. The difference of medians 'girls minus boys' ranged from -8 at age 1;7 to +19 one month later. The gender by age interaction was not significant, $F(12,2333) = 1.1, p=0.36.$ On average, a child understood 21 words (95% confidence interval: 20-22) more for each month, and girls understood 8 words (2-14) more than boys.
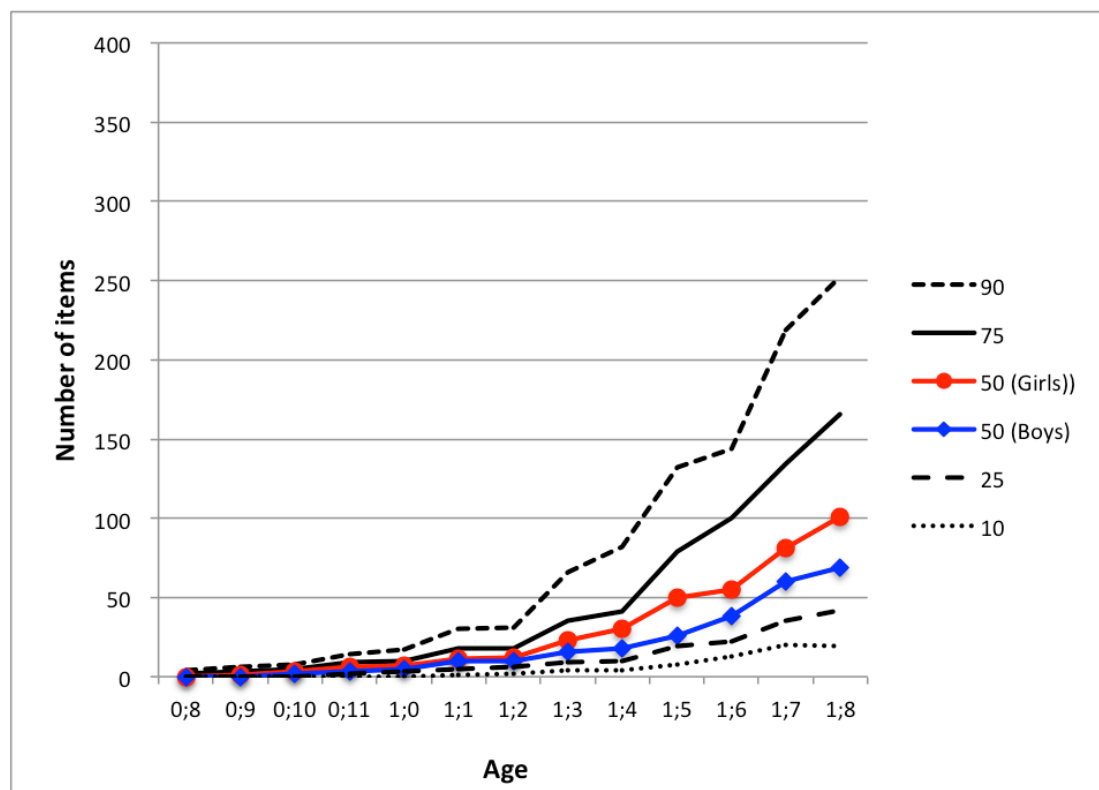

Figure 2. Percentiles of productive vocabulary size by age and gender (CDI:WG; 395 items)

