Author(s): Määttä, Sira; Laakso, Marja-Leena; Ahonen, Timo; Tolvanen, Asko; Westerholm, Jari; Aro, Tuija

Title: Continuity From Prelinguistic Communication to Later Language Ability : A Follow-Up Study From Infancy to Early School Age

Year: 2016

Version:

Please cite the original version:

Määttä, S., Laakso, M.-L., Ahonen, T., Tolvanen, A., Westerholm, J., & Aro, T. (2016). Continuity From Prelinguistic Communication to Later Language Ability : A Follow-Up Study From Infancy to Early School Age. Journal of Speech, Language, and Hearing Research, 59(December), 1357-1372. https://doi.org/10.1044/2016_JSLHR-L-15-0209

Continuity from prelinguistic communication to later language ability: a follow-up study from infancy to early school age

Journal:	*Journal of Speech, Language, and Hearing Research*
Manuscript ID	JSLHR-L-15-0209.R2
Manuscript Type:	Research Article
Date Submitted by the Author:	08-Apr-2016
Complete List of Authors:	Määttä, Sira; Niilo Mäki Institute, Research; University of Jyväskylä, Department of Psychology Laakso, Marja-Leena; University of Jyväskylä, Department of Education Ahonen, Timo; University of Jyväskylä, Department of Psychology Tolvanen, Asko; University of Jyväskylä, Department of Psychology Westerholm, Jari; Niilo Mäki Institute, Statistics; Aro, Tuija; University of Jyväskylä, Department of Psychology
Keywords:	Communication, Language, Development

Continuity from prelinguistic communication to later language ability: a follow-up study from infancy to early school age

Sira Määttä

Niilo Mäki Institute & University of Jyväskylä

Marja-Leena Laakso, Timo Ahonen Asko Tolvanen,

University of Jyväskylä

Jari Westerholm,

Niilo Mäki Institute

Tuija Aro

University of Jyväskylä & Niilo Mäki Institute

Author Note

Sira Määttä, Niilo Mäki Institute & Dept. of Psychology, University of Jyväskylä; Marja-Leena Laakso, Dept. of Education, University of Jyväskylä; Timo Ahonen, Dept. of Psychology, University of Jyväskylä; Asko Tolvanen, Dept. of Psychology, University of Jyväskylä; Jari Westerholm, Niilo Mäki Institute, Jyväskylä; Tuija Aro, Dept. of Psychology, University of Jyväskylä & Niilo Mäki Institute.

This research was supported in part by a grant from the Finnish National Doctoral Programme of Psychology.

Correspondence concerning this article should be addressed to Sira Määttä, Niilo Mäki Institute, PL 35, 40014 University of Jyväskylä, Finland. E-mail: sira.maatta@nmi.fi

Abstract

Purpose: This longitudinal study examined the development of prelinguistic skills, and continuity of communication and language from prelinguistic stage to school age.

Method: Prelinguistic communication of 427 Finnish children was followed repeatedly from 6 to18 months of age (*n*=203-322 at ages 6, 9, 12, 15, and 18 months), and its associations with language ability at ages 2;0 (*n*=104), 3;0 (*n*=112), 4;7 (*n*=253), 5;3 (*n*=102) and 7;9 (*n*=236) were examined using latent growth curve modeling.

Results: Prelinguistic development across several skills emerged as a rather stable intra-individual characteristic during the first two years of life. Continuity from prelinguistic development to later language ability was indicated. The common level and growth of prelinguistic skills were significant predictors of language ability between ages 2;0-7;9, the percentage explained varying between 10.5-53.3%. A slow pace of development across multiple skills, in particular, led to weaker language skills.

Conclusions: The results support the idea of a developmental continuum from prelinguistic to linguistic ability, and the dimensional view of language ability, by indicating that individual variations in early communication skills show consistency that extends beyond the toddler years. Our results also advocate developmental surveillance of early communication by emphasizing the significance of growth in predicting language development.

Keywords: communication, language, development

Continuity from prelinguistic communication to later language ability: a follow-up study from infancy to early school age

The development of communication skills starts long before children are able to use language as their primary means of communication. The prelinguistic stage refers to the period during which children use mainly nonverbal means of communication, and spans intentional preverbal communication and the transition to first words (Watt, Wetherby, & Shumway, 2006). The first two years of life is an important period in the development of these early communication skills. The way infants communicate prelinguistically is thought to form a developmental continuum with later, more language-based, communication (Bruner, 1983). Accordingly, Bates (2004) and Rescorla (2009, 2013) have suggested that several interrelated but distinct early socio-cognitive skills serve as building blocks for later language. That is, language emerges from the interactions of these early socio-cognitive processes (Bates, 2004), and differences in language ability stem, in part, from the differential endowment of these language-subserving skills (Rescorla, 2009).

These early socio-cognitive skills (i.e., prelinguistic means of communication) include joint attention, gestures, early vocalizations, first words, language comprehension, and play (Watt et al., 2006). Although these skills and their connections to later language outcome have been rather extensively studied (for a review, see McCathren, Warren, & Yoder, 1996), studies examining several of these skills together, their co-development over time and the implications that this co-development has on later language, are lacking. Thus, it is not possible to ascertain whether it is the general level or pace of early communication development (i.e., the variation that is shared across skills) that is predictive of later language ability, or whether a specific skill, or a combination of skills, at a specific age makes a unique contribution to language development over and above that of the general level. The present study utilizes a longitudinal design to address how prelinguistic communication skills, assessed repeatedly with a multifaceted parental screener covering relevant areas of social, speech, and symbolic skills, develop during the first two years of life. The connections between this development and that of later language and communication are followed up to school age.

The development of prelinguistic skills

Language and communication development is characterized by substantial inter- and intra-individual variation in the acquisition of different skills. However, despite this variation, both continuity (i.e., group-mean-level consistency) and stability (i.e., consistency in the relative standing of individuals over time) in language development have been reported (Bornstein, Hahn, Putnick, Suwalsky, 2013; Bornstein & Putnick, 2012; Fenson et al., 1994; Thal, Bates, Goodman, & Jahn-Samilo, 1997). The dimensional view of language ability (Rescorla, 2009, 2013) argues for the stability of individual differences in language skills by suggesting that the rank order of children is partly determined by differential endowment. According to Rescorla (2009, 2013) this endowment, that is, a spectrum of language ability which she compares to that of intelligence, derives from variation in several language-subserving socio-cognitive skills and is, at least partly, constitutionally based. These skills, such as auditory perception and processing, verbal working memory, and joint reference are assumed to form the base from which prelinguistic communication and later, language ability develops (Bates, 2004; Rescorla, 2013).

Research on stability in the development of language and communication has tended to focus more on older ages. For example, according to Bornstein and colleagues (2012, 2013), when multiple domains, measures, and sources are used across age, child language emerges as a stable characteristic of individual differences. The development of communication in the prelinguistic period has been less extensively studied. The developmental sequences of separate prelinguistic skills such as gestures (Bates & Dick 2002) and joint attention (Carpenter, Nagell, & Tomasello, 1998) have been studied, but several prelinguistic skills have been included in the same analysis in only a few studies, while even fewer have tracked the co-development of these skills in the prelinguistic period.

The few studies that have examined multiple prelinguistic skills simultaneously have reported significant correlations between measures of different skills (e.g., Laakso, Poikkeus, Katajamäki, &

Lyytinen, 1999; Watt, et al., 2006), indicating that these measures partially tap the same underlying functions (i.e. language endowment). For example, in Laakso et al. (1999) parental report on gestures and concurrently observed joint attention correlated significantly (*r* =.21-.26) at age 14 months. Watt et al. (2006) explored the concurrent correlations of several prelinguistic skills and reported that gestures (*r* =.29-.46) and joint attention (*r* =.29-.47) were significantly correlated with most of the other measures, especially early in the second year of life (a total of 22/36 correlations were significant, *r* =.01-.61 at 14 months, and 11/36, *r* =.00-.75 at 20 months).

Darrah, Hodge, Magill-Evans, and Kembhavi (2003), Reilly et al. (2006), Watt et al. (2006), and Wetherby, Allen, Cleary, Kublin, and Goldstein (2002) have examined the development of the social, speech, and symbolic skills of children using the Infant-Toddler Checklist (ITC) or the Behavioral Sample of the Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP, Wetherby & Prizant, 2002). These studies have all reported significant longitudinal correlations between assessments of joint attention, gestures, vocalizations, first words, and comprehension in the prelinguistic period (13-21 months *r* =.46-.55 for the total score in Darrah et al., 2003; 8-12 months *β* =.56 for the total score in Reilly et al, 2006; 14-20 months, *r* =.39-.59 in separate skills in Watt et al., 2006; 13-20 months, *r* =.77-89 for the different composites, and *r* = .85-.91 for the total score in Wetherby et al., 2002). These correlations indicate stability in individual differences in the development of these skills over time whereas concurrent reports on increases in raw scores indicate fast growth in these skills (Watt et al., 2006; Wetherby et al., 2002).

General level and pace of development or skill- and age-specific associations?

It has been shown that a major predictor of communication and language status at a given age is the level of skills at an earlier age, both in the prelinguistic stage (Reilly et al., 2006, 2007) and during later language development (Bornstein & Putnick, 2012). However, it has also been suggested that the pace of development, rather than the level at any given age, might be more predictive of later development (e.g. Rowe, Raudenbush, & Goldin-Meadow, 2012). This view has received support from studies examining vocabulary development (Rescorla, Mirak, & Singh, 2000), and early precursors of literacy (Lyytinen et al., 2006). According to Rowe et al. (2012), it is plausible that “the rates of growth likely contain more information about the child’s language acquisition potential than their ability at one point in time” (p. 510). This could be the case, especially with respect to early communication development, where growth depends also on the acquisition of new skills as well as augmenting existing skills (Reilly et al., 2006).

Another relevant question is whether it is the general (i.e., common) level or growth of prelinguistic development across several skills or a specific skill at a specific age that best predicts later language ability. Research has established links between several early socio-cognitive and prelinguistic communication skills and later language ability: for example, gaze following (Brooks & Meltzoff, 2008), and other forms of joint attention (Beuker, Rommelse, Donders, & Buitelaar, 2013), gestures (Colonnesi, Stams, Koster, & Noom, 2010), deferred imitation (Heimann, et al., 2006), verbal comprehension and symbolic play (Bruce, Kornfält, Radeborg, Hansson, & Nettelbladt, 2003), and frequency of intentional communication and reciprocity (Paavola, Kemppinen, Kumpulainen, Moilanen, & Ebeling, 2006). However, the comparison of studies is difficult, as their measures, ages and methodology vary considerably. In addition, multiple prelinguistic skills have been rarely addressed in the same study.

The results reported thus far suggest that the proportion of shared variance between the different prelinguistic measures is large relative to the unique contribution of single skills (Watt et al., 2006; Wetherby et al., 2002). For example in the studies by Wetherby and colleagues (2002; Watt et al., 2006), the joint contribution of gaze following, joint attention, gestures, vocalizations, first words, comprehension, and play was large relative to the unique contribution of any of these skills for the predicted variance. This led the authors to conclude that judgments about the relative importance of any particular skill in predicting language outcome should be guarded, and that using an array of prelinguistic measures strengthens their predictive value.

However, despite the large shared variance, the studies by Wetherby and colleagues on the different composites of the ITC have found specific associations between the social and symbolic composites and receptive language (*pr* =.23-.62), the speech composite and expressive language (*pr* = .17-.59, Wetherby et al., 2002), and the social composite and later autism spectrum disorder (Wetherby, Brosnan-Maddox, Peace, & Newton, 2008). Heimann et al. (2006) tracked the development of deferred imitation (early memory) and joint attention between ages 6-14 months and found that deferred imitation at 9 months was the single strongest predictor of gestures (β =.53) at age 14 months. In addition, Bruce et al. (2003) reported significant unique contributions of verbal comprehension (*r* =-.58) and symbolic play (*r* = -.40) at age 18 months to language difficulties at age 4;5. However, despite a rather comprehensive assessment of language at the follow-up, they reported the outcome results as frequency of difficulties, thus rendering generalization to typical development difficult. In Lyytinen, Poikkeus, Laakso, Eklund, and Lyytinen (2001), symbolic play at 14 months of age was found to correlate significantly (*r* =.28-.39) with later receptive language skills in typically developing children. However, when early comprehension was controlled for in regression models, symbolic play did not uniquely predict language outcome (β =.13). Similarly, Salley, Panneton and Colombo (2013) found that visual attention and joint attention made unique contributions to later vocabulary size (β = 0.278), but when baseline communication was controlled for, joined attention was no longer a significant predictor (β = 0.093). Thus, the results on the unique predictive ability of different prelinguistic skills remain inconclusive.

The predictive relations of different prelinguistic skills with language development have also been found to show age-specificity. Watt et al. (2006) studied the different variables of the CSBS DP and found that the predictive relations varied according to age: Early in the second year of life, specific predictive associations were found between gestures and receptive language, and joint attention and expressive language. Late in the second year, inventory of consonants contributed uniquely to expressive language. Comprehension was predictive of later receptive and expressive language throughout the second year. Others (Brooks & Meltzoff, 2008; Heimann et al., 2006; Rose, Feldman, Jankowski, & Van Rossem, 2008) have also suggested that the age at which early skills are assessed might affect the way they relate to later abilities, that is, the predictive power of a certain skill might vary depending on whether the skill is just emerging or already more established.

Goals of the present study

This study addresses the question of continuity and stability both within the prelinguistic period and from the prelinguistic period to linguistic development. We examined the development of early communication skills by repeatedly following, during the first two years of life, several of the relevant developmental areas suggested by previous research. In addition, we explored the longitudinal associations of this early development with later language ability. Following the premises of the dimensional view of language (Rescorla, 2013) and the suggestions of Bornstein and Putnick (2012) and Conti-Ramsden and Durkin (2012), a multiage, multidomain, multimeasure, and multisource approach was adopted in the follow-up procedures of the present study. The complex nature of language requires that both multiple dimensions of language and measures of working memory be used in the assessment (Conti-Ramsden & Durkin, 2012). The associations between prelinguistic development and subsequent language outcomes were studied in separate but largely overlapping subsamples at five consecutive time points (at age 2, 3, 4;7, 5;3 and first grade, mean age 7;9, range 7;2 – 8;4). Three areas of language development (expressive, receptive, communicative/pragmatic), along with verbal working memory, were explored using several different measures, including both parental report and psychometric testing.

