JYVÄSKYLÄ STUDIES IN HUMANITIES

72

Lea Nieminen

A Complex Case

A Morphosyntactic Approach to Complexity in Early Child Language

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN HUMANITIES 72

Lea Nieminen

A Complex Case

A Morphosyntactic Approach to Complexity in Early Child Language

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Villa Ranan Paulaharjun salissa
kesäkuun 21. päivänä 2007 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities of the University of Jyväskylä,
in the Building Villa Rana, Paulaharju Hall, on June 21, 2007 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2007

A Complex Case

A Morphosyntactic Approach to Complexity in Early Child Language

JYVÄSKYLÄ STUDIES IN HUMANITIES 72

Lea Nieminen

A Complex Case

A Morphosyntactic Approach to Complexity in Early Child Language

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2007

Editors
Maisa Martin
Department of Languages, University of Jyväskylä
Irene Ylönen, Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

URN:ISBN:9789513928582
ISBN 978-951-39-2858-2 (PDF)

ISBN 978-951-39-2840-7 (nid.)
ISSN 1459-4331

Jyväskylä University Printing House, Jyväskylä 2007

ABSTRACT

Nieminen, Lea
A complex case: a morphosyntactic approach to complexity in early child language
Jyväskylä: University of Jyväskylä, 2007, 296 p.
(Jyväskylä Studies in Humanities
ISSN, 1459-4331; 72)
ISBN 978-951-39-2858-2 (PDF), 978-951-39-2840-7 (nid.)
Diss.

This study investigates morphosyntactic complexity in children's utterances from a multidimensional perspective. Results from Mean Length of Utterance and Index of Productive Syntax were compared with the results of Utterance Analysis, a method developed for the purposes of this study, to discover what aspects of absolute complexity are reflected by each of the scales.

A second purpose of the study was to compare children with high genetic risk of dyslexia (N= 20) with their controls (N= 20) in order to determine whether morphosyntactic complexity in their utterances differed on a group level. The data were gathered in a play situation when the children were 30 months old. MLU and IPSyn analyses were carried out on the 80 longest utterances from each child. For the Utterance Analyses three subgroups showing contradictory results in MLU and IPSyn were selected.

The study revealed three major findings. The first, concerning morphosyntactic development in general, suggests that children usually start with morphological elaboration first in one and then in two components in an utterance. Only after that does the elaboration of components start to deepen vertically. The more components there are in an utterance the more similar is children's performance in elaboration.

The second finding, concerning methodology, suggests that different scales of measurement reveal different aspects of complexity. MLU focuses on linear length and IPSyn on the inventory of resources, but the relationship between these two methods uncovers developmental trends in morphosyntactic elaboration. Only Utterance Analysis can provide a complete picture of multidimensional structural complexity.

The third finding concerns the comparison between risk and control children. It seems that children with a high genetic risk of dyslexia do have the same morphosyntactic resources as their controls, but they use them differently. Risk children concentrate more on elaborating only a single component per utterance whereas control children spread elaboration to more components.

Keywords: structural complexity, absolute approach to complexity, first language acquisition, acquisition of morphosyntax, Index of Productive Syntax, Mean Length of Utterance

Author's address Lea Nieminen
Department of Languages
P.O. Box 35 (F)
FIN-40014 University of Jyväskylä
FINLAND

Supervisors Professor Matti Leiwo
Department of Languages
P.O. Box 35 (F)
FIN-40014 University of Jyväskylä
FINLAND

Professor Maisa Martin
Department of Languages
P.O. Box 35 (F)
FIN-40014 University of Jyväskylä
FINLAND

Reviewers Professor emeritus Ann Peters
Graduate Chair
Department of Linguistics
University of Hawai`i
1890 East West Road, Rm 569
Honolulu, HI 96822
USA

Ph.D Matti Miestamo
Helsinki Collegium for Advanced Studies
P.O. Box 4 (Fabianinkatu 24)
FI-00014 University of Helsinki
FINLAND

Opponent Professor emeritus Ann Peters
Graduate Chair
Department of Linguistics
University of Hawai`i
1890 East West Road, Rm 569
Honolulu, HI 96822
USA

FOREWORD

Now that the long process of writing a dissertation is finally coming to an end it is time to thank all those people who have made my winding journey through this process possible. I am very grateful to my supervisor, Professor Matti Leiwo, for giving me the opportunity to participate in such an interesting and ambitious research project as Early Language Development and Dyslexia, heading towards internationally significant goals. My other supervisor, Professor Maisa Martin, deserves especially warm thanks for being so supportive, encouraging and open-minded and ready for brainstorming during the final stages of the research.

I am very grateful to the Jyväskylä Longitudinal Study of Dyslexia project, headed by Professor Heikki Lyytinen, for making their unique data available. I would like to warmly thank all those who conducted the data collection and the families that were followed in this research project.

It has been an honour to have two such distinguished evaluators as Emeritus Professor Ann Peters and Dr Matti Miestamo reading my dissertation. I am especially grateful to Professor Peters for her encouraging comments on all aspects of the work. Similarly, Dr Miestamo deserves special thanks for his insightful suggestions concerning the theoretical issues. Thanks also to David Underwood for checking the language of the text.

It has been a privilege to share a room with a group of exceptionally intelligent and creative fellow researchers: I am deeply grateful to Ulla Richardson, Pirjo Kulju, Päivi Torvelainen, Minna Suni, Heidi Vaarala, Annekatrin Kaivapalu and Niina Lilja for being so understanding and encouraging over the years. Our long discussions at times of both despair and joy have meant a great deal. I also want to thank Professor Ritva Takkinen and Professor Hannele Dufva for the strength they have given me.

I would like to thank the Department of Languages for providing me with facilities and especially the Finnish language and Finnish sign language staff in the Fennicum building for creating the warm atmosphere essential for a long-term project. This work has been financially supported by the Ellen and Artturi Nyyssönen Foundation and the Central Finland Cultural Foundation.

Last but not least I want to thank my nearest and dearest ones. Without the strength my husband has given me through all these years I would never have achieved this goal. His endless support in times of trouble and sincere joy in moments of success have made it all worthwhile. I want to thank our children for being the inspiration of the whole study. After all, it was their steps through the early stages of language acquisition that made me interested in this subject. I therefore dedicate this work to my loving husband Jouko and our wonderful children Joonas, Milja and Jiri.

Vaajakoski May 21, 2007 Lea Nieminen

TABLES

FIGURES

APPENDICES

CONTENTS

1 INTRODUCTION

1.1 Complexity in studies on child language acquisition

This study is concerned with complexity and its variation in children's language. In this study structural complexity is investigated in relation to the morpho-syntactic structure of Finnish, high genetic risk of dyslexia, and the methods commonly used to evaluate complexity in children's speech. All the results are based on spontaneous speech data collected at the age of 30 months from 40 normally developing children acquiring Finnish as their first language. Half of the children involved in this study had a high genetic risk of dyslexia.

Complexity has been an important part of acquisition studies, even though it is not always explicitly discussed. The concept of language development itself seems to include the notion of complexity, since it is most often understood in terms of distinct areas of language growing into more complex, fine-grained and hierarchical systems. Studying language development – no matter whether it is normal or impaired – often means documenting growth towards complexity. Comparing children in terms of their linguistic development means more or less comparing their positions on a complexity scale. Thus, complexity is somehow a built-in feature of acquisition