

Supplemental Appendix S1. Detailed description of the participants, procedures, and measures.

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 and 6 years. The services are focused on health promotion, risk assessment, and disease prevention. Visits are made to the clinic 10–15 times during the first 2 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 [CSBS-DP], 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 3 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 years;months (parent report), 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 ages 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 also enables comparison of at-risk and typically developing children. At ages 4;7 and in first grade, all the originally participating families were contacted. The group differences in the early ITC scores are summarized in Supplemental Table SM1.

Demographic information by subsamples is presented in Supplemental 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%) of 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 (which is 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 7-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 3 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 daycare centers or in family daycare, the former of which is more commonly used (84% vs. 16%; Kekkonen, 2014). Rates of daycare attendance vary according to the child's age. Around 30% of 1-year-olds, 50% of 2-year-olds, 70% of 3-year-olds, 75% of 4-year-olds, and 80% of 5-year-olds are in daycare. Family daycare is more common in the youngest age groups. Children have a right to attend preschool education the year before their compulsory education starts (i.e., the year they turn 6). Preschool education is provided in daycare centers and primary schools. The majority (98%) of children attend preschool education (Statistics Finland, 2013). Compulsory schooling starts in the year of the child's 7th 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 ranges = 100–102 and 230–234) and those who did not (n ranges = 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 3-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 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 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 in 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–Revised (PPVT-R; Dunn & Dunn, 1981). The PPVT-R 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 attention-deficit/hyperactivity disorder (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 3-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 α s 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 Durkin (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 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) of the 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 (e.g., 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 and Gathercole (2006) and Petruccelli, Bavin, and 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) of the 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-II; Bishop, 2003; Norbury, Nash, Baird, & Bishop, 2004). The CCC-II 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 five questions on difficulties and two questions on strengths (reversed scale). Parents rate the frequency of their child’s language and communication behaviors on a 4-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 α s for the separate subscales have been reported to be above .66 (Bishop, 2003). The α s in the current sample ranged between .57–.87 for the separate subscales and the α s 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 (LGM) 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 multiple linear regression (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 because 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 Supplemental Table SM4.

The goodness-of-fit of the estimated latent growth curve (LGC) models was evaluated using several fit indexes (χ^2 test, the comparative fit index [CFI], the Tucker–Lewis index [TLI], the root mean square error of approximation [RMSEA], and standardized root mean square [SRMR] 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 LGC Model for Early Communication Development

All correlations between the three ITC composites at different age stages are shown in Supplemental 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 of the main article 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 of the main article), 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 of the main article) 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 Supplemental 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—Second Edition*. London, United Kingdom: Harcourt Assessment.
- Boyle, W., Lindell, A. K., & Kidd, E. (2013). 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: Erlbaum.
- Dunn, L. M., & Dunn, D. M. (1981). *Peabody Picture Vocabulary Test—Revised*. Circle Pines, MN: AGS.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S., . . . Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1–185.
- 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, NY: Guilford.
- Henrichs, J., Rescorla, L., Schenk, J. J., Schimdt, H. G., Jaddoe, V. W. V., Hofman, A., . . . Tiemeier, 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, PA: Lea & Febiger.
- Kekkonen, M. (2014). Perheiden lastenhoitojärjestelyt ja tyytyväisyys päivähoitopalveluihin [The daycare 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, Finland: 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, Finland: Gaudeamus 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, Finland: Yliopistopaino.
- McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In R. B. Cattell & J. Nesselrode (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561–614). New York, NY: Plenum.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user’s guide* (6th ed.). 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
- Petrucelli, 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, Finland: Suomen Yliopistopaino Oy.
- Statistics Finland. (2013). [Homepage]. Retrieved from <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, United Kingdom: Psychology Press.
- Wechsler, D. (1995). *WPPSI-R Käsikirja* [Wechsler Preschool and Primary Scale of Intelligence– Revised: Manual for the Finnish adaptation]. Helsinki, Finland: Psykologien kustannus Oy.
- Wechsler, D. (1999). *WISC-III Käsikirja* [Wechsler Intelligence Scale for Children–Third Edition: Manual for the Finnish adaptation]. Helsinki, Finland: Psykologien kustannus Oy.
- Welch, G. W. (2007). *Model fit and interpretation of non-linear latent growth curve models* (Doctoral dissertation). Retrieved from the University of Pittsburgh repository.
- Wetherby, A. M., & Prizant, B. M. (2002). *Communication and Symbolic Behavior Scales: Developmental Profile–First Normed Edition*. Baltimore, MD: Brookes.