Specifically, we asked:

1. How is the development of prelinguistic communication skills depicted when three relevant areas of development (social, speech, symbolic) are assessed longitudinally between 6 and 18 months of age? In particular, we explored the stability of individual differences over time, and whether development in these three areas is mainly overlapping (i.e., can be depicted by a model of common level and growth)

or distinct (i.e., skill-specificity), and whether the course of development shows age-related differences (i.e., age-specificity). Based on previous findings on early communication skills (Laakso et al., 1999) and the ITC (Watt et al., 2006; Wetherby et al., 2002), we expected the three studied areas to show significant shared variance and also, to some extent, skill- and age-specific variance. In addition, we expected to find stability in development over time.

2. Which aspects of early development (i.e., common level and growth or skill- or age-specific features) best predict later language ability? In accordance with the dimensional view (Rescorla, 2009, 2013), and earlier findings (e.g., Bornstein et al., 2013), we expected to find indications of continuity and stability on the aggregate level (i.e., common level and growth predicting later abilities). Based on the findings of Wetherby and colleagues (2002; Watt et al., 2006) in somewhat older samples, we hypothesized that social and symbolic skills would show more predictive power early on in development, whereas the role of speech skills might be more pronounced later on.

Method

Brief summaries of the methods follow; for a more detailed description of the participants, procedures, and measures, see the online Supporting Material.

Participants and procedure

The participants of the present study represent subsets of a community-based sample collected in a longitudinal study of early language and communication development conducted between the ages of 6 months and 8 years (see Määttä, Laakso, Tolvanen, Ahonen & Aro, 2012, 2014). Altogether, 508 children (50.2 % boys, 49.8 % girls) aged 6 to 24 months participated in the study. All of the families were Caucasian, and all of the children spoke Finnish as their native language. At the initial assessment, mothers' mean age was 29.8 years (*SD* = 5.4), and fathers' 32.1 years (*SD* = 6.3). Educational attainment was assessed with a 7-point scale ranging from no vocational education (0) to a higher-level university degree (6). The mean educational level was 3.9 (*SD* = 2.0) for mothers and 3.6 (*SD* = 2.0) for fathers.

The early questionnaire (ITC) data were collected repeatedly every three months until the children were 24 months of age. The total number of questionnaires filled in by the parents depended on the age of their child at recruitment and on how many of the subsequent forms they completed. In the present study we used the data gathered on the children across ages 6 – 18 months. This yielded a total sample of 427 children (*n* = 229 at 6 months, *n* = 203 at 9 months, *n* = 322 at 12 months, *n* = 305 at 15 months, and *n* = 279 at 18 months of age). Of these 427 children 25.8 % had data from all five data points, 9.6 % from four, 29.0 % from three, 23.4 % from two and 12.2 % from one data point. The last two measurements of the early data, collected at 21 and 24 months, were excluded from the analyses due to skewed and kurtic distributions (ceiling effect).

Subgroups of the original participants were followed after the ITC data collection phase at the age of 2, 3, 4;7, and 5;3 years, and in the spring term of first grade (mean age 7;9, range 7;2-8;4). The numbers of participants are described in Figure 1, subsample differences in the ITC scores in Table SM1, and the demographic data of the different subsamples along with information on Finnish families and family services in the online Supporting Material and Table SM2. In the follow-ups at ages 2, 3, and 5;3, we were not able to collect information from all the families, owing to time and resource limitations. Thus, the subsamples were constructed so as to ensure that a sufficient number of at-risk children would participate. At-risk status was defined as slow communication development in ITC at ages 12, 15 and 18 months following the criteria suggested by Wetherby & Prizant (2002; follow-ups at ages 2 and 3 years) or as a high score (90th percentile) in a parent report symptom questionnaire at age 4;7 (follow-up at age 5;3). In the follow-ups at age 4;7 and first grade, all the originally participating families, excluding those who had declined to participate in further follow-ups in the previous data collection phases, were attempted to contact. For the families that were not reached, we were unable to find a valid address.

Insert Figure 1 about here.

At age 2, a small subset of families (n = 143) was invited to fill in the MacArthur-Bates Communicative Inventories (MBCDI; Fenson et al., 1994; Lyytinen, 1999). The participants included 65 children who were identified as being at risk based on their scores in the ITC, using the norms and 10th percentile cut-off reported by the original authors (Wetherby & Prizant, 2002). These children performed either in the lowest 10 percent in the social or symbolic composites at 12 or 15 months of age or within the lowest 10 percent in the speech composite at 15 or 18 months of age. The rest of the sample (n = 78) performed above the 10th percentile in all three composites at all ages. In total, 104 families (72.7%, at risk n = 44, no risk n = 60) returned the questionnaire. The mean age of the children at the time of the completion of the questionnaire was 25.3 months (SD = 1.2, range 24-30 months). The ITC composite scores at 12, 15, and 18 months were compared between the children who had data at age 2 (n = 99, 101, and 98 respectively) and those who did not, (n = 223, 204, and 181). Effect sizes were calculated using partial eta squared (η_p^2). Significant differences between the children emerged in the speech composite at 15 months, $F(1, 302) = 17.557$, $\eta_p^2 = .055$, $p = .000$ after controlling for multiple comparisons (Table SM1).

At age 3, the same subset of families was invited for individual assessments of vocabulary. Of these families, 112 (at-risk n = 56, no risk n = 56) agreed to participate. The mean age of the children at the time of the assessment was 36.7 months (SD = 0.8, range 36-41). When comparing the children with data at age 3 (n = 109, 108, and 105 at 12, 15, and 18 months of age respectively) and those without (n = 213, 196, and 174), small but significant differences in the social composite at 15 months, $F(1, 303) = 12.282$, $\eta_p^2 = .039$, $p = .001$, and speech composite at 15 months, $F(1, 302) = 15.346$, $\eta_p^2 = .048$, $p = .000$ and 18 months, $F(1, 277) = 12.948$, $\eta_p^2 = .045$, $p = .001$ of age remained after controlling for multiple comparisons (Table SM1). In both 2 and 3 year data comparisons, the significant differences resulted from the participating children having lower mean and showing larger variation than the children without follow-up data. In the present study all the available data from the assessments at ages 2 and 3 years were used.

When the children were aged 4 years 7 months all the originally participating families were sent a questionnaire concerning their child's language and communication skills. Of the 508 families, 473 (93.1%) were reached, and 296 (62.6% of reached; 58.3% of the original sample; total attrition rate 41.7%) returned the questionnaire. The mean age of the children at the time of the completion of the questionnaire was 56.9 months (*SD* = 4.0, range 52-69). There were no significant differences in ITC scores at ages 12, 15, and 18 months between the children who participated in this follow-up and those who did not. In the present study, children who had early data only from ages 21 and 24 months were excluded, and thus data from 253 children were used.

At age 5 years 3 months (5;3), a subsample of 102 children were invited for individual follow-up assessment. Primarily, children with full datasets from the previous assessment points (early questionnaire data, vocabulary data from either age 2 or 3 or both, and questionnaire data from age 4;7, *n* = 70) were selected to ensure adequate data for studying development over time. The sample was supplemented with children whose parents reported concerns related to language and communication, hyperactivity, or executive functions in the previous follow-up stage at age 4;7 (*n* = 32). The mean age of the children at the time of the assessment was 62.3 months (*SD* = 0.5, range 61-65 months). No significant differences were observed in the ITC scores at ages 12, 15, and 18 months between the children who participated in this follow-up (*n* = 93, 97, and 88, respectively) and those who did not (*n* = 229, 208, and 191). In the present study, all the available data were used.

The final follow-up was conducted during the spring term of the first grade (mean age = 93.3, *SD* = 3.9, range 86-103 months). All the originally participating families were sent a questionnaire on their child's language and communication skills. Altogether, 453 families (89.2%) were reached and 273 (60.3 % of reached; 53.7 % of the original sample; total attrition rate 46.3%) returned the questionnaire. There were no significant differences in ITC scores at ages 12, 15, and 18 months between the children who participated in this follow-up and those who did not. In the present study, children with early data only from ages 21 and 24 months were excluded, resulting in a sample of 236 children.

Measures

Parents completed questionnaires every three months between the ages 6 to 18 months and at the follow-ups at ages 2 years, 4;7, and first grade. Face-to-face assessments were administered at the ages of 3 years, and 5;3 (see Table 1.).

Insert Table 1 about here.

Early development measure. The Finnish version of the Infant-Toddler Checklist (ITC) of the Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP, Laakso, Poikkeus, & Eklund, 2011; Wetherby & Prizant, 2002) was used to obtain parental estimates of their children's early communication skills. The ITC is one of the most comprehensive parent-report screening tools for prelinguistic and early language skills currently available (for a review of methods, see Crais, 2011). The questionnaire covers three composites of development that address several relevant aspects of prelinguistic communication, such as emotion and eye gaze, gestures, and communication (social), sounds and words (speech), and understanding and object use (symbolic). Wetherby and Prizant (2002) report Cronbach's alphas (*α*s) ranging from .87 to .99 for the three composites combined over the age span of 6 to 24 months, which indicates a high degree of internal consistency. In the present data, the Cronbach's *α*s over the age span of 6 to 18 months ranged from .80 to 89, and by age (6, 9, 12, 15, and 18 months; *n*s = 191-320) from .68 to .73 for the social composite, from .47 to .63 for the speech composite, and from .38 to .58 for the symbolic composite.

The ITC has been shown to be able to detect developmental growth and produce relatively stable rankings of children over short periods of time (Reilly et al., 2006; Wetherby et al., 2002), although indications of instability in ITC rankings both between and within individuals have also been reported

(Darrah et al., 2003). Longitudinal connections between ITC scores and receptive and expressive language at 2 and 3 years of age (Wetherby, Goldstein, Cleary, Allen, Kublin, 2003), and between the ITC and later communication difficulties, including autism spectrum disorders (Wetherby et al, 2008) have been reported. However, studies extending the follow-up period beyond the toddler years remain scarce (for exceptions, see Määttä et al., 2012, 2014; Reilly et al., 2006).

Follow-up measures

Measures at 2 years (n = 104). The vocabulary scale, sum of noun and verb inflections, and maximum sentence length subscales of the Finnish version of the MacArthur-Bates Communicative Development Inventories Words and Sentences (MBCDI; Fenson et al., 1994; Lyytinen, 1999) was used as a measure of early expressive vocabulary. Cronbach's α for the vocabulary scale was .95.

Measures at 3 years (n = 112). Children's receptive vocabulary was assessed with the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1981) and expressive vocabulary with Boston naming (Kaplan, Goodglass, & Weintraub, 1983). Cronbach's αs were .94 for PPVT and .82 for Boston naming.

Measures at 4;7 (n = 253). Children's language-related difficulties were assessed using the questionnaire Five to Fifteen (FTF; Kadesjö et al., 2004). As the FTF is a symptom questionnaire based on parent report the results are regarded to represent parental concerns as opposed to clinically evaluated difficulties. The language domain of the FTF includes three subscales that cover comprehension, expressive, and communication skills. The Cronbach's αs for the scales were .66 for comprehension, .87 for expressive and .71 for comprehension.

Measures at 5;3 (n = 98-102). The language tasks were selected to measure a range of language-based skills that tap different dimensions of language in both the receptive and expressive domains. The Similarities subtest (SI, WPPSI-R; Wechsler, 1995) was used to assess verbal abstract reasoning and conceptualization abilities. Single-word receptive vocabulary was assessed with a short version of Peabody Picture Vocabulary Test – Revised (PPVT; Dunn & Dunn, 1981). As a measure of receptive grammar, we used the Korpilahti Auditory Sentence Comprehension test (SC; Korpilahti, 1996), which assesses the ability to process semantic and syntactic information in sentences of increasing complexity. The Verbal Fluency subtest of NEPSY-II (VF; Korkman, Kirk, & Kemp, 2008) was used to assess verbal fluency and vocabulary through the ability to generate words within specific semantic categories.

The memory tasks were selected to measure different subsystems of Baddeley's (2003) model of working memory. The phonological loop (i.e., auditory short-term memory) was assessed with the Digit Span- forwards subtest (DSf, WISC-III; Wechsler, 1999) and the Repetition of Nonsense Words task (NWR, NEPSY; Korkman, Kirk, & Kemp, 1997). The Digit Span- backwards subtest (DSf, WISC-III; Wechsler, 1999) was used to assess the central executive and the Sentence Repetition task (SR, NEPSY-II; Korkman et al., 2008) to assess the episodic buffer.

Measures in first grade (mean age 7;9, range 7;2-8;4, n = 236). The Finnish version of the Children's Communication Checklist-II (CCC-2; Bishop, 2003; 2014) was used to assess children's language and communication difficulties. The questionnaire includes four subscales evaluating language-related abilities (speech, syntax, semantics, coherence; Cronbach's $\alpha = .91$) and four subscales concentrating on pragmatics (inappropriate initiations, stereotyped language, use of context, nonverbal communication; $\alpha = .92$).

Data analyses

The repeated measures of early communication skills (the three composites of the ITC: social, speech, and symbolic) were analyzed using a type of second-order multivariate Latent Growth Curve modeling (LGC, Bollen & Curran, 2006; factor-of-curves, Duncan, Duncan, & Strycker, 2006). The analyses were performed using the Mplus statistical package (version 7; Muthén & Muthén, 1998-2010). The missing data function in Mplus enables all the observations in the data to be used in estimating the parameters of the models. Because some of the variables were skewed, the robust MLR estimation method was used (Muthén & Muthén, 1998-2010). The goodness-of-fit of the estimated LGC models was evaluated using the χ^2 test (p >.05), the Comparative Fit Index (CFI ≥ .95), the Tucker-Lewin Index (TLI ≥ .95), the Root Mean Square Error of Approximation (RMSEA < .06), and

Standardized Root Mean Square Error of Approximation (SRMR < .08) (Hu & Bentler, 1999; Muthén & Muthén, 1998-2010). Instead of as definitive cut-off criteria, the values of the fit indices were used as guidelines for evaluating the model fit (for a critical discussion, see Marsh, Hau, & Wen, 2004). Greater weight was given to the other fit indices than to chi-square, as the chi-square value is known to be sensitive to large sample sizes (Miles & Shevlin, 2007). The model modification indices, alongside with theoretical considerations, were utilized in specifying the model.