Figure 2 presents selected percentiles of productive vocabulary size in infants by age and gender, comprising 395 items. Concerning the chronological relationship between vocabulary comprehension and production, it is clear from our data that comprehension precedes production: It was not until age 1;1 (4 months after comprehension) that half of the children were reported to have said at least 10 words. The productive vocabulary increased significantly with age, $F(12,2333) = 110.2$, $p<0.0001$ and gender, $F(1,2333) = 16.5$, $p=0.0001$. Girls started with a median of 0 and 1 at age 0;8 and 0;9, respectively, and reached a median score of 101 words at age 1;8. Boys began at the same level – median 0 at both 0;8 and 0;9, median of 2 words at 0;10 – and reached by a somewhat less steep increase a median level of 69 words at 1;8. Here, the gender by age interaction term was significant, $F(12,2333) = 2.1$, $p=0.043$. On average, a child said 9 words (8-10) more for each month, and girls said 8 words (4-12) more than boys.
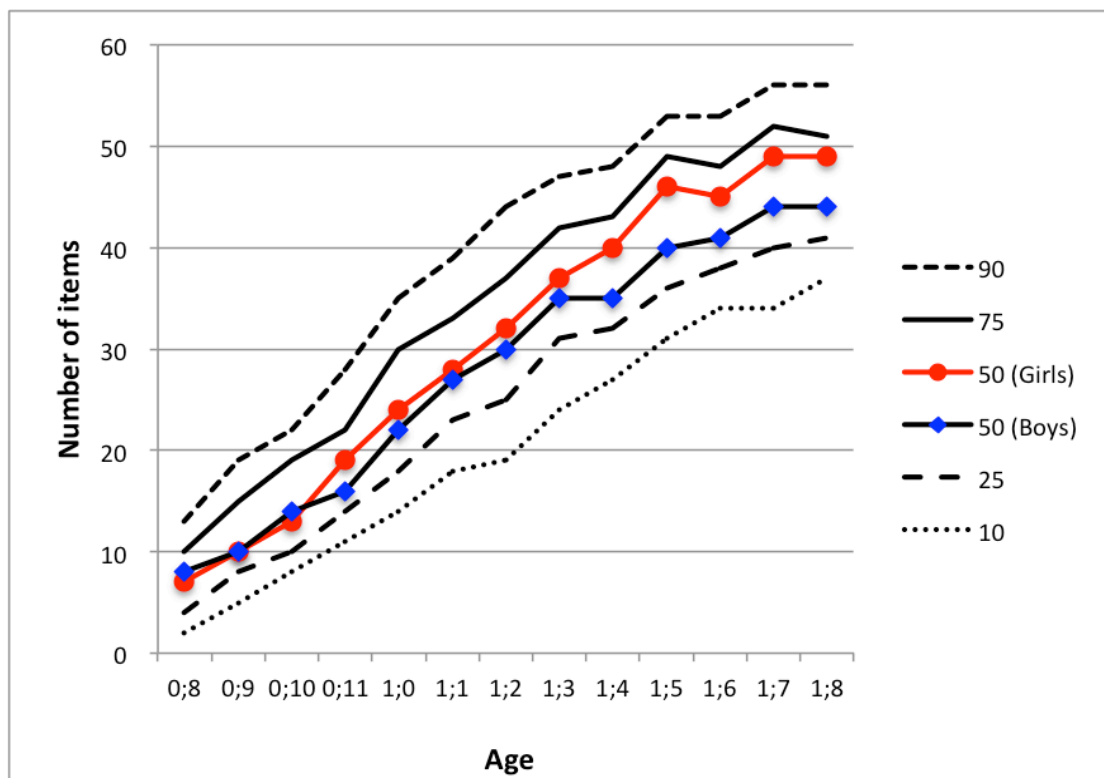
Figure 3. Percentiles of total number of communicative gestures by age and gender
(CDI:WG, 63 items)

Figure 3 presents selected percentiles of the total communicative gestures score in

infants by age and gender, comprising 63 items. The total number of communicative

gestures which the children were reported to use increased clearly with age,

F(12,2322) = 491.3, p<0.0001. There might, however, be a ceiling effect since most

of the curves seemed to flatten somewhat after age 1;5: whereas the (overall) median

increased from seven gestures at age 0;8 to 42 gestures at age 1;5, it only increased by

further four gestures up to 46 at age 1;8. The main effect of gender was significant,

F(1,2322) = 44.9, p<0.0001: the girls' median was larger than the boys' median at all

times after age 0;10 with the largest difference of six gestures at age 1;5 (before age

0;10, the difference was at most one gesture in favour of the boys). Since the gender

effect increased with age, the interaction term was significant, *F(12,2322) = 2.3,*

p=0.007. On average, a child used 3 gestures (3.2-3.4) more for each month, and girls

used 3 gestures (2-3) more than boys.
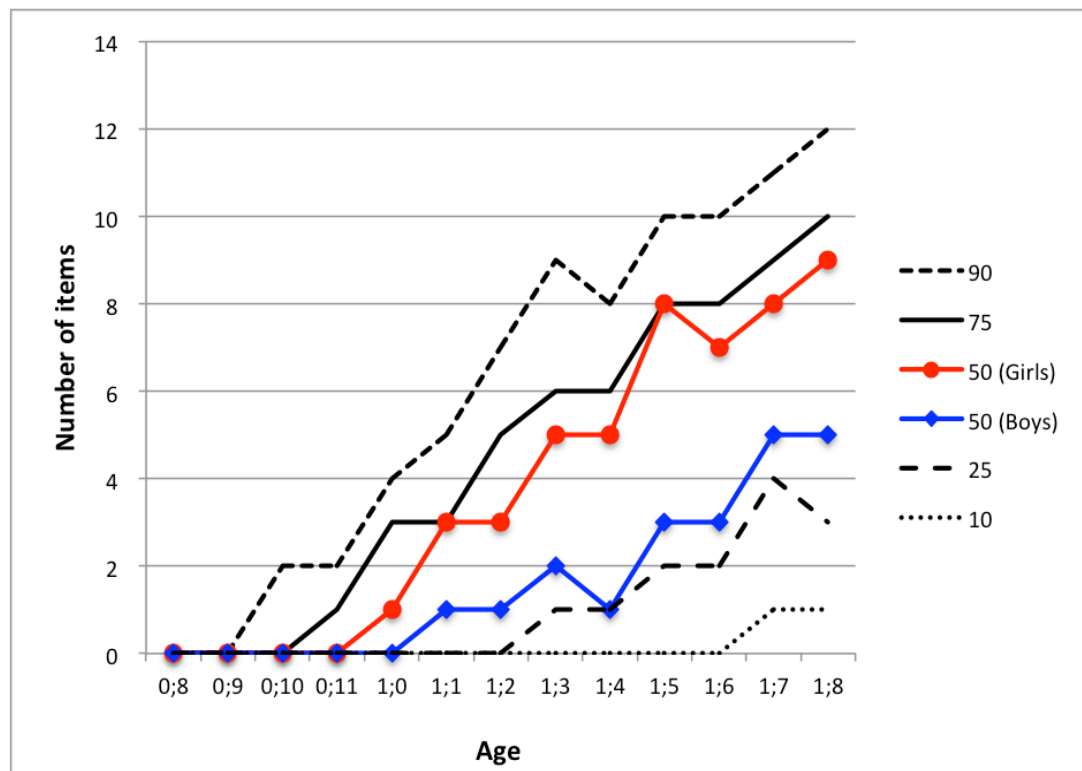
Figure 4. Percentiles of 'Pretending to be a parent' score by age and gender (CDI:WG;13 items)