In the analysis, the growth curves were first applied simultaneously to each ITC composite, estimating the initial level of each composite (i.e., the level), the average rate of growth (i.e., the slope), and individual variation in the initial level and growth. These first-order factors described individual differences within each ITC composite. Second-order common factors (common level and common growth) were then added to describe commonality (i.e., to model the correlation structure) among the first-order factors. The associations of early communication development (the early LGC model) with later language ability were explored by regressing the follow-up measures on the common level and growth factors. Skill- and age-specific connections were tested by building the specific pathways suggested by the model modification indices. The regressions were run separately for each follow-up stage. Raw scores were used in all the analyses.

Results

A latent growth curve (LGC) model for early communication development

The means and standard deviations of the three ITC composites (social, speech, symbolic) between ages 6 and 18 months are shown in the upper part of Table 2. All three composites showed marked growth throughout the assessment period and all the successive measurements within the composites correlated significantly with each other (social r =.47 - .72, p <.001; speech r =.34 - .79, p <.001, and symbolic r =.40 - .68, p <.001; for a full correlation matrix, see Table SM3) with a large effect size (Cohen, 1992). However, there were also notable differences between individuals throughout the period. That is, overlap in the scores was observed between the different age stages – the highest performing children at only 6 months of age scored almost as high as the lowest performing children at ages 15 and 18 months (95 % confidence intervals).

Insert Table 2 about here.

A LGC model for each of the three early communication composites was estimated simultaneously (see Figure 2). Altogether, 427 children were included in the analysis. The coverage of the elements in the covariance matrix varied from 31.6 to 75.4 % (see Table SM4). Due to the sequential nature of the data, all the successive measurements were allowed to correlate with each other within the composites. Following the suggestions of the modification indices and visual inspection of the individual growth curves, nonlinear growth was estimated: the first and last factor loadings on the growth factors of each communication composite were fixed, and the factor loadings at ages 9, 12, and 15 months were estimated freely. The level and growth factors of the three composites were allowed to correlate, and the correlations were high and significant both between the level factors (r =.57-.81, p <.001), and the growth factors (r = .48 - .67; p <.001 - .010).

A second-order factor structure was then added to the previous model in order to model the correlation structure between the first-order factors. The three first-order level factors were set to load onto the second-order level factor (common level) and the three first-order growth factors were set to load onto the second-order growth factor (common slope; see Figure 2). Because of high correlations between the residuals of the different composites at ages 9, 12 and 15 months, specific factors by age were added to explain the residual covariance. The loadings of the three composites were set equal across the three measurements. The model fitted the data well: $\chi^2(73) = 87.405$, p =.120 , CFI =.991 , TLI =.987, RMSEA = .021 and SRMR =.083.

Insert Figure 2 about here.

All loadings on the second-order level and growth factors were significant and positive, with small differences in the magnitude of the loadings between the different composites. The second-order level factor explained (R^2) 74 % of the variance in the first-order level factor of the social composite, 72 % of that of the speech composite, and all of the variance in the level of the symbolic composite (due to a small negative residual variance, the residue of the first-order level factor of the symbolic composite was set to zero). The second-order growth factor explained 88 % of the variance in the first-order growth factor of the social composite, and 91 % of the symbolic composite. For the speech composite, the percentage explained was somewhat smaller (66 %), although significant. The second-order level and growth factors correlated negatively (r = -.48), indicating that the rate of growth was steeper for children who started at a lower initial level. The residuals of the first-order level factors of the social and speech composites, and the first-order growth factor of the speech composite were significant (.26 - .34, p = .010-.031). This indicates that, despite good model fit, there was skill-specific variation that was not explained by the common level and growth factors. In addition, the presence of the age-specific factors at ages 9, 12 and 15 months suggest that there is also age-specific variation, not captured by the growth model.

In sum, the LGC model of early communication skills suggested that there is a large amount of shared variance in the development of early social, speech, and symbolic skills. That is, individuals tend to be on a similar level (i.e. to have similar ranking relative to others) across the different skills, and the rate of development tends also to be similar across the skills, especially in social and symbolic composites. However, despite the notable commonalities, there is also significant skill- and age-specific variation, as indicated by the significant residual variances of the skill-specific factors, and the emergence of the age-specific factors.

Early communication development and later language and communication skills

The longitudinal associations of the LGC model for early communication development with later language and communication development were explored separately for each follow-up measurement at ages 2, 3, 4;7, and 5;3 years, and in first grade. The analysis was performed in two steps. First, the follow-up measures were regressed on the second-order factors (i.e., common level and growth). Second, in order to explore possible skill- and age-specific pathways, the specific associations suggested by the model modification indices were tested. For a summary of the model fit indices see Table SM5 in the online Supporting Material. The regression coefficients together with the tested specific associations are summarized in Table 3.

Insert Table 3 about here.

LGC and vocabulary at 2 years (n = 104). The MBCDI vocabulary, inflections and maximum sentence length (MSL) were used as the outcome measures. These data were available for 24 % of the children in the early LGC model. The resulting model fitted the data well: $\chi^2(112) = 153.548$, p =.0056, CFI = .979, TLI = .971, RMSEA = .029, SRMR = .091. The level and growth of early communication skills explained 32.4 % (β_L =.32, β_G =.64) of the variance in the MBCDI vocabulary, and the growth of early communication skills alone explained 41.1 % (β_G=.69) and 46.7 % (β_G=.74) of the variances of the MBCDI inflections and MSL, respectively.

LGC and vocabulary at 3 years (n = 112). At three years of age, the Boston naming, and PPVT were administered to a subsample of the children (26 % of the children in the LGC model). The model fitted the data well: $\chi^2(99) = 119.435$, p = .0793, CFI = .988, TLI = .984, RMSEA = .022 ,and SRMR = .083. For Boston naming, the common growth in early communication skills explained 27.6 % (β_G=.58) of the variance. For the PPVT, both the common level and growth of early communication skills together explained 10.5 % (β_L =.29, β_G=.33) of the variance.

LGC and parental concerns of language development at age 4;7 (n = 253). Parents reported language related difficulties in the areas of expressive and receptive language and communication skills using the FTF when the children were aged 4;7. These data were available for 59 % of the children in the LGC model. The model fitted the data well: $\chi^2(112) = 136.459$, $p = .0579$, CFI = .987, TLI = .982, RMSEA = .022, and SRMR = .078. Together, the common level and growth of early communication skills explained 15.0 % (β_L=-.28, β_G=-.43), 19.8 % (β_L=-.26, β_G=-.50), and 20.9 % (β_L=-.32, β_G=-.51) of the variances of the parent-reported concerns in the areas of receptive and expressive language, and communication, respectively.

LGC and language and verbal working memory skills at 5;3 (n = 102). Two factors were constructed from the tasks administered at the age of 5;3 months to a subsample of the children (24 % of the children in the LGC model). The language factor included the Similarities, PPVT, Verbal Fluency, and Sentence Comprehension tasks. The memory factor included the Digit Span forwards and backwards, Nonword Repetition, and Sentence Repetition tasks. The two parts of the Digit Span task were allowed to correlate. The resulting model fitted the data well: $\chi^2(207) = 258.160$, p =.0090, CFI =.973, TLI =.967, RMSEA = .024,and SRMR =.089. The common growth factor of early communication skills explained 33.4 % (β_G=.65) of the variance in the language factor, and 53.3 % (β_G =.74) of the variance in the memory factor. The at-risk status was added to the model as a covariate in order to control for the possible effects it may have on the follow-up outcome. However, the connections were not significant (language β =.02, p =.85; memory β =-.09, p =.29).

LGC and communication skills in the first grade (n = 236). Parents reported strengths and difficulties in language and communication using the Children’s Communication Checklist-II when their children were in the first grade. These data were available for 55 % of the children in the LGC model. Two factors were constructed from the CCC-II subscales based on their content. The language factor included the subscales Speech, Syntax, Semantics, and Coherence. The communication factor included the subscales Inappropriate initiation, Stereotyped language, Use of context, and Non-verbal communication. Correlations were allowed within the factors for Speech and Syntax, and Stereotyped language and Non-verbal communication. The correlation between the factors was .87 (p <.000). The resulting model fitted the data well: $\chi^2(206) = 236.476$, p =.0714, CFI =.991, TLI =.988, RMSEA = .018, and SRMR =.071. The common level and growth factors explained 48.2 % (β_L=-.31, β_G=-.78) and 38.9 % (β_L=-.24, β_G=-.70) of the variances in the language and communication factors.

Skill- and age-specific pathways. All in all, there were 108 possible specific pathways (12 outcome measures x 9 specific factors), and thus the significance level was set at $p < .001$. Of these possible pathways, 17 were tested based on the model modification indices (see Table 3 and online Supporting Material). None of these pathways were significant at the .001 significance level, while three of these pathways approached significance: the growth factor of the speech composite to the memory factor at age 4;7 (p =.002) and to the language factor in first grade (p =.007), and the age-specific factor at 15 months of age to language in first grade (p =.007).

Summary of the common and specific connections. The initial level and, in particular, the growth of early communication skills were significant predictors of later language ability. The children who had a higher initial level of communication skills showed better productive vocabulary at 2 years and better receptive vocabulary at 3 years, and their parents reported fewer language- and communication-related concerns at age 4;7 and in first grade. Children who had a faster rate of growth in early communication skills during the period from 6 to 18 months showed better vocabulary skills at ages 2 and 3 years, had fewer parent reported concerns about language and communication development at 4;7 and in first grade, and showed better language and verbal working memory skills at age 5;3. The percentage explained by the level and growth factors varied from 10.5 to 53.3 %. The model modification indices suggested several skill- and age-specific pathways, but none of these pathways were significant at the .001 significance level, and thus, no specific paths were added to the regression models.

Discussion

The aim of this study was to explore the co-development of several early communication skills during the prelinguistic period, and the associations between this development and later language ability. Early communication skills showed fast growth throughout the ages from 6 to 18 months. There were large individual differences in the development of these skills and these differences showed rather high stability throughout the prelinguistic period. The development in different early communication skills showed a large amount of shared variance which was indicated by the significant and high loadings on the common level and growth factors. However, despite the notable commonalities in development across skills, significant skill- and age-specific variance was also present. The follow-ups were conducted at several time points using multiple measures and sources. The level and especially the growth of early communication skills were significant predictors of later language ability, explaining between 10.5 to 53.3% of the outcome variances. No reliable skill- or age-specific connections were found. The results support a continuum from prelinguistic to linguistic ability (Bruner, 1983), and the dimensional view of language ability (Rescorla, 2009, 2013), by indicating that the individual variations in early language endowment show consistency that extends far beyond the toddler years.

The development of early communication skills between 6 and 18 months of age

Marked growth was evident in all three areas of early communication development (social, speech, and symbolic) across the age span from 6 to 18 months. Also evident was large interindividual variation in the development of these early skills, as shown by the overlap in scores across the different ages. However, despite the change in the mean scores at group level and large interindividual variation, both continuity and stability were present. The high correlations between the successive measurements suggested continuity in individual differences in these skills over age, whereas the significant loadings of the measurements at each age on the skill-specific level and growth factors indicated stability. That is, within a composite, the relative standings of individuals in their development were rather consistent over time.

As expected, the LGC model suggested a large amount of shared variance in the development of early social, speech, and symbolic skills. The level factors of the three composites loaded significantly on the common level factor, indicating that individuals tended to have a similar ranking relative to others across the three composites. In other words, individuals performing high in one composite were also likely to perform high in the other two. Likewise, the three growth factors of the three composites loaded significantly on the common growth factor, indicating that the relative pace of development was similar across the composites. For example, individuals who showed slow development in one skill composite tended to be slow in the other two as well. Development in the social and symbolic composites, in particular, seemed to go side by side.

Thus, when multiple early communication skills were assessed repeatedly with a parental screener, early communication ability emerged as a rather continuous and stable characteristic of individual differences during the prelinguistic period, that is, the first two years of life. This is in line with previous results on early communication skills obtained by observational methods (Watt et al., 2006). However, the common level and growth factors did not explain all of the variation in development, as shown by the significant residual variances of the skill-specific factors and the emergence of age-specific factors. This indicates that notable skill- and age-specific variation was also present in early communication development, an issue we will turn to later on.

The development of prelinguistic communication skills and later language ability

Both common level and growth of early communication were significant predictors of later language and communication ability. Together, the common level and growth factors explained 10.5 to 53.3 % of the variance in the follow-up measures. These percentages are comparable to those obtained by Wetherby et al. (2002; 2003), who found that the ITC, filled in between 12 and 24 months, explained 20-51 % of the variances in receptive and expressive language outcomes at 2 and 3 years of age. Our results consolidate and expand these results by suggesting that the predictive relation between early communication skills and later language ability is present as early as at 6 months of age, and holds longitudinally up until 8 years of age.

The connection was stronger for the common growth factor, which was significantly connected to all of the follow-up language and communication measures from age 2 years to first grade. A faster rate of growth in early communication skills consistently led to better language ability and fewer parent-reported difficulties later on. Our results support the suggestion that the pace of development, rather than the level at any given age, better reflects the language acquisition potential of the child, and thus might predict later language development more accurately (Rowe et al., 2012). A slow pace of development might be an indicator of risk for later language difficulties and could be useful in identifying which children might go on to have persistent language difficulties (Lyytinen et al., 2006; Rescorla et al., 2000).

The amount of variance explained by the early communication model did not decrease over time, nor was it consistently the largest when the same source of information was used (parental reports). In addition, it did not seem to be dependent on whether the follow-up sample was based on a selected subsample (ages 2, 3, and 5;3) or the full sample (ages 4;7 and first grade). Thus, our results cannot be accounted for solely by the temporal closeness of the assessments, shared source variance, or sampling procedures. Interestingly, the strongest predictive relation was found between the growth of early skills and later verbal working memory capacity. Although based on a selected subsample of children at age 5;3, we do not believe this finding results from sampling issues, since there were no significant differences in early communication skills (the ITC) between the children who participated in the follow-up at 5;3 and those who did not, and since the at-risk status was not a significant covariate. Instead of being an isolated cognitive skill, language has been shown to be rather inextricably linked to a set of processes shared with other realms of cognition early in life (including memory, attention, and processing speed; e.g., Rose, Feldman, & Jankowski, 2009). Memory processes are considered to be an important underlying component of language development (Gathercole, 2006; Heimann et al., 2006), and have been shown to yield a significant level of consistency over time (Rescorla, 2013). Our results seem to fit in with these findings. Thus, the development and application of infant measures that can tap early memory skills more specifically than the ITC in order to provide more information about the co-development of early language and memory, might prove worthwhile in trying to predict language outcome.

Skill- and age-specific factors and later language ability

The common level and growth factors captured the variance shared by the three composites across age. However, as indicated by the significant residual variances of the skill-specific factors and the emergence of age-specific factors at 9, 12 and 15 months of age, significant skill- and age-specific variation was also apparent in early communication development. The contribution of these factors to later language development was examined by testing the specific pathways suggested by the model modification indices. Several such pathways were tested, but, contrary to our expectations, none of them reached significance. These results seem to be in line with those of previous studies reporting that when other aspects of communicative development are controlled for, the unique contributions of specific skills diminish (Lyytinen et al., 2001; Salley et al., 2013) and that the amount of shared variance is notably large relative to the unique contributions (Watt et al., 2006; Wetherby et al., 2002). However, despite the large body of research on prelinguistic predictors of language development, studies that have considered multiple concurrent predictors and their unique contributions to later development remain scarce and the results are not able to lead to firm conclusions. While the tested pathways failed to reach significance in the present study, they nonetheless raise interesting topics for future research.

There are several possible reasons we did not find any reliable specific associations. As indicated by the significant correlations found between early socio-cognitive skills in earlier studies (Laakso et al., 1999; Watt et al., 2006; Wetherby at al., 2002), these skills are highly interrelated, and thus specific connections might be difficult to discern. It is possible that a broadband screener is only able to describe the common trends in development, and that to be able to capture the more specific processes of development, a more fine-tuned measurement is needed. For example, Laakso et al. (1999) found that parental ratings of intentional communication yielded general associations to later language ability, whereas the associations from observed joint attention to later language varied depending on the specific aspects of the joint attention behaviors under observation. Watt et al. (2006) also found, using observational methods, that some skills make a unique contribution to language outcome despite sharing a substantial amount of variance. In addition, Brooks and Meltzoff (2008) found that observed pointing in an experimental setting, but not parental report of pointing, was a significant predictor of vocabulary growth. Specific predictive relations might also be stronger when the focus of assessment is centered more on domain-general skills, such as attention (Salley et al., 2013) and memory (Heimann et al., 2006; Rose et al., 2009), which have been shown to show discreteness already early on in life (Rose, Feldman, & Jankowski, 2005).

Strengths, limitations and further directions

This study examined the continuity and stability of language and communication development from 6 months to first grade. So far, few studies have examined developmental continuity and predictive relations starting from such an early age and extending over a notably long follow-up period (however, see Reilly et al., 2006). The use of a rather large community-based sample, repeated assessment of early communication skills during the prelinguistic period, the inclusion of social and symbolic abilities in addition to oral communication in the early assessments, and a diverse assessment of language and communication with the inclusion of working memory measures in the follow-ups are clear assets of the study.

Although the present study established that strong longitudinal associations exist between prelinguistic development and later language ability, several important limitations must be noted. First, despite its initial size, our study is limited by the nature of the sample. Due to sampling decisions and attrition, the number of children having data at each follow-up ranged between 23 – 26% in selected subsamples, and between 60 - 63% in population follow-up samples. The variation in the subsample sizes and measures makes comparison of the coefficients of determination (R^2) challenging. However, these values did not systematically vary according to the coverage of the initial sample or the source of information (parent report or psychometric assessment). It is also important to bear in mind that in some of the follow-up samples (2;0, 3;0, and 5;3) children with possible risks for language difficulties were slightly oversampled and thus, inferences based on these follow-ups should be interpreted with caution. However, although there were some differences in the ITC composite scores between the children participating and not participating in these follow-ups, these differences did not seem to be systematic across the follow-ups.

Second, our parental questionnaire on early communication skills was comprehensive in its contents, but nevertheless short and designed for screening purposes. Parent-report measures of communication and language skills have been shown to be reliable and valid, and to correlate with concurrent and subsequent behavioral measures (e.g., Feldman et al., 2005; Laakso et al., 1999). However, more direct assessments of the possible (socio)-cognitive precursors of language, such as working memory and processing efficiency (e.g., Fernald & Marchman, 2012), attentional capacity (Rose et al., 2009), joint attention and gestures (Beuker et al., 2013), and symbolic play (Bruce et al., 2003) in infancy, would aid in better understanding the processes that underlie the emergence and further development of language. Our results suggest that the role of early working memory especially should be studied further. From a clinical perspective, however, studying feasible and implementable methods is essential.

Third, we did not control for possible confounding variables related to the child (e.g., gender, nonverbal ability) or the family (e.g., parental age and education) that could affect the estimation of stability (see, Bornstein et al., 2013; McKean et al, 2015). It has been shown that these variables show little explanatory power in the early stages of development (5-6% in Reilly et al., 2006). At later ages, they have been shown to contribute more to later language status (19-21% in Reilly et al., 2010), but to show only modest discrimination between children with and without low language ability. For the current study, data related to birth and family are, however, reported in the Supporting material along with descriptions of Finnish society. Fourth, it is likely that there are other important risk or resilience factors that contribute to later language ability that we did not assess in this study. As language develops in social interaction, factors related to the social environment might prove useful (Bruner, 1983; Paavola et al., 2006). For example, maternal responsivity has been shown to mediate the relation between early communication and later language (Yoder & Warren, 1999). In addition, it is likely that the use of speech and language therapy services affect children's outcomes (Law, Garrett, Nye, 2003). Unfortunately, we did not have this information for the whole sample, and thus could not control for it.

Finally, it is important to bear in mind that the results obtained in this study reflect predictive relations at the group level. Studying stability and prediction at the individual level was not within the scope of this study. However, examination of the persistency of at-risk status and estimations of sensitivity, specificity and other predictive values is a natural next step in our research. The sensitivity and specificity estimates obtained in previous studies using the ITC have been relatively good (81-89 % and 70-79 %) up to three years of age (Wetherby et al., 2003).

The present study contributes to the literature on early language development by adducing further evidence for the link between prelinguistic communication and later language ability. The results support the dimensional view of language ability (Rescorla, 2009, 2013) by showing that instead of a specific skill, the combined development of several early communication skills (i.e., the early language endowment) is more predictive of later language ability. Our results conform to the views of Bates (2004) and Rescorla (2009, 2013) that language ability builds up from the interactions of several interrelated early communication skills that do not map one-on-one to later abilities. Our results also advocate developmental surveillance of early communication skills by showing that the pace of development, rather than the level at any given age, is more predictive of later development. In addition to providing a better understanding of developmental processes in typical development, these results have important implications for early screening procedures. By assessing multiple early communication skills concurrently, and by following the development in these skills over short periods of time, we might be able to identify at-risk children more accurately and at an earlier age.

References

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36*, 189-208. doi: 10.1016/S0021-9924(03)00019-4

Bates, E. A. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? *Brain and Language, 88*, 248-253. doi: 10.1016/S0093-934X(03)00102-0

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology, 40*, 293–310. doi:10.1002/dev.10034

Beuker, K. T., Rommelse, N. N. J., Donders, R., & Buitelaar, J. K. (2013). Development of early communication skills in the first two years of life. *Infant Behavior and Development, 36,* 71-83. doi: 10.1016/j.infbeh.2012.11.001

Bishop, D. V. M. (2003). *The Children's Communication Checklist* (Vol 2). London: Harcourt Assessment.

Bishop, D. V. M. (2014). *The Children's Communication Checklist* (Vol 2). [Manual for the Finnish adaptation]. Helsinki: Hogrefe Psykologien Kustannus.

Bollen, K. A., & Curran, P. J. (2006*). Latent curve models: A structural equation perspective*. New Jersey: John Wiley & Sons.

Bornstein, M. H., Hahn, C.-S., Putnick, D. L., & Suwalsky, J. T. D. (2013). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development*. Advance online publication. doi: 10.1111/cdev.12192

Bornstein, M. H. & Putnick, D. L. (2012). Stability of language in childhood: a multiage, multidomain, multimeasure, and multisource study. *Developmental Psychology, 48,* 477-491. doi: 10.1037/a0025889

Brooks, R. & Meltzoff, A. N. (2008) Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. *Journal of Child Language, 35*, 207-220. doi:10.1017/S030500090700829X

Bruce, B, Kornfält, R., Radeborg, K., Hansson, K., & Nettelbladt, U. (2003). Identifying children at risk for language impairment: screening of communication at 18 months. *Acta Paediatric, 92,* 1090-1095. doi: 10.1080/08035250310004414

Bruner, J. (1983). *Child's talk: Learning to use language*. Oxford: Oxford University Press.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*, 1-143.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*, 98–101. doi:10.1111/1467-8721.ep10768783

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: a meta-analysis. *Developmental review, 30*, 352-366. doi:10.1016/j.dr.2010.10.001

Conti-Ramsden, G., & Durkin, K. (2012). Language development and assessment in the preschool period. Neuropsychology Review, 22, 384–401. doi:10.1007/s11065-012-9208-z

Crais, E. R. (2011). Testing and beyond: Strategies and tools for evaluating and assessing infants and toddlers. *Language, Speech, and Hearing Services in Schools, 42,* 341-364. doi: 10.1044/0161-1461(2010/09-0061)

Darrah, J., Hodge, M., Magill-Evans, J., & Kembhavi, G. (2003). Stability of serial assessments of motor and communication abilities in typically developing infants – implications for screening. *Early Human Development, 72*, 97– 110. doi: 10.1016/S0378-3782(03)00027-6

Duncan, T. E., Duncan, S. C., & Strycker L. A. (2006). An introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications. (2nd ed.) Mahwah, NJ: Lawrence Erlbaum Associates.

Dunn, L. M & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised*. Circle Pines. MN: American Guidance Service.

Feldman, H. M., Dale, P. S., Campbell, T. F., Colborn, D. K., Kurs-Lasky, M., Rockette, H. E., &

Paradies, J. L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development, 76*, 856-868. doi: 10.1111/j.1467-8624.2005.00882.x

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*, 5 (Serial No. 242).

Fernald, A. & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83*, 203-222. doi: 10.1111/j.1467-8624.2011.01692.x

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513-543. doi: 10.1017.S0142716406060383

Heimann, M., Strid, K., Smith, L., Tjus, T., Ulvund, S. E., & Meltzoff, A. N. (2006). Exploring the relation between memory, gestural communication, and the emergence of language in infancy: A longitudinal study. *Infant and Child Development, 15,* 233-249. doi: 10.1002/icd.462

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi: 10.1080/10705519909540118

Kadesjö, B., Janols, L.-O., Korkman, M., Mickelsson, K., Strand, G., Trillingsgaard, A., & Gillberg, C. (2004). The FTF (Five to Fifteen): The development of a parent questionnaire for the assessment of ADHD and comorbid conditions. *European Child and Adolescent Psychiatry, 13 (Suppl. 3)*, 3–13. doi:10.1007/s00787-004-3002-2

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test.* Philadelphia: Lea & Febiger

Korkman, M., Kirk, U., & Kemp, S. L. (1997). *Lasten neuropsykologinen tutkimus (NEPSY: A Developmental Neuropsychological Assessment.*). Helsinki, Finland: Psykologien kustannus Oy.

Korkman, M., Kirk, U., & Kemp, S. L. (2008). *NEPSY-II: Lasten neuropsykologinen tutkimus (NEPSY-II: A Developmental Neuropsychological Assessment, Second edition).* Helsinki, Finland: Psykologien kustannus Oy.

Korpilahti, P. (1996). *Lausetesti (Test for the language comprehension of sentences).* Helsinki: University Press.

Laakso, M.-L., Poikkeus, A.-M., & Eklund, K. (2011). *Lapsen esikielellisen kommunikaation ja kielen ensikartoitus* [The Infant-Toddler Checklist of Children's Prelinguistic Communication and Language]. Jyväskylä, Finland: Niilo Mäki Instituutti.

Laakso, M.-L., Poikkeus, A.-M., Katajamäki, J., & Lyytinen, P. (1999). Early intentional communication as a predictor of language development in young toddlers. *First Language, 19,* 207-231. doi: 10.1177/014272379901905604

Law, J., Garrett, Z., & Nye, C. (2003). Speech and language therapy interventions for children with primary speech and language delay or disorder. The Cochrane Database of Systematic Reviews. Issue 3. Art. No: CD004110. Retrieved from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004110/full

Lyytinen, H. Erskine, J., Tolvanen, A., Torppa, M., Poikkeus, A.-M., & Lyytinen, P. (2006). Trajectories of reading development: a follow-up from birth to school age of children with and without risk for dyslexia. *Merrill-Palmer Quarterly, 52*, 514-546. doi: 10.1353/mpq.2006.0031

Lyytinen, P. (1999). *Varhaisen kommunikaation ja kielen kehityksen arviointimenetelmä* [Finnish Manual for Communicative Development Inventories]. Jyväskylän yliopiston lapsitutkimuskeskus ja Niilo Mäki Instituutti. Jyväskylä: Yliopistopaino

Lyytinen, P., Poikkeus, A.-M., Laakso, M.-L., Eklund, K., & Lyytinen, H. (2001). Language development and symbolic play in children with and without familial risk for dyslexia. *Journal of Speech, Language, and Hearing Research, 44,* 873-885. doi: 10.1044/1092-4388(2001/070)

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches for setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11,* 320-341. doi:

10.1207/s15328007sem1103_2

McCathren, R. B., Warren, S. F., & Yoder, P. J. (1996). Prelinguistic predictors of later language development. In K. N. Cole, P. S. Dale, & D. J. Thal (Eds.), Assessment of communication and language (pp. 57–75). Baltimore, MD: Brookes

McKean, C., Mensah, F. K., Eadie, P., Bavin, E. L., Bretherton, L., Cini, E., & Reilly, S. (2015). Levers for language growth: characteristics and predictors of language trajectories between 4 and 7 years. PLoS ONE 10(8): e0134251. doi:10.1371/journal.pone.0134251

Miles, J. & Shevlin, M. (2007). A time and a place for incremental fit indices. *Personality and Individual Differences, 42*, 869-874. doi: 10.1016/j.paid.2006.09.022

Muthén, L. K. and Muthén, B. O. (1998-2010). *Mplus User's Guide.* Sixth Edition. Los Angeles, CA: Muthén & Muthén

Määttä, S., Laakso, M.-L., Tolvanen, A., Ahonen, T., & Aro, T. (2012). Developmental trajectories of early communication skills. *Journal of Speech, Language, and Hearing Research, 55*, 1083-1096. doi: 10.1044/1092-4388(2011/10-0305

Määttä, S., Laakso, M.-L., Tolvanen, A., Ahonen, T., & Aro, T. (2014). Children with differing developmental trajectories of prelinguistic communication skills: Language and working memory at age 5. *Journal of Speech, Language, and Hearing Research, 57*, 1026-1039. doi: 10.1044/2014_JSLHR-L-13-0012

Paavola, L., Kemppinen, K., Kumpulainen, K., Moilanen, I., & Ebeling, H. (2006). Maternal sensitivity, infant co-operation and early linguistic development: Some predictive relations. *European Journal of Developmental Psychology, 3*, 13-30. doi: 10.1080/17405620500317789

Reilly, S., Eadie, P., Bavin, E. L., Wake, M., Prior, M., Williams, J., … Ukoumunne, O. C. (2006). Growth of infant communication between 8 and 12 months: A population study. *Journal of Paediatrics and Child Health, 42,* 764–770. doi:10.1111/j.1440-1754.2006.00974.x

Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., … Ukoumunne, O. C. (2007).