Figure 4 presents selected percentiles of 'Pretending to be a parent' score in infants by age and gender, comprising 13 items. Results in this subcategory are comparable to results from the overall gestures score (see above), only more pronounced. The score increased with age, *F(12,2199) = 144.7, p<0.0001,* and the main effect of gender as well as the interaction was significant *(Gender: F(1,2199) = 290.4, p<0.0001, interaction: F(12,2199) = 14.2, p<0.0001).* On average, a child used 1 gesture (0.58-0.64) from this subcategory more for each month, and girls used 2 gestures (1.8-2.2) more than boys.
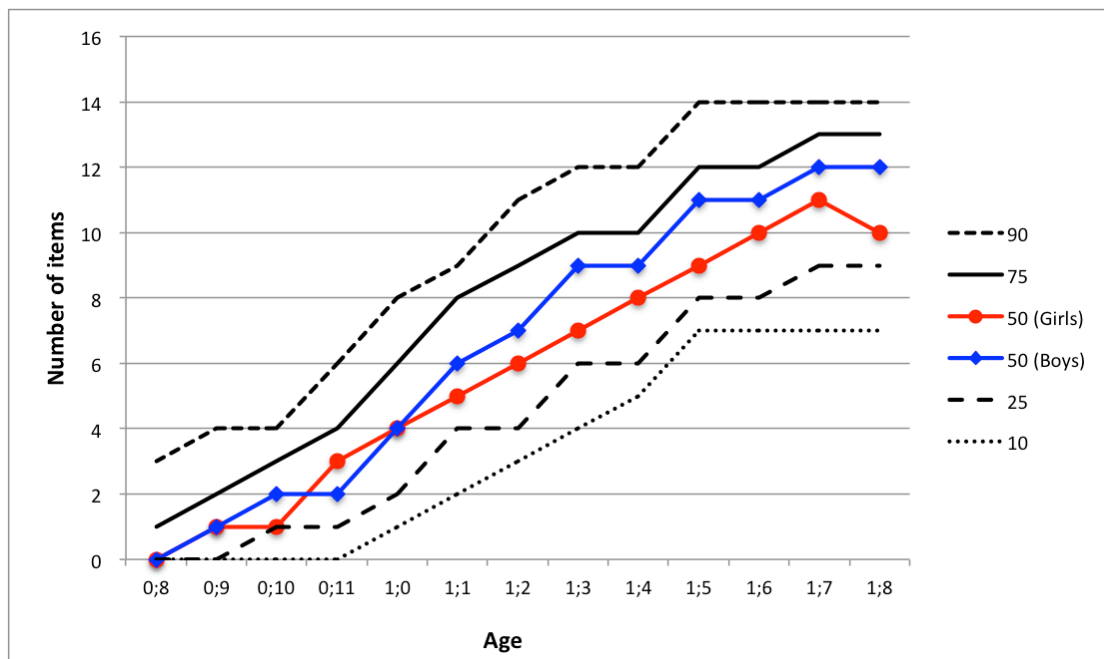
Figure 5. Percentiles of 'Imitation of other adult activities' score by age and gender (CDI:WG; 15 items)

Figure 5 presents selected percentiles of 'Imitation of other adult activities' score in infants by age and gender, comprising 15 items. In this subcategory, the interaction term did not reach significance, $F(12,2199) = 1.6, p=0.0795$ – still, the number of gestures increased significantly with age, $F(12,2199) = 367.2, p<0.0001,$ and varied significantly between the two genders, $F(1,2199) = 57.1, p<0.0001$. Note that here the gender difference was in the opposite direction from what we have seen in the other measures: here, the boys had a higher score than the girls. On average, a child used 1 gesture (0.92-0.97) from this subcategory more for each month, and girls used 1 gesture (0.6-1.0) less than boys.