Predicting language at 2 years of age: A prospective community study. *Pediatrics, 120*, e1441-e1449. doi: 10.1542/peds.2007-0045

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., … Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Finding from Early Language in Victoria Study. *Pediatrics, 126*, e1530-e1537. doi: 10.1542/peds.2010-0254

Rescorla, L., (2009). Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research, 52*, 16–30. doi:10.1044/1092-4388(2008/07-0171)

Rescorla, L. (2013). Late-talking toddlers. A 15-year follow up. In L. A. Rescorla & P. S, Dale (Eds.), *Late Talkers – Language, Development, Intervention, and Outcomes* (pp. 219-239). Baltimore, MD: Brookes.

Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: Lexical development from 2;0 to 3;0. *Journal of Child Language, 27,* 293–311. doi:10.1017/S030500090000413X

Rose, S. A, Feldman, J. R., & Jankowski, J. J. (2005). The structure of infant cognition at 1 year. *Intelligence, 33*, 231-250. doi: 10.1016/j.intell.2004.11.002

Rose, S. A, Feldman, J. R., Jankowski, J. J., & Van Rossem, R. (2008). A cognitive cascade in infancy: Pathways from prematurity to later mental development. *Intelligence, 36*, 367-378. doi: 10.1016/j.intell.2007.07.003

Rose, S. A, Feldman, J. R., & Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Development, 80*, 134-150. doi: 10.1111/j.1467-8624.2008.01250.x

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development, 83,* 508-525. doi: 10.1111/j.1467-8624.2011.01710.x.

Salley, B., Panneton, R. K., & Colombo, J. (2013). Separable attentional predictors of language outcome. *Infancy, 18*, 462-489. doi: 10.1111/j.1532-7078.2012.00138.x

Thal, D. J., Bates, E., Goodman, J., & Jahn-Samilo, J. (1997). Continuity of language abilities: An exploratory study of late- and early-talking toddlers. *Developmental Neuropsychology, 13,* 239–273. doi:10.1080/87565649709540681

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic predictors of language outcome at 3 years of age. *Journal of Speech, Language, and Hearing Research, 49,* 1224-1237. doi: 10.1044/1092-4388(2006/088)

Wechsler, D. (1995). *WPPSI-R Käsikirja* [Wechsler Preschool and Primary Scale of Intelligence: Revised. Manual for the Finnish Adaptation]. Helsinki: Psykologien kustannus Oy.

Wechsler, D. (1999). *WISC-III Käsikirja* [Manual for the Wechsler Intelligence Scale for Children–Third Edition. Manual for the Finnish Adaptation]. Helsinki: Psykologien kustannus Oy.

Wetherby, A. M., Allen, L., Cleary, J., Kublin, K., & Goldstein, H. (2002). Validity and reliability of the Communication and Symbolic Behavior Scales Developmental Profile with very young children. *Journal of Speech, Language, and Hearing Research, 45,* 1202-1218. doi: 10.1044/1092-4388(2002/097)

Wetherby, A. M., Brosnan-Maddox, S., Peace, V., & Newton, L. (2008).Validation of the Infant-Toddler Checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism, 12,* 487-511. doi: 10.1177/1362361308094501

Wetherby, A. M., Goldstein, H., Cleary, J., Allen, L., Kublin, K. (2003). Early identification of children with communication disorders: Concurrent and predictive validity of the CSBS Developmental Profile. *Infants and Young Children, 16*, 161-174. Retrieved from http://journals.lww.com/iycjournal/

Wetherby, A. M., & Prizant, B. M. (2002). *Communication and Symbolic Behavior Scales: Developmental Profile* (1st normed ed.). Baltimore, MD: Brookes.

Yoder, P. J. & Warren, S. F. (1999). Maternal responsivity mediates the relationship between prelinguistic intentional communication and later language. *Journal of Early Intervention, 22*, 126-136. doi: 10.1177/105381519902200205

Figure captions.

Figure 1. Flowchart of the study participants. MBCDI = MacArthur-Bates Communicative Development Inventories; PPVT = Peabody Picture Vocabulary Test.
[1] The numbers of participants in the present study are given in parentheses.
[2] The *n* for PPVT was 111.
[3] Risk in language, attention or hyperactivity based on FTF 90th percentile cut-off.

Figure 2. Latent growth curve model for early communication skills. Standardized estimates are presented. The first time points of the slope factors are fixed to 0 and, along with nonsignificant paths, are omitted from the figure. Lev = level; Slo = slope (growth); Soc = social composite; Spe = speech composite; Sym = symbolic composite; Sf = specific factor. Numbers after soc, spe, sym, and sf represent age in months.

Supplemental material description

Supplemental material includes a more detailed description of the study participants, measures, data, and analysis methods.

Table 1.

Communication and Language Measures Used in the Study in Each Age Stage

Age	***n***	**Source**	**Measure**	**Components**	**Scoring**
6, 9, 12, 15, 18, 21, and 24 months	203-322	Parent report	Communication and Symbolic Behavior Scales – Infant Toddler Checklist (ITC)	Social	Sum of 13 ordinal items on the ITC measure (max. 26)
				Speech	Sum of five ordinal items on the ITC measure (max. 14)
				Symbolic	Sum of six ordinal items on the ITC measure (max. 17)
2 years	104	Parent report	MacArthur-Bates Communicative Development Inventories – Words and Sentences Inventory	Vocabulary	Number of words from a pre-specified list of 595 that the child says
				Inflections	Sum of noun and verb inflections the child uses (max. 16)
				Maximum sentence length	Average morpheme length of three sentences
3 years	112	Individual assessment	Boston naming	Expressive vocabulary	Sum of 60 binary items
			Peabody Picture Vocabulary Test	Receptive vocabulary	Sum of 166 binary items
4;7	253	Parent report	Five to Fifteen (FTF) – Language subscales	Comprehension	Mean of five ordinal items on the FTF measure
				Expressive	Mean of 13 ordinal items on the FTF measure
				Communication	Mean of three ordinal items on the FTF measure
5;3	98-102	Individual assessment	WPPSI-R – Similarities	Verbal reasoning	Sum of 12 binary and eight ordinal items (max. 28)
			Peabody Picture Vocabulary Test-Revised	Receptive vocabulary	Sum of 30 binary items
			Korpilahti Auditory Sentence Comprehension Test	Receptive grammar	Sum of 30 binary items
			NEPSY-II – Verbal Fluency	Verbal productivity	Sum of semantically correct words produced in 60 seconds
			WISC-III – Digit Span	Working memory	Sum of 12 (forward) and 10 (backward) binary items
			NEPSY – Repetition of Nonsense Words	Working memory	Sum of 16 binary items
			NEPSY-II – Sentence Repetition	Working memory	Sum of 17 ordinal items (max. 34)
1st grade	236	Parent report	Children's Communication Checklist - II	Language: Speech, Syntax, Semantics, Coherence	Four subscales, sum of seven ordinal items (5 addressing deficits, 2 strengths) in each scale (max. 112)
				Communication: Inappropriate initiation, Stereotyped language, Use of context, Non-verbal communication	Four subscales, sum of seven ordinal items (5 addressing deficits, 2 strengths) in each scale (max. 112)

Note. WPPSI-R = Wechsler Preschool and Primary Scale of Intelligence –Revised; NEPSY(-II) = A Developmental Neuropsychological Assessment (- Second edition); WISC-III = Wechsler Intelligence Scale for Children – Third edition.

Table 2

Means and Standard Deviations of the Early Communication and Follow-Up Measures

Early communication measures	Max.	**6 mo.** *n* = 229 *M (SD)*	**9 mo.** *n* = 203 *M (SD)*	**12 mo.** *n* = 322 *M (SD)*	**15 mo.** *n* = 305 *M (SD)*	**18mo.** *n* = 279 *M (SD)*
ITC Social	26	9.99 (2.78)	14.30 (3.63)	19.61 (3.42)	21.73 (3.01)	23.06 (2.59)
Speech	14	3.03 (1.75)	5.83 (2.30)	7.64 (2.34)	9.49 (2.26)	11.17 (2.12)
Symbolic	17	3.72 (1.60)	6.32 (1.81)	9.60 (2.31)	12.88 (2.16)	14.94 (1.79)
Follow-up measures	Max.	**2 y.** *n* = 104 *M (SD)*	**3 y.** *n* = 111-112 *M (SD)*	**4;7 y.** *n* = 253 *M (SD)*	**5;3 y.** *n* = 98-102 *M (SD)*	**1st grade** *n* = 236 *M (SD)*
MBCDI Vocabulary	595	325.10 (154.10)				
Inflections	16	10.36 (4.90)				
MSL		7.0 (3.62)				
Boston naming	60		14.88 (5.30)			
PPVT	166		23.87 (12.15)			
FTF[a] Comprehension	2			0.23 (0.29)		
Expressive	2			0.21 (0.28)		
Communication	2			0.26 (0.39)		
SI	28				16.42 (4.04)	
PPVT-R	30				16.67 (3.63)	
SC	30				21.25 (4.17)	
VF					15.56 (6.08)	
DS	22				6.55 (2.41)	
NWR	16				9.53 (2.72)	
SR	34				21.25 (4.68)	
CCC-II[a] Language	112					5.07 (8.50)
Communication	112					7.75 (9.06)

Note. All means are calculated from nonstandardized sum scores with the exception of the FTF subscales that are calculated from scale means due to missing items (nine subjects, maximum of three missing values). ITC = Infant-Toddler Checklist; MBCDI = MacArthur-Bates Communicative Development Inventories; MSL = Maximum Sentence Length; PPVT = Peabody Picture Vocabulary Test; FTF = Five to Fifteen; SI = Similarities; SC = Sentence Comprehension; VF = Verbal Fluency; DS = Digit span; NWR = Nonword Repetition; SR = Sentence Repetition; CCC-II = Children's Communication Checklist –Second Edition.

[a] = higher value represents more difficulties

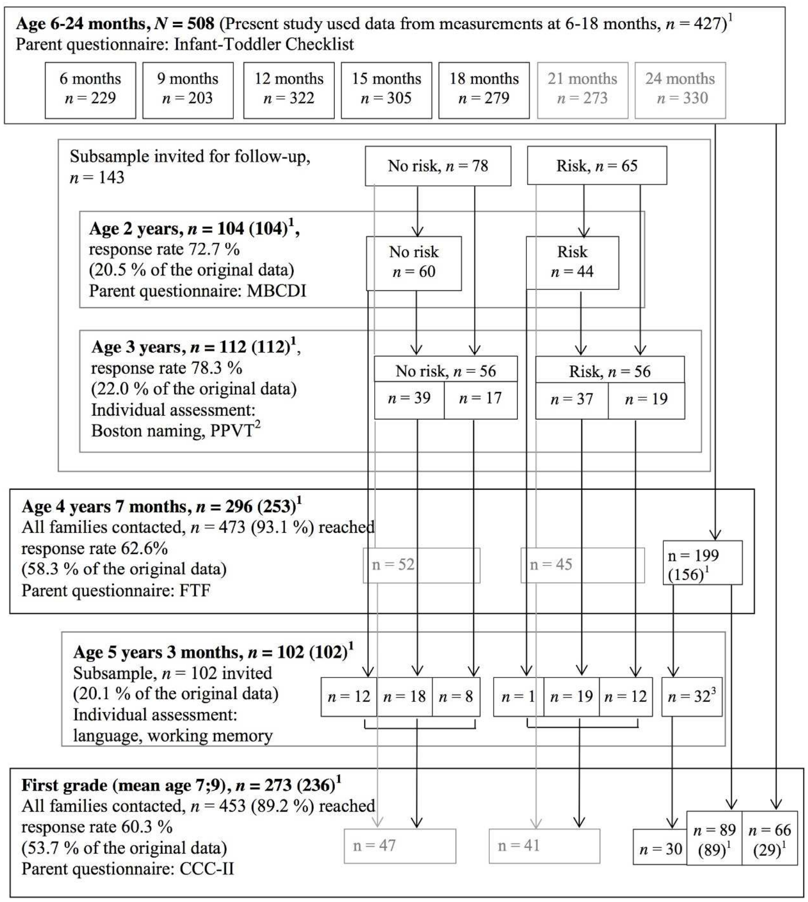
Table 3.