Developmental trends: Toddlers
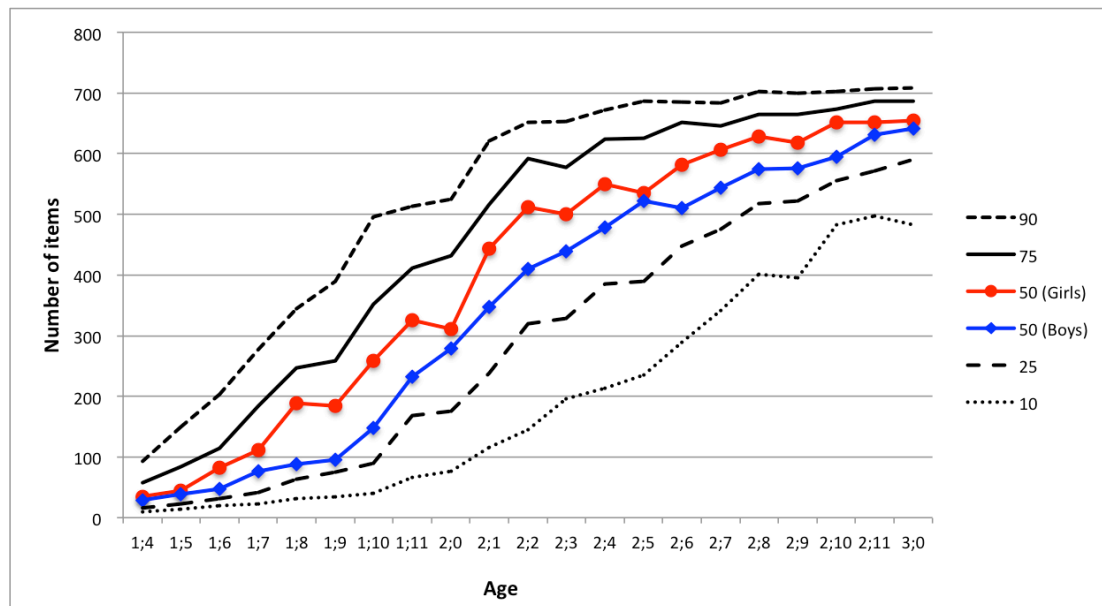
Figure 6. Percentiles of productive vocabulary size by age and gender (CDI:WS; 731 items)

Figure 6 presents selected percentiles of productive vocabulary size in toddlers by age and gender, comprising 731 items. Not surprisingly, vocabulary production continued to grow with age, $F(20,4173) = 382.9, p<0.0001,$ although there was a potential ceiling effect towards the end of the observation period where children used more than 600 words of the 731 words included in the checklist. A median girl started out with 35 words at age 1;4, reaching 654 words at 3 years of age. A median boy was a bit slower, beginning with 29 words at age 1;4 but using roughly the same number of words at the end of the observation period. However, there was a main gender effect, $F(1,4173) = 152.8, p<0.0001,$ which was clearly visible in the middle of the observation period. In the beginning, up to age 1;7, the difference between the girls' and boys' median was rather small. In the following year, up to age 2;7, the difference was constantly large (ranging from 61 words at age 2;1 to 132 at age 2;0, though with a minimum of 13 at age 2;5). Towards the end, the boys seemed to catch up with the

girls (whose results show a ceiling effect). The age by gender interaction was significant, $F(20,4173) = 2.6, p=0.0002$. On average, a child said 32 words (31-32) more for each month, and girls said 56 words (47-65) more than boys.
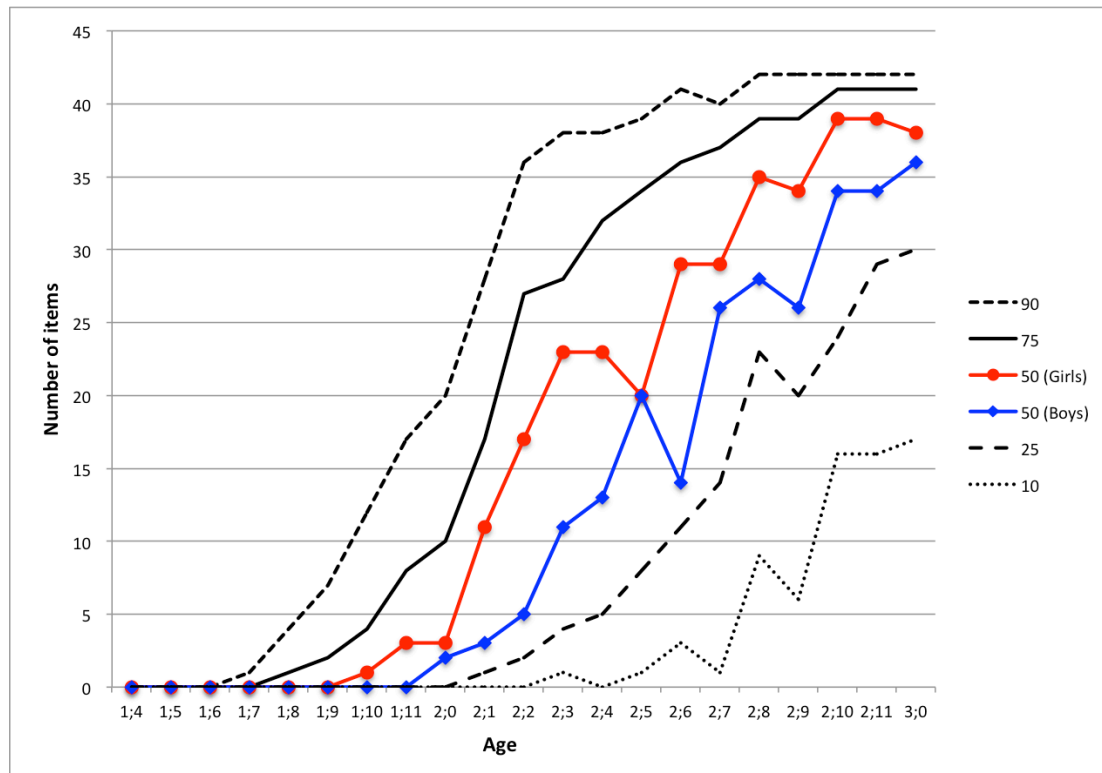
Figure 7. Percentiles of complexity score by age and gender (CDI:WS; 42 items)

Figure 7 presents selected percentiles of the complexity score in toddlers by age and gender, comprising 42 items. The same tendencies as for vocabulary production can be observed for the complexity score. There was a significant increase with age, $F(20,4173) = 285.7, p<0.0001$, gender plays a role as well, $F(1,4173) = 104.2, p<0.0001$, and the age by gender interaction was significant, $F(20,4173) = 3.6, p<0.0001$. On average, a child used 2 (1.9-2.0) more complex items for each month, and girls used 3 (2.8-4.1) more complex items than boys. There was a noticeable drop for both the girls' and boys' median at around 2;5 years of age – whether this is

simply due to random variation or can be connected to an intermediate step in the

grammatical development raises questions for further research.

Emergence of grammatical markers.

Table 4 summarizes our findings on the emergence of grammatical markers based on

the *CDI: Words and Sentences*. For word combinations and for each of the

grammatical markers of possession, definiteness, plural, present tense, and past tense,

there are three measures: the age when 25%, 50%, and 75% have started to use the

marker, that is, where both answer categories 'sometimes' and 'often' are interpreted

as 'yes' (= the child has started using this). Missing values were excluded. In addition,

a logistic regression was performed for each marker. Age (continuous) and gender

were included as covariates. Significance level was 5%. For each grammatical

marker, the age-adjusted odds ratio for gender was highly significant, *p <.0001*,

ranging from 1.5 (95% confidence interval 1.3 - 1.7) for "plural" to 2.6 (2.2 - 3.2) for

"possession".

TABLE 4. EMERGENCE OF GRAMMATICAL MARKERS

Marking	Age when **25%** have started[1]	Age when **50%** have started	Age when **75%** have started	Odds Ratio (95% CI)[2]
	All (Girls, Boys)	All (Girls, Boys)	All (Girls, Boys)	
Combining words	*Ex: mamma bil, mer kake*	1;6 (1;6, 1;6)	1;7 (1;7, 1;7)	1;10 (1;8, 1;11) 1.9 (1.6-2.3)
Possession	*-s/sin, sitt*	1;6 (1;6, 1;7)	1;9 (1;8,1;10)	2;0 (1;10, 2;0) 2.6 (2.2-3.2)
Definiteness	*-en/-a/-et*	1;8 (1;8,1;10)	1;11 (1;10,1;11)	2;2 (2;1, 2;3) 1.9 (1.6-2.2)
Plural	*-er*	1;10 (1;9,1;10)	2;1 (1;11,2;0)	2;4 (2;2, 2;5) 1.5 (1.3-1.7)
Present tense	*-er*	1;10 (1;9,1;10)	2;1 (2;1, 2;2)	2;4 (2;2, 2;5) 1.8 (1.5-2.1)
Past tense	*-a,-et, -te,-de*	1;10 (1;9,1;10)	2;0 (2;0, 2;0)	2;5 (2;3, 2;7) 2.0 (1.7-2.3)

[1] Earliest month where at least 25% of the answers were "sometimes" or "often".

[2] Based on logistic regression models with gender and (continuous) age as covariates.

Discussion

In this paper we have presented findings from a large population-based study of early

communicative development in children learning Norwegian, building on CDI data

from 6,574 children aged between 0;8 and 3;0 years, collected via the Internet. We

started with reliability analyses, and then reported developmental trends for

communicative gestures, comprehension and production of vocabulary, grammatical

complexity and grammatical skills, and also for gender variation on these measures.

We found high internal consistency and satisfactory test-retest correlations.

Undoubtedly, our reliability sample is rather small and self-elected, implying

potentially biased results. Non-biasedness cannot be proven in a post-hoc sample. The

only way to ensure non-biasedness would be a (large) randomized study. However,

since nobody can be forced to participate in a study, all (reliability) samples are essentially self-elected. The Norwegian findings are in good agreement with both American and Danish results (Fenson et al., 1994, 2007; Bleses et al., 2008a; Wehberg, 2008), suggesting that the Norwegian CDI data are not essentially different. Whether reliability (i.e. that a measurement is stable when repeated) is a meaningful concept at all when applied to a questionnaire which is designed to capture (rapid) development over time, is, of course, a matter of opinion.