Summary of the Regression Analyses Predicting Later Language Ability from Early Communication Development

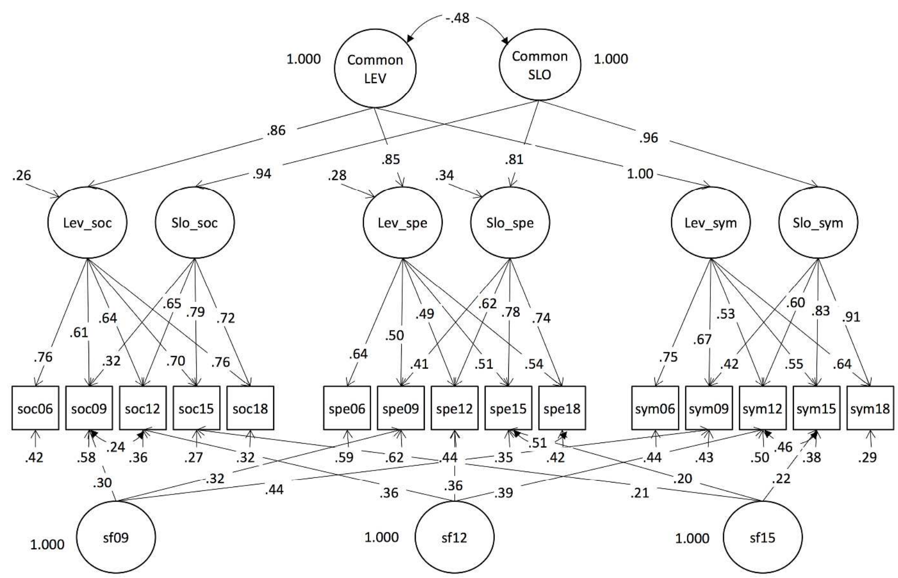
			Common			Specific associations[a]								
						Social		Speech		Symbolic		Age specific		
Follow-up measure	Age (y)	*n*	Level	Growth		Level	Growth	Level	Growth	Level	Growth	Sf 09	Sf 12	Sf 15
			β		R^2 %	β		β		β		β		
MBCDI Vocabulary	2	104	0.32^{*}	0.64^{***}	32.4									0.29
Inflections	2	104	0.12	0.69^{***}	41.1		-0.28	0.22						-0.26
MSL	2	104	0.17	0.74^{***}	46.7									
Boston naming	3	112	0.17	0.58^{***}	27.6				0.24					
PPVT	3	111	0.29^{*}	0.33^{***}	10.5									
FTF Comprehension	4;7	253	-0.28^{**}	-0.43^{***}	15.0									
Expressive	4;7	253	-0.26^{*}	-0.50^{***}	19.8		0.28		-0.23					
Communication	4;7	253	-0.32^{**}	-0.51^{***}	20.9									
Language (SI, PPVT, SC, VF)	5;3	102	0.24	0.65^{***}	33.0									
Memory (DSf, DSb, NWR, SR)	5;3	102	0.03	0.74^{***}	53.3		-0.50	0.47	0.45					-0.47
CCC-II Language	1st gr	236	-0.31^{*}	-0.78^{**}	48.2			-0.24	-0.20					0.23
Communication	1st gr	236	-0.24^{*}	-0.70^{***}	38.9			0.24	0.21					-0.17

Note. MBCDI = MacArthur-Bates Communicative Development Inventories; MSL = Maximum Sentence Length; PPVT = Peabody Picture Vocabulary Test; FTF = Five to Fifteen; SI = Similarities; SC = Sentence Comprehension; VF = Verbal Fluency; DS = Digit Span; NWR = Nonword Repetition; SR = Sentence Repetition; CB = Corsi Block task; CCC-II = Children's Communication Checklist – Second Edition; Sf = Specific factor (and age in months).
a. Each specific pathway was tested separately. The significance level was set to p<. 001.
*<.05. **<.01. ***.001.

Running head: CONTINUITY OF COMMUNICATION AND LANGUAGE DEVELOPMENT



1
215x279mm (300 x 300 DPI)



2
279x215mm (300 x 300 DPI)

Continuity from prelinguistic communication to later language ability: a follow-up study from infancy to school age

Appendix S1

Method

Participants and procedure

The original sample of 508 children was recruited through community-based child health care clinics in the city of Jyväskylä, Central Finland. Child health care clinics provide free services for all families with children between ages 0 to 6. The services are focused on health promotion, risk assessment, and disease prevention. Visits are made to the clinic 10 to 15 times during the first two years of life, and thereafter annually or at 18 month intervals. The clinics are regularly attended by over 95 % of Finnish parents and their children (for a more detailed description of pre- and postnatal care for families, see Callister, Lauri, & Vehviläinen-Julkunen, 2000). All the clinics in the area (population base close to 100,000, and age cohort of about 900 at the time) volunteered to participate in the study. The Infant-Toddler Checklist (ITC, part of the Communication and Symbolic Behavior Scales – Developmental Profile, Wetherby & Prizant, 2002) was introduced to the families by the nurses at the clinics. Children were eligible for participation if aged between 6 and 24 months at time of recruitment.

After giving their consent and completing the first ITC questionnaire, parents were asked to fill in a new questionnaire every three months until the child was 24 months of age (i.e., a maximum of seven times; at ages 6, 9, 12, 15, 18, 21, and 24 months). The sample sizes for the measurement points were $n = 229$ at 6 months, $n = 203$ at 9 months, $n = 322$ at 12 months, $n = 305$ at 15 months, $n = 279$ at 18 months, $n = 273$ at 21 months, and $n = 330$ at 24 months of age. For the majority of the sample (67.9%), data were available from at least three measurement points. The total number of forms filled in by parents depended on their child's age at the time of completing the first questionnaire and on how many of the subsequent questionnaires they completed. In the present study, the data from the measurements conducted between 6 and 18 months were used, yielding a total sample of 427 children.

After the early questionnaire data collection phase, subgroups of the original 508 participants were followed at ages 2 and 3 years (parent report and individual assessments), 4;7 (parent report), and 5;3 (individual assessment), and in the spring term of the first grade (age range 7;5 – 8;4, parent report). During the follow-ups at 2, 3, and 5;3, only a small subset of families were contacted due to time and resource limitations. Thus, the subgroups were constructed to include a sufficient number of children showing possible risks for language and communication development. This form of data collection enables also comparison of at-risk and typically developing children. At ages 4;7 and first grade, all the originally participating families were contacted. The group differences in the early ITC scores are summarized in table SM1. .

Demographic information by subsamples is presented in Table SM2. The Finnish population of 5.4 million is relatively homogeneous in ethnicity, culture, religion, and language. All the participating children were Caucasian and spoke Finnish as their native language. Data related to birth and family were collected at the initial recruitment stage (data available for 472 - 485 children). Fourteen children (2.9 %) had been born preterm (i.e., gestational age less than 36 weeks). Sample mean birth weight was 3.5 kg (*SD* = 0.6, range 1.1 -5.4). Slightly over half (*n* = 267, 56.6 %) the children were firstborns. At time of recruitment, 19 (3.9 %) families reported single parenthood. This is a markedly lower percentage than in the general population (= 14% of families with children aged 0-7 years during 2003, when the initial data were collected; Statistics Finland, 2013). However, the percentage of single parent families can be expected to be lower among the families of young infants.

Parental education was classified using a seven-point scale ranging from a basic level, 0 (*no* *vocational education*), to advanced educational training, 6 (*higher-level university degree*). The sample was fairly representative of the Finnish population (Statistics Finland, 2013). The distribution of family educational level in the sample was as follows: 7% of mothers and 6% of fathers (general population 6%) had no vocational education, 58% of mothers and 66% of fathers (general population 64%) had at least some vocational degree, and 35% of mothers and 29% of fathers (general population 29%) had a master's or higher university degree. Finnish families are typically dual-earner families with both parents working full time (Salmi & Lammi-Taskula, 2014). State-funded parental leave lasts up to 10 months of age, after which child home-care allowances are provided for the first 1 to 3 years. Around 40% of mothers with children under the age of three years and 80% of mothers of children aged between 3 to 6 years work outside the home (Salmi & Lammi-Taskula, 2014). Child care is provided in day care centers or in family day care, the former of which is more commonly used (84% vs. 16%; Kekkonen, 2014). Rates of day care attendance vary according to the child's age. Around 30% of one-year-olds, 50% of two-year-olds, 70% of three-year-olds, 75% of four-year-olds, and 80% of five-year-olds are in day care. Family day care is more common in the youngest age groups. Children have a right to attend pre-school education the year before their compulsory education starts (the year they turn 6). Pre-school education is provided in day care centers and primary schools. The majority (98%) of children attend pre-school education (Statistics Finland, 2013). Compulsory schooling starts in the year of the child's seventh birthday.

There were small but significant differences in demographics between the children who had data from the last two follow-ups at age 5;3 and first grade (*n* range 100-102 and 230-234) and those who did not (*n* range 373-394 and 241-263): The participants in the last two follow-ups had slightly older and more educated mothers (mother's age: 30.9 vs. 29.6 at 5;3, $p = .031$, $\eta_p^2 = .009$; 30.4 vs. 29.4 in first grade, $p = .039$, $\eta_p^2 = .009$; mother's education: 4.1 vs. 3.7 at 5;3, $p = .050$, $\eta_p^2 = .008$; 4.1 vs. 3.6 in first grade, $p = .002$, $\eta_p^2 = .020$). However, only maternal education in the sample at first grade remained significantly different after correcting for multiple comparisons (Bonferroni correction, nine comparisons). These results are in line with previous observations reported by longitudinal studies of language that attrition tends to be lower among children with older and more educated mothers (e.g., Henrichs et al., 2011; Reilly et al., 2010). No other significant differences between the subsamples were found.

By the end of the study, two children (information was available for 338 children) had received a diagnosis of language impairment and three children were reported as having broader developmental difficulties. In addition, based on parent report, health care providers had observed indications of delayed language development in 17 children (5.0 %). Parents reported the use of speech and language therapy services for language-related difficulties (excluding articulation and stuttering problems) for 11 (3.3 %) children. The discrepancy between the number of children with diagnosed language impairment and those attending speech and language therapy services is probably due to the service structure in Finland. Children do not need a formal diagnosis to be eligible for specialist services. Families are referred to these services if any concerns arise during their annual check-ups at their local child health care clinics. Very often, the first step is to see whether a more intensive follow-up together with family guidance or a few visits to a speech and language therapist is enough before referring the child for further assessments and formal diagnostic procedures.

Measures

Early communication measure. Early communication skills were assessed using the Finnish version of the ITC of the CSBS-DP (Laakso, Poikkeus, & Eklund, 2011; Wetherby & Prizant, 2002). The ITC is a parent-report screening tool that consists of 24 questions designed to measure relevant prelinguistic milestones of early communication and language development in children aged 6 to 24 months. The questions are organized into three composites and cover several areas of development, such as emotion and use of eye gaze, communication, and gestures (social composite, 13 questions);

sounds and words (speech composite, five questions); and understanding and object use (symbolic composite, six questions). The ratings are either on a three-point scale (0 = not yet, 1 = sometimes, 2 = often) or on scales that describe a series of numbers or ranges affording 0 to 4 points (e.g., 0 = none, 1 = 1-3, 2 = 4-10, 3 = 11-30, 4 = over 30). The Cronbach's αs over the age span of 6 to 18 months ranged from .80 to 89, and by age (6, 9, 12, 15, and 18 months; *n*s = 191-320) from .68 to .73 for the social composite, from .47 to .63 for the speech composite, and from .38 to .58 for the symbolic composite. The variations in the alpha values by age are probably due to the fact that the questions for each age are the same, meaning that some of the questions might behave differently at different age stages (such as the number of words spoken or understood).

Measures at 2 years of age. Children's expressive vocabulary was assessed with the Finnish version of the MacArthur-Bates Communicative Development Inventories Words and Sentences (Fenson et al., 1994; Lyytinen, 1999). The checklist contains four subscales that measure vocabulary, use of language, noun and verb inflections, and word combinations in children aged 16 to 30 months. Three of these subscales were used in this study. In the vocabulary scale, the parent indicates which of the predefined 595 words they have heard their child produce spontaneously. The words include nouns, verbs and adjectives that are commonly used by children of this age. A total number of words is calculated for each child. In the inflections scale, the parent indicates which of the 16 inflections (e.g., plural, verb tenses) are present in the child's spontaneous speech. The sum of the noun and verb inflections that the child uses is calculated for each child. In the third section, the parent writes verbatim the three longest sentences they have heard their child produce. Average sentence length, measured as morphemes, is calculated based on these three sentences.

Measures at 3 years of age. The children's single-word receptive vocabulary was assessed with the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1981). The PPVT consists of 166 words accompanied by black-and-white line drawings. The child hears a word and selects the picture that corresponds to the word from an array of four pictures. Total score of correct answers was used in the analyses.

Boston naming (Kaplan, Goodglass, & Weintraub, 1983) was used as a measure of single-word expressive vocabulary. The task consists of 60 pictures that the child has to name. If the child does not produce a word for the picture, he/she is prompted with a semantic cue. If the child fails to produce the word, a phonological cue is given (e.g., the first two sounds of the word). The total number of correct productions is calculated from the words the child produces either spontaneously or with the semantic cue.

Measures at 4;7. Language and communication related concerns were assessed with the Five to Fifteen questionnaire (FTF; Kadesjö et al. 2004). The FTF is a parent questionnaire developed for the elicitation of symptoms and problems typical of ADHD and its comorbidities. The FTF comprises 181 statements related to behavioral or developmental problems. The language domain of the questionnaire consists of 21 questions divided into three subscales. The comprehension subscale (five questions) measures difficulties in understanding words, explanations and stories. The expressive subscale (13 questions) measures difficulties in fluency, word retrieval and complexity of speech. The communication subscale (three questions) measures difficulties in social communication and narration. Ratings are made on a three-point scale (0 = does not apply, 1 = applies sometimes or to some extent, 2 = definitely applies). Due to missing values for some items, the means of the subscales were used in the analyses. The Finnish validation of the FTF for 5-year-olds (n = 769) reported the reliability of the whole language domain to be .89 (Korkman, Jaakkola, Ahlroth, Pesonen, & Turunen, 2004). Cronbach's αs of .84 for comprehension, .84 for expressive, and .75 for communication have been reported (Kadesjö et al., 2004). In the present data the corresponding values were .66, .87, and .71, respectively.

Language measures at 5;3. The language measures were selected to cover various areas of language ability in both the expressive and receptive domains, as suggested by Conti-Ramsden and Durking (2012).