As could be expected, in development we found that vocabulary comprehension preceded vocabulary production, and that both use of gestures, comprehension and production of vocabulary, grammatical complexity and grammatical skills increased with age, gestures producing the least consistently reliable results overall. In general, the findings for Norwegian seem to be in accordance with findings from other languages based on the same instrument, both concerning the extensive variation between children, and concerning general developmental trends (Bates & Goodman, 1999; Bleses et al., 2008b; Dale & Goodman, 2005; Law & Roy, 2008). Based on visual inspection of Figures 2, 6, and 7, it seems safe to conclude that Norwegian children go through two developmental spurts, one in productive vocabulary during the second year of life, and one in morphosyntactic development during the third year. The acceleration in vocabulary growth starts on average at around 50 words, but with extensive variation both concerning the number of words at the start of the 'spurt' (the earliest children start at 30-40 words) and as to when it starts (a span of 8 months, from 1;2 to 1;10). Whether all children go through such a spurt is of course a different question that cannot be answered here. Concerning grammatical skills, they seem to take off on average

around 1;7 with combination of words, followed by inflections of nouns and verbs over the next 5 months. Again, there is extensive variation between children as to when the development starts and how long it takes for the grammatical skills to be mastered. An evident and interesting next step will be to look more directly at the correlation between vocabulary size and the emergence of grammar in these children, as suggested by Bates and Goodman (1997, 1999).

For all measures included we found statistically significant gender differences, mainly in the direction of girls outperforming boys. Our results indicate that gender differences come out even more consistently in our data than in data from other languages, as reviewed in Eriksson et al. (2012). For receptive vocabulary, the female advantage was small. For productive vocabulary, it was more pronounced, starting already at the infant stage and increasing with age. Towards 3 years the boys seemed to catch up – however, since there is a ceiling effect on the present measure, we do not know whether this is actually the case (i.e., girls could be producing more words than are captured in this instrument). In communicative gestures, in grammatical complexity, as well as in acquisition of morphology the girls also clearly outperformed the boys – in morphology, the acquisition of the possession marker was particularly salient in this respect. The clearest gender difference was found in the category 'Pretending to be a parent', where the girls by far outperformed the boys, starting already at around the age of 1;0 and clearly increasing with age.

Interestingly, the same gender difference did not show up in the test items measuring 'Imitation of other adult activities', which includes a spectrum of activities from using a hammer and a saw to vacuum cleaning, watering flowers, driving a car and reading a book. Here, on the contrary, we found a small but significant advantage

for the boys. Although these results fit relatively well with more anecdotal observations, we have not seen them reported on in other CDI studies, and it is interesting to see how early these differences start to show. We do not know whether they are related to the bias in parents' educational background in our study, nor have we looked at how such findings are correlated with the gender differences in early vocabulary development. These are interesting questions for further study. Overall, the consistent gender differences we obtained indicate that separate norms for girls and boys should be considered when the instrument is used for screening for language delay.

As noted earlier, knowledge of early communicative development in Norwegian children is sparse and mostly based on single or multiple case studies. The present investigation is the first population-based study of children acquiring this language. We are now in a position to state with more confidence both rate of development and extent of variation in typically developing monolingual children acquiring Norwegian, at least within the areas covered by the CDI. We can also supplement the few earlier studies that exist. Recall for instance that Anderssen (2005, 2007) examined acquisition of prenominal definiteness markers, suffixal definiteness markers and pronouns in three children learning a Northern Norwegian dialect aged 1;8 - 3;3. She found that the definiteness suffix was used already at age 1;8, and that it was used in more than 80% of the obligatory contexts by two years of age. As evidenced in Table 4, our findings support those of Anderssen quite closely.

CDIs are currently being used as assessment tools in many languages, and the Norwegian CDI is no exception. However, in particular for this function the limitations of the study must be noted. Although the study is large and population-

based, the response rate was only 37%, and there is a bias with respect to educational

level of the parents. These are factors that need to be taken into consideration.

One of the advantages of the MacArthur-Bates CDI-instrument is that is has

been adapted to so many languages, thus forming a good basis for cross-linguistic

comparisons. With the Norwegian adaptation of the CDI-instrument and the norming

study reported from here, there exist adaptations to all the Scandinavian languages,

i.e. Danish, Swedish and Norwegian, as well as comparable data sets. These data sets

form a good starting point for more in-depth investigations of similarities and

differences in the acquisition of lexical and grammatical skills in these three

languages, and also for research on underlying factors – both linguistic and extra-

linguistic – contributing to the differences.

References

Anderssen, M. (2005). *The acquisition of compositional definiteness in Norwegian.*
Unpublished Doctoral dissertation, University of Tromsø, Tromsø.

Anderssen, M. (2007). The acquisition of compositional definiteness in Norwegian.
Nordlyd, 34(3), 252-275.

Aukrust, V. G. (1992). *Fortellinger fra stellerommet: to-åringer i barnehage: en
studie av språkbruk - innhold og struktur.* Unpublished Doctoral dissertation,
University of Oslo, Oslo.

Bates, E., Dale, P.S., & Thal, D (1995). Individual differences and their implications
for theories of language development. In P.Fletcher & B. MacWhinney (Eds.),
Handbook of child language (pp. 96-151). Oxford: Basil Blackwell.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes, 12* (5/6), 507-584.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29-79). Mahwah, NJ: Lawrence Erlbaum.

Berglund, E., Eriksson, M., & Johansson, I. (2001). Parental reports of spoken language skills in children with Down syndrome. *Journal of Speech, Language, and Hearing Research, 44*(1), 179.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008a). The Danish Communicative Developmental Inventories: validity and main developmental trends. *Journal of Child Language, 35*(3), 651-669.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008b). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language, 35*, 619-650.

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*(2), 159-200.

Caselli, M. 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*(1), 69-111.

Caselli, M. C., Monaco, L., Trasciani, M., &Vicari, S. (2008). Language in Italian children with Down syndrome and with specific language impairment. *Neuropsychology, 22*(1), 27-35.

Caselli, M. C., Vicari, S., Longobardi, E., Lami, L., Pizzoli, C., & Stella, G. (1998).

Gestures and words in early development of children with Down syndrome.

Journal of Speech, Language, and Hearing Research, 41(5), 1125-1135.

Chilosi, A. M., Cipriani, P., Bertuccelli, B., Pfanner, L., & Cioni, G. (2001). Early

cognitive and communication development in children with focal brain

lesions. *Journal of Child Neurology, 16*(5), 309-316.

Dale, P. S., & Goodman, J. C. (2005). Commonality and individual differences in

vocabulary growth. *Beyond nature-nurture: Essays in honor of Elizabeth

Bates* (pp. 41-78). Mahwah, NJ: Lawrence Erlbaum.

Dale, P. S., & Penfold, M. (2011). Adaptations of the MacArthur-Bates CDI into non-

U.S. English languages. *2011*(1114). Retrieved from

http://www.sci.sdsu.edu/cdi/documents/AdaptationsSurvey7-5-11Web.pdf

Devescovi, A., Caselli, M. C., Marchione, D., Pasqualetti, P., Reilly, J., & Bates, E.

(2005). A crosslinguistic study of the relationship between grammar and

lexical development. *Journal of Child Language, 32*(4), 759-786.

Eriksson, M., Marschik, P.B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg,

S.,… & Gallego, C. (2012). Differences between girls and boys in emerging

language skills: Evidence from 10 language communities. *British Journal of

Developmental Psychology, 30*, 326-343.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J.S., & Thal, D. (2000).

Measuring variability in early child language: Don't shoot the messenger.

Child Development, 71, 323-328.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., . . . Stiles, J. (1994). *Variability in early communicative development* (Vol. 59): Society for research in child development.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: Users guide and technical manual*. Paul H. Brookes Publishing Co.

Hall, S., Rumney, L., Holler, J., & Kidd, E. (2013). Associations among play, gesture and early spoken language acquisition. *First Language, 33*, 294-312.

Ingvaldsen, S. (2001). *Barns tidlige språkutvikling: en undersøkelse med vekt på relasjonen mellom forskjellige spesifikke mål.* Unpublished MA thesis, University of Oslo, Oslo.

Janson, H. (2003). Influences on participation rate in a natural Norwegian child development screening questionnaire study. *Acta Pædiatrics, 92*, 91-96.

Janson, H. & Smith, L. (2003). *Norsk manualsupplement til Ages and Stages Questionnaires*. Oslo: R.BUP, Regionsenter for barne- og ungdomspsykiatri, Helseregion Øst/Sør.

Janson, H. & Squires, J. (2004). Parent-completed developmental screening in a Norwegian population sample: a comparison with US normative data. *Acta Pædiatrics, 93*, 1525-1928.

Jørgensen,, R. N., Dale, P. S., Bleses, D., & Fenson, L. (2009). CLEX: A cross-linguistic lexical norms database. *Journal of Child Language, 36*, 1-10

Kristoffersen, K. E., Simonsen, H. G., Bleses, D., Wehberg, S., Jørgensen, R. N., Eiesland, E. A., & Henriksen, L. Y. (2013). The use of the Internet in

collecting CDI data - an example from Norway. *Journal of Child Language, 40*, 567-585.

Kristoffersen, K. E., Smith, L., Wium, K., Torkildsen, J. v. K., Simonsen, H. G., & von Tetzchner, S. (2008). MacArthur-Bates communicative development inventories (Norwegian adaptation): University of Oslo.

Law, J. & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health, 13*(4), 198-206.

Lorenzen, V. & Pedersen, H. (2001). *Barns tidlige språk- og kommunikasjonsutvikling i en sosialkognitiv ramme.* Unpublished MA thesis, University of Oslo, Oslo.

Maital, S. L., Dromi, E., Sagi, A., & Bornstein, M. H. (2000). The Hebrew Communicative Development Inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language, 27*(1), 43-67.

McBride-Chang, C., Tardif, T., Cho, J.-R., Shu, H., Fletcher, P., Stokes, S. F., . . . Leung, K. (2008). What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied Psycholinguistics, 29*(3), 437-462.