The Similarities (SI) subtest of the Wechsler Preschool and Primary Scale of Intelligence – Revised (Wechsler, 1995) was used to assess verbal abstract reasoning and conceptualization abilities. The test comprises three parts: In the first part, the child sees a stimulus picture and is asked to select a compatible picture from an array of four pictures (six items); in the second part, the child completes a sentence with an appropriate word (six items); and in the third part the child describes how two things are alike (eight items).

A 30-item shortened version of the Peabody Picture Vocabulary Test – Revised (PPVT-R; Dunn & Dunn, 1981) was used to assess the child's single-word receptive vocabulary. The items were selected on the basis of data drawn from another Finnish study, the Jyväskylä Longitudinal Study of Dyslexia (see Lyytinen et al., 2004; Lyytinen, Erskine, Tolvanen, Torppa, Poikkeus, & Lyytinen, 2006) where the full-scale version of the PPVT-R was administered to the control group.

The Korpilahti Auditory Sentence Comprehension test (SC; Korpilahti 1996) was used as a test for receptive grammar. The test assesses the ability to process semantic and syntactic information in sentences. The test comprises 30 sentences that increase in complexity and make increasing demands on verbal reasoning and auditory short-term memory. After each sentence the child is presented with three pictures and asked to choose the one that goes best with the sentence.

The Verbal Fluency, Semantic categories test (VFS; NEPSY-II; Korkman, Kirk, & Kemp, 2008) assesses verbal productivity and vocabulary. The child is asked to generate as many words as possible within specific semantic categories (animals, foods) in 60 s.

Working memory measures at 5;3. The working memory measures were selected to cover the relevant subsystems of Baddeley's (2003; 2012) model of working memory, following the conceptualizations of Archibald & Gathercole (2006) and Petruccelli, Bavin, & Bretherton (2012).

The Digit Span subtest of the Wechsler Intelligence Scale for Children – Third Edition (WISC-III; Wechsler, 1999) comprises two parts: In the first part, the child repeats a dictated series of digits verbatim (forward part), and in the second part the child repeats the series backwards (backward part). The series begin with two digits and increases in length with two trials at each length. As the forward part is regarded as tapping the phonological loop and the backward part as tapping both the phonological loop and the central executive (e.g., Vance, 2008), the two parts were treated as separate measures in the analyses.

Nonword repetition (NWR) ability was assessed with the Repetition of Nonsense Words test (NEPSY; Korkman, Kirk, & Kemp, 1997). In this test, the child imitates 16 nonwords that increase in length from one ("nas") to six ("skrikoflunaflistrop") syllables. The nonwords conform to the phonotactic rules of Finnish but are low in word likeness and phonotactic frequency. The test is regarded as tapping the phonological loop along with other language-related processes such as speech perception, phonological encoding and assembly, and articulation (Coady & Evans, 2008).

In the Sentence Repetition task (SR; NEPSY-II; Korkman et al., 2008), the child is read 17 sentences that increase in complexity and length, and asked to recall each sentence verbatim immediately after it is presented. The task requires the integration of information from phonological short-term memory with long-term linguistic knowledge, and thus is regarded as being a measure of the episodic buffer, which is responsible for storing chunks of such integrated information (Baddeley, 2000; Boyle, Lindell, & Kidd, 2013).

Measures in the first grade. The children's language and communication difficulties were assessed with the Finnish version of the Children's Communication Checklist – Second Edition (CCC-2; Bishop, 2003; Norbury, Nash, Baird, & Bishop, 2004). The CCC-2 is a parent questionnaire used to screen for general language impairments and pragmatic language impairment in children aged 4 to 16 years. The questionnaire includes four subscales that measure language abilities (speech, syntax, semantics, and coherence) and four areas of pragmatics (inappropriate initiations, stereotyped language, use of context, and nonverbal communication). The two additional subscales (social relations and interests) were omitted in this study. Each scale comprises 5 questions on difficulties, and 2 questions on strengths (reversed scale). Parents rate the frequency of their child's language and communication behaviors on a four-point scale (0 = less than once a week, 1 = at least once a week, not every day, 2 = once or twice a day, 3 = several times a day/always). The Cronbach's αs for the separate subscales have been reported to be above .66 (Bishop, 2003). The αs in the current sample ranged between .57-.87 for the separate subscales and the αs for the combined language scales and combined pragmatics scales were .91 and .92 respectively.

Data analyses

The development of early communication skills was analyzed using a type of second-order multivariate Latent Growth Curve modeling called the factor-of-curves model (Duncan, Duncan, & Strycker, 2006, pp. 68-70; McArdle, 1988). Multivariate LGM is used to determine if development on one behavior covaries with development in other behaviors and it provides a "more dynamic view of the correlates of change, as development in one variable can be associated with development in another variable" (Duncan et al., 2006, p.63). In the factor-of-curves model it is examined whether a second-order factor adequately describes the covariances among lower order developmental functions (Duncan et al., 2006, p.68).

The analyses were performed using the Mplus statistical package (version 7; Muthén & Muthén, 1998-2010). The estimation method was the robust MLR which corresponds to the full-information maximum likelihood (FIML). In FIML there does not need to be the same number of items, observations, or variables for every individual as the log-likelihoods are written for each individual based on the individual's observed data (e.g. see Enders, 2010, pp. 88-92; Graham & Coffman, 2012, p. 282). The use of FIML over other methods such as listwise deletion is recommended as FIML preserves key relationships among variables and better estimates the variability in the data yielding more valid results (see, Jeličić, Phelps, & Lerner, 2009). Thus, despite having different amount of data at different age stages, all available data between the ages 6 to 18 months was used (n = 203-322 at different ages, n = 427 in total) as it leads to improved accuracy of parameter estimates (Enders, 2010, p.92). The coverage of the elements in the covariance matrix is presented in Table SM4.

The goodness-of-fit of the estimated LGC models was evaluated using several fit indexes (χ^2 test, the Comparative Fit Index, the Tucker-Lewin Index, the Root Mean Square Error of Approximation, and Standardized Root Mean Square Error of Approximation; Hu & Bentler, 1999; Muthén & Muthén, 1998-2010). Specifications to the model were done based on the model modification indices and theoretical considerations. Modification indices above 4 were taken into account and each of them was considered from a theoretical standpoint. Only those indices that were deemed appropriate both statistically and theoretically were added to the model.

All analyses were conducted with raw data. As the follow-up subsamples were only partially overlapping, the regression analyses were conducted separately for each follow-up.

Results

A latent growth curve (LGC) model for early communication development

All correlations between the three ITC composites at different age stages are shown in Table SM3. All the successive measurements within the ITC composites correlated significantly with each other, as was expected due to the sequential nature of the data, and thus, were allowed to correlate with each other within the composites in the LGC model.

A LGC model for each of the three ITC composites (social, speech, symbolic) was estimated simultaneously. Based on visual inspection of the individual growth curves and the model modification indices, nonlinear growth was estimated. In the model specifications all loadings on first order intercepts were fixed (at 1), while in the loadings of the first-order slopes, the first and last time points were fixed (at 0, and at 4), and age 9, 12, and 15 month loadings were estimated freely (*1, *2, *3). The modeling of unspecified trajectories using a two-factor model (only intercept and slope instead of a specified model) was chosen as the unspecified model might be able to provide better model fit and is somewhat easier to interpret. That is, the fitting of a quadratic and a cubic slope factor (i.e., a specified model), would lead to 9 and 12 first-order factors, respectively, which would lead to an unnecessarily complex model that would be more difficult to interpret and might lead to convergence problems. In addition, it has been suggested that unless there are solid theoretical justifications for another model, using unspecified model is recommended (see a simulation study by Welch, 2007). In this type of modeling, instead of a predefined shape of growth (i.e., adding a quadratic or cubic factor), the data is allowed to determine the shape of growth (Duncan et al., 2006, pp.31-35).

The correlations between the first-order level factors and between the first-order growth factors were significant (r =.57-.81, p <.001 between the social, speech, and symbolic level factors, and r = .48 - .67; p <.001 - .010 between the social, speech, and symbolic growth factors). Thus, a second-order factor structure (common level and common slope) was added to the model to describe these relationships between the composite-specific first-order factors (i.e., explain the covariances among the first-order factors; Duncan et al., 2006, pp. 68-69). The symbolic composite was used as the reference scaling for the second order structure (fixed at 1; Duncan et al., 2006, p. 69; McArdle, 1988) and the other factor loadings were estimated freely.

The residual correlations were strong between the different measures at the same time point (i.e., social, speech, and symbolic at age 9 month, age 12 months, and age 15 months) indicating that there is some age-specificity in development at these ages that is not captured by the first- and second-order factors. Thus, specific age factors were added to explain this between-individual variation that is specific to the time points measured and not related to development over the measured time period. These specific factors were not allowed to correlate with each other or with the first- and second-order factors. The model fitted the data well: $\chi^2(73) = 87.405$, p =.120 , CFI =.991 , TLI =.987, RMSEA = .021 and SRMR =.083.

Figure 2 depicts the LGC model and reports the standardized estimates. These estimates should be interpreted to depict effect sizes. In line with the observed means across the 6 to 18 month period (reported in Table 2), the LGC model showed growth throughout the measured time period in all three ITC composites, which was indicated by increases in the model produced mean values over time. The correspondence between the observed and the model estimated mean values was good. The first-order loadings on growth factors represent the individual differences present at a certain time point. Thus, a higher standardized loading for example at age 15 months compared to the loading at age 18 months in social and speech composites (see Figure 2) indicate that the largest individual differences are present at this age.

Early communication development and later language and communication skills

The model fit indices for the longitudinal models between the early LGC model and the follow-up measurements at ages 2, 3, 4;7, 5;3 and first grade are summarized in Table SM5.

The model modification indices suggested several skill- and age-specific pathways from the level of the speech composite, from the growth factor of the social and speech composites, and from the age-specific factor at 15 months of age. More specifically, the suggested pathways included: from the growth factor of the social composite to MCDI inflections at 24 months (p = .040), FTF expressive language at 55 months (p = .217), and the memory factor at 63 months (p = .274); from the level of the speech composite to MCDI inflections at 24 months (p = .136), the memory factor at 63 months (p = .019), and the first-grade language (p = .034) and communication (p = .024) factors; and from the growth factor of the speech composite to expressive vocabulary at 36 months (p = .050), FTF expressive language at 55 months ($p = .063$), the memory factor at 63 months ($p = .002$), and the first-grade language ($p = .007$) and communication ($p = .023$) factors. Age-specific paths were suggested from the specific age factor at 15 months to MCDI vocabulary ($p = .012$) and inflections ($p = .024$) at 24 months, the memory factor at 63 months ($p = .090$), and the first-grade language ($p = .007$) and communication ($p = .011$) factors. None of these paths were significant at the .001 level, and thus no specific paths were added to the regression models.

References

Archibald, L.M.D. & Gathercole, S.E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders, 41*, 675-693. doi: 10.1080/13682820500442602

Baddeley, A. (2000). The episodic buffer: a new component of working memory? Trends in Cognitive Sciences, 4, 417-423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36,* 189-208. doi: 10.1016/S0021-9924(03)00019-4

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1-29. doi: 10.1146/annurev-psych-120710-100422

Bishop, D. V. M. (2003). *The Children's Communication Checklist* (Vol 2). London: Harcourt Assessment.

Boyle, W., Lindell, A.K., & Kidd, E. Investigating the role of verbal working memory in young children's sentence comprehension. Language Learning, 63, 211-242. doi: 10.1111/lang.12003

Callister, L.C., Lauri, S., & Vehviläinen-Julkunen, K. (2000). A description of birth in Finland. *The American Journal of Maternal/Child Nursing, 25,* 146-150.

Coady, J.A. & Evans, J.L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*, 1-40. doi: 10.1080/13682820601116485

Conti-Ramsden, G. & Durkin, K. (2012). Language development and assessment in the preschool period. *Neuropsychology Review, 22,* 384-401. doi: 10.1007/s11065-012-9208-z

Duncan, T. E., Duncan, S. C., & Strycker L. A. (2006). An introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications. (2nd ed.) Mahwah, NJ: Lawrence Erlbaum Associates.

Dunn, L.M & Dunn, L.M. (1981). *Peabody Picture Vocabulary Test – Revised*. Cirle Pines. MN: American Guidance Service.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Publications.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. & Pethick, S. (1994). *Variability in early communicative development*. *Monographs of the Society for Research in Child De6elopment*, 59, 5 (Serial No. 242).

Graham, J. W. & Coffman, D. L. (2012). Structural equation modeling with missing data. In R. H. Hoyle (Ed.), Handbook of Structural Equation Modeling (pp. 277-295). New York: Guilford Press.

Henrichs, J., Rescorla, L., Schenk, J.J., Schimdt, H.G., Jaddoe, V.W.V., Hofman, A., … Tiemeierc, H. (2011). Examining continuity of early expressive vocabulary development: The Generation R study. Journal of Speech, Language, and Hearing Research, 54, 854-869. doi: 10.1044/1092-4388(2010/09-0255)

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi: 10.1080/10705519909540118

Jeličić, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. Developmental Psychology, 45, 1195-1199. doi: 10.1037/a0015665

Kadesjö, B., Janols, L.-O., Korkman, M., Mickelsson, K., Strand, G., Trillingsgaard, A., & Gillberg, C. (2004). The FTF (Five to Fifteen): The development of a parent questionnaire for the assessment of ADHD and comorbid conditions. *European Child and Adolescent Psychiatry, 13*

(Suppl. 3), 3–13. doi:10.1007/s00787-004-3002-2

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test.* Philadelphia: Lea & Febiger

Kekkonen, M. (2014). Perheiden lastenhoitojärjestelyt ja tyytyväisyys päivähoitopalveluihin. [The day care arrangements of families and their contentment with the services]. In J. Lammi-Taskula & S. Karvonen (eds). Lapsiperheiden hyvinvointi 2014. [The well-being of families with children in 2014]. National Institute of Health and Welfare. Tampere: Suomen Yliopistopaino Oy.

Korkman, M., Jaakkola, M., Ahlroth, A., Pesonen, A.-E., & Turunen, M.-M. (2004). Screening of developmental disorders in five-year-olds using the FTF (Five to Fifteen) questionnaire: A validation study. European Child and Adolescent Psychiatry, 13(Suppl. 3), 31–38. doi:10.1007/s00787-004-3005-z

Korkman, M., Kirk, U., & Kemp, S.L. (1997). *Lasten neuropsykologinen tutkimus (NEPSY: A Developmental Neuropsychological Assessment.*). Helsinki, Finland: Psykologien kustannus Oy.

Korkman, M., Kirk, U., & Kemp, S. L. (2008). *NEPSY-II: Lasten neuropsykologinen tutkimus (NEPSY-II: A Developmental Neuropsychological Assessment, Second edition).* Helsinki, Finland: Psykologien kustannus Oy.

Korpilahti, P. (1996). *Lausetesti; Test for the language comprehension of sentences.* Helsinki: University Press.

Laakso, M.-L., Poikkeus, A.-M., & Eklund, K. (2011). *Lapsen esikielellisen kommunikaation ja kielen ensikartoitus* [The Infant-Toddler Checklist of Children's Prelinguistic Communication and Language]. Jyväskylä, Finland: Niilo Mäki Instituutti.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Kulju, P., Laakso, M-L., … Viholainen, H. (2004). Early Development of children at familial risk for dyslexia – Follow-up from birth to school age. *Dyslexia, 10*, 146-178. doi: 10.1002/dys.274

Lyytinen, H. Erskine, J., Tolvanen, A., Torppa, M., Poikkeus, A-M.,, & Lyytinen, P. (2006).

Trajectories of reading development: a follow-up from birth to school age of children with and without risk for dyslexia. *Merrill-Palmer Quarterly, 52*, 514-546. doi: 10.1353/mpq.2006.0031

Lyytinen, P. (1999). *Varhaisen kommunikaation ja kielen kehityksen arviointimenetelmä (Finnish Manual for Communicative Development Inventories)*. Jyväskylän yliopiston lapsitutkimuskeskus ja Niilo Mäki Instituutti. Jyväskylä: Yliopistopaino

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In R.B. Cattell & J. Nesselroade (Eds.), Handbook of multivariate experimental psychology (2nd ed., pp. 561-614). New York: Plenum.

Muthén, L. K. and Muthén, B. O. (1998-2010). *Mplus User's Guide.* Sixth Edition. Los Angeles, CA: Muthén & Muthén

Norbury, C. F., Nash, M., Baird, G., & Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist – 2. *International Journal of Language and Communication Disorders, 39,* 345-364. doi: 10.1080/13682820410001654883

Petruccelli, N., Bavin, E.L., & Bretherton, L. (2012). Children with specific language impairment and resolved late talkers: Working memory profiles at 5 years. *Journal of Speech, Language, and Hearing Research, 55,* 1690-1703. doi: 10.1044/1092-4388(2012/11-0288)

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., … Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Finding from Early Language in Victoria Study. *Pediatrics, 126*, e1530-e1537. doi: 10.1542/peds.2010-0254

Salmi, M. & Lammi-Taskula, J. (2014). Lapsiperheiden vanhemmat työelämässä. [Parents in the working life]. In J. Lammi-Taskula & S. Karvonen (eds). Lapsiperheiden hyvinvointi 2014. [The well-being of families with children in 2014]. National Institute of Health and Welfare, Tampere: Suomen Yliopistopaino Oy.

Statistics Finland (2013). Retrieved from Statistics Finland website: http://www.tilastokeskus.fi

Vance, M. (2008). Short-term memory in children with developmental language disorder. In C.F. Norbury, J. B. Tomblin, & D. Bishop (Eds.), *Understanding Developmental Language Disorders: from theory to practice*, (pp. 23-38). East Sussex: Psychology Press.

Wechsler, D. (1995). *WPPSI-R Käsikirja (Wechsler Preschool and Primary Scale of Intelligence: Revised. Manual for the Finnish Adaptation*). Helsinki: Psykologien kustannus Oy.

Wechsler, D. (1999). *WISC-III Käsikirja (Manual for the Wechsler Intelligence Scale for Children– Third Edition. Manual for the Finnish Adaptation*). Helsinki: Psykologien kustannus Oy.

Welch, G. W. (2007). Model Fit and Interpretation of Non-Linear Latent Growth Curve Models. (Doctoral dissertation). Retrieved from http://d-scholarship.pitt.edu/8769/1/WelchGreg Dissertation_ 080707.pdf

Wetherby, A. M., & Prizant, B. M. (2002). *Communication and Symbolic Behavior Scales: Developmental Profile* (1st normed ed.). Baltimore, MD: Brookes.

Table SM1

Comparisons of the ITC scores at ages 12, 15, and 18 months of age between the children participating and not participating in the follow-ups

	ITC at age 12 months				ITC at age 15 months				ITC at age 18 months			
	n (data/ no data)	Soc	Spe	Sym	n (data/ no data)	Soc	Spe	Sym	n (data/ no data)	Soc	Spe	Sym
			F				F				F	
2 years	99/223	*n.s.*	*n.s.*	*n.s.*	101/204	*n.s.*	17.6*	*n.s.*	98/181	*n.s.*	*n.s.*	*n.s.*
3 years	109/213	*n.s.*	*n.s.*	*n.s.*	108/197	12.3*	15.3*	*n.s.*	105/174	*n.s.*	12.9*	*n.s.*
4;7	206/116	*n.s.*	*n.s.*	*n.s.*	208/97	*n.s.*	*n.s.*	*n.s.*	191/88	*n.s.*	*n.s.*	*n.s.*
5;3	93/229	*n.s.*	*n.s.*	*n.s.*	97/208	*n.s.*	*n.s.*	*n.s.*	88/191	*n.s.*	*n.s.*	*n.s.*
1st grade	190/132	*n.s.*	*n.s.*	*n.s.*	193/112	*n.s.*	*n.s.*	*n.s.*	171/108	*n.s.*	*n.s.*	*n.s.*

Note. ITC = Infant-Toddler Checklist; Soc = Social composite; Spe = Speech composite; Sym = Symbolic composite.
*p < .001

Table SM2

Demographic Information for the Original Sample, the Sample in the Early LGC Model, and the Follow-Up Subsamples

	Original sample	Early LGC model	Follow-up 2 years	Follow-up 3 years	Follow up 4;7	Follow-up 5;3	Follow-up first grade
N	508	427	104	111	253	102	236
Males/females, %	50.2/49.8	50.8/49.2	54.8/45.2	56.8/43.2	51.4/48.6	52.0/48.0	51.7/48.3
Preterm birth (< 36 wk), *n* (%)	14 (2.9)	14 (3.4)	3 (2.9)	5 (4.5)	6 (2.4)	3 (2.9)	6 (2.6)
Birth weight, *M* (*SD*), kg	3.5 (0.6)	3.5 (0.6)	3.6 (0.6)	3.6 (0.6)	3.5 (0.5)	3.5 (0.6)	3.6 (0.5)
Birth order, first born, *n* (%)	267 (56.6)	216 (54.8)	49 (50.0)	51 (48.6)	134 (57.8)	50 (51.0)	115 (52.5)
Single parents[a], *n* (%)	19 (3.9)	15 (3.6)	2 (2)	3 (2.7)	7 (2.9)	0 (0)	5 (2.1)
Parent's education[a], *M* (*SD*)							
Mother	3.9 (2.0)	3.8 (1.9)	4.1 (1.8)	4.1 (1.9)	3.9 (1.9)	4.1 (2.0)	4.1 (1.9) [b]
Father	3.6 (2.0)	3.6 (1.8)	3.5 (1.8)	3.8 (1.7)	3.6 (1.9)	3.8 (1.8)	3.6 (1.8)
Parent's age[a] *M* (*SD*), y							
Mother	29.8 (5.4)	29.7 (5.4)	30.4 (5.1)	30.6 (5.2)	30.0 (5.3)	30.9 (5.2)	30.4 (5.2)
Father	32.1 (6.3)	32.0 (6.4)	32.5 (6.0)	32.7 (5.7)	32.3 (6.1)	32.5 (5.8)	32.4 (6.3)

Note. Coverage of the demographic data varied between 472-494 in the original sample, 394-416 in the early LGC model sample, 98-104 at 24 months, 105-111 at 36 months, 232-249 at 55 months, 98-102 at 63 months, and 219-235 in the first grade. The percentages are calculated from the available data. LGC = latent growth curve.

[a] At time of initial recruitment.

[b] Significantly different compared to the original sample (p < .01)

Table SM3.

Correlations between the ITC composites at different ages.

		Soc06	Soc09	Soc12	Soc15	Soc18	Spe06	Spe09	Spe12	Spe15	Spe18	Sym06	Sym09	Sym12	Sym15
Soc09	*r*	.471***													
	n	202													
Soc12	*r*	.386***	.560***												
	n	142	135												
Soc15	*r*	.302***	.477***	.619***											
	n	151	142	277											
Soc18	*r*	.278***	.404***	.594***	.721***										
	n	149	140	234	246										
Spe06	*r*	.304***	.236***	.182*	.070	.115									
	n	229	202	142	151	149									
Spe09	*r*	.225***	.281***	.162	.220**	.148	.340***								
	n	202	203	135	142	140	202								
Spe12	*r*	.220**	.406***	.439***	.342***	.403***	.257**	.484***							
	n	142	135	322	277	234	142	135							
Spe15	*r*	.099	.216**	.345***	.431***	.385***	.180*	.425***	.529***						
	n	151	142	276	304	246	151	142	276						
Spe18	*r*	.098	.203*	.330***	.380***	.384***	.132	.407***	.468***	.786***					
	n	149	140	234	246	279	149	140	234	246					
Sym06	*r*	.509***	.368***	.188*	.109	.099	.369***	.256***	.168*	.072	.086				
	n	229	202	142	151	149	229	202	142	151	149				
Sym09	*r*	.352***	.493***	.419***	.322***	.250**	.278***	.389***	.385***	.274***	.289***	.402***			
	n	202	203	135	142	140	202	203	135	142	140	202			
Sym12	*r*	.181*	.309***	.501***	.445***	.334***	.250**	.251**	.443***	.388***	.339***	.189*	.496***		
	n	142	135	322	277	234	142	135	322	276	234	142	135		
Sym15	*r*	.156	.287***	.466***	.530***	.386***	.149	.234**	.339***	.477***	.406***	.067	.457***	.666***	
	n	151	142	277	305	246	151	142	277	304	246	151	142	277	
Sym18	*r*	.157	.366***	.482***	.553***	.487***	.138	.284***	.363***	.429***	.465***	-.009	.331***	.501***	.683***
	n	149	140	234	246	279	149	140	234	246	279	149	140	234	246

Note. Soc = Social composite; Spe = Speech composite; Sym = Symbolic composite. Numbers after soc, spe, and sym represent age in months.
*<.05. **<.01. ***.001.

Table SM4

The coverage of the elements in the covariance matrix in the LGC model and follow-up assessments (% of the total sample of 427).

Age (months)	**Early communication**	Age in months					Age (years; months)				
		6	9	12	15	18	2	3	4;7	5;3	7;9
6	ITC	53.6									
9	ITC	47.3	47.5								
12	ITC	33.3	31.6	75.4							
15	ITC	35.4	33.3	64.9	71.4						
18	ITC	34.9	32.8	54.8	57.6	65.3					
(years)	**Follow-up measures**										
2	MBCDI Vocabulary	8.7	8.2	23.2	23.7	23.0	24.4				
2	Inflections	8.4	8.2	23.0	23.4	22.7	24.1				
2	MSL	8.2	8.0	22.2	22.7	22.0	23.4				
3	Boston naming	10.8	10.8	25.5	25.3	24.6		26.2			
3	PPVT	10.5	10.5	25.3	25.1	24.4		26.0			
4;7	FTF	27.7	25.3	43.8	44.3	40.6			63.0		
5;3	SI	11.0	10.8	20.8	21.8	19.7				23.0	
5;3	PPVT-R	11.0	10.8	21.3	22.2	20.1				23.4	
5;3	SC	11.5	11.2	21.8	22.7	20.6				23.9	
5;3	VF	11.0	10.8	20.6	21.5	19.4				22.7	
5;3	DSf	11.5	11.2	21.5	22.5	20.4				23.7	
	DSb	11.0	10.8	20.8	21.8	19.7				23.2	
5;3	NWR	11.0	10.8	20.8	21.8	19.7				23.0	
5;3	SR	11.2	11.0	21.3	22.2	20.1				23.4	
7:9	CCC-II	27.8	25.6	40.9	41.6	36.9					58.8

Note. ITC = Infant-Toddler Checklist; MBCDI = MacArthur-Bates Communicative Development Inventories; MSL = Maximum Sentence Length; PPVT = Peabody Picture Vocabulary Test; FTF = Five to Fifteen; SI = Similarities; SC = Sentence Comprehension; VF = Verbal Fluency; DS = Digit span; NWR = Nonword Repetition; SR = Sentence Repetition; CCC-II = Children's Communication Checklist –Second Edition.

Table SM5

The Model Fit Indices for the LGC Model and the Regression Models of the Follow-Up Measurements.

Age	LGC model	χ^2	df	p	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR
12 to 18 months	Early communication development	87.405	73	0.1198	0.991	0.987	0.021	0.000 0.037	0.083
	Outcome variables								
2 years	MCDI: vocabulary, inflections, MSL	153.548	112	0.0056	0.979	0.971	0.029	0.017 0.040	0.091
3 years	Boston, PPVT	119.435	99	0.0793	0.988	0.984	0.022	0.000 0.035	0.083
4;7	FTF expressive, receptive, communication	136.459	112	0.0579	0.987	0.982	0.022	0.000 0.033	0.078
5;3	Psychometric tests: Language, memory	258.160	207	0.0090	0.973	0.967	0.024	0.018 0.033	0.089
1st grade	CCC-II: Language, communication	236.476	206	0.0714	0.991	0.988	0.018	0.000 0.028	0.071

Note. A nonsignificant chi-square test (p >.05), CFI and TLI values at or above .95, RMSEA below .06, and SRMR below .08 serve as guidelines for determining good model fit (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004). LGC = latent growth curve; CFI = comparative fit index; TLI = Tucker-Lewin index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square error of approximation; MCDI = MacArthur-Bates Communicative Development Inventories; MSL = Maximum Sentence Length; PPVT = Peabody Picture Vocabulary Test; FTF = Five to Fifteen; CCC-II = Children's communication Checklist –Second Edition.