O'Toole, C. & Fletcher, P. (2010). Validity of a parent report instrument for Irish speaking toddlers. *First Language, 30*, 199-217.

Plunkett, K. & Strömquist, S. (1992). The acquisition of Scandinavian languages. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 3, pp. 457-556). Hillsdale, NJ: Lawrence Erlbaum.

Scarborough, H. S. (2009). Connecting early language and literacy to later reading (dis) abilities: Evidence, theory, and practice. In F. Fletcher-Campbell, J. Soler

& G. Reid (Eds.), *Approaching Difficulties in Literacy Development: Assessment, Pedagogy and Programmes* (pp. 23-38). Los Angeles: Sage.

Schjølberg, S., Lekhal, R., Wang, M. V., Zambrana, I. M., Mathiesen, K. S., Magnus, P., & Roth, C. (2009). *Forsinket språkutvikling. En foreløpig oversikt basert på data fra Den norske mor og barn undersøkelsen*. Oslo: Norwegian Institute of Public Health.

Simonsen, H. G. (1986). Om toåringers fonologi – system og prosesser. In P. E. Mjaavatn & L. Smith (Eds.), *Barnespråk* (pp. 210-233). Trondheim: NAVFs senter for barneforskning.

Simonsen, H. G. (1990). *Barns fonologi: system og variasjon hos tre norske og et samoisk barn.* Unpublished Doctoral dissertation, University of Oslo, Oslo.

Stolt, S., Haataja, L., Lapinleimu, H., & Lehtonen, L. (2008). Early lexical development of Finnish children: A longitudinal study. *First Language, 28*, 259-279.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development, 70*(3), 620-635.

Thal, D. J., Reilly, J., Seibert, L., Jeffries, R., & Fenson, J. (2004). Language development in children at risk for language impairment: Cross-population comparisons. *Brain and Language, 88*(2), 167-179.

Tomasello, M. (2003). *Constructing a language. A usage based theory of language acquisition.* Cambridge, Ma. & London: Harvard University Press.

Torkildsen, J. K. (2008). *Lexical processing in typically and atypically developing toddlers: insight from event-related brain potentials* (Vol. PhD). Oslo: UniPub.

Torkildsen, J. K., Friis Hansen, H., Svangstu, J. M., Smith, L., Simonsen, H. G.,
    Moen, I., & Lindgren, M. (2009). Brain dynamics of word familiarization in
    20-month-olds: Effects of productive vocabulary size. *Brain and Language,
    108*(2), 73-88.

Torkildsen, J. K., Sannerud, T., Syversen, G., Thormodsen, R., Simonsen, H. G.,
    Moen, I., . . . Lindgren, M. (2006). Semantic organization of basic-level words
    in 20-month-olds: An ERP study. *Journal of Neurolinguistics, 19*(6), 431-454.

Torkildsen, J. K., Svangstu, J. M., Hansen, H. F., Smith, L., Simonsen, H. G., Moen,
    I., & Lindgren, M. (2008). Productive vocabulary size predicts event-related
    potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive
    Neuroscience, 20*(7), 1266-1282.

Torkildsen, J. K., Syversen, G., Simonsen, H. G., Moen, I., & Lindgren, M. (2007).
    Electrophysiological correlates of auditory semantic priming in 24-month-
    olds. *Journal of Neurolinguistics, 20*(4), 332-351.

Trudeau, N. & Sutton, A. (2011). Expressive vocabulary and early grammar of 16- to
    30-month-old children acquiring Quebec French. *First Language, 31*, 480-
    507.

Vanvik, A. (1971). The phonetic-phonemic development of a Norwegian child. *Norsk
    tidsskrift for sprogvidenskap, 24*, 269-325.

Wehberg, S. (2008). *Numbers on words – analyses of the Danish Longitudinal CDI
    study.* Unpublished PhD thesis, University of Southern Denmark, Odense.

Westergaard, M. R. (2005). *The development of word order in Norwegian child
    language: the interaction of input and economy principles in the acquisition of
    V2.* Unpublished PhD thesis, University of Tromsø, Tromsø.

Appendix

Overview of the participants at each monthly stage, divided by gender

Participants - *Words and Gestures*

	Boys	Girls	Total
Months	N	N	N
8	56	68	124
9	83	76	159
10	92	110	202
11	81	84	165
12	97	88	185
13	92	102	194
14	86	100	186
15	81	110	191
16	84	89	173
17	88	88	176
18	105	90	195
19	103	92	195
20	120	94	214

Participants - *Words and Sentences*

	Boys	Girls	Total
Months	N	N	N
16	66	69	135
17	77	92	169
18	80	102	182
19	110	95	205
20	77	112	189
21	130	140	270
22	136	124	260
23	113	98	211
24	105	90	195
25	117	110	227
26	84	123	207
27	90	98	188
28	91	96	187
29	105	91	196
30	104	112	216
31	110	100	210
32	91	102	193
33	106	105	211
34	90	93	183
35	81	130	211
36	95	75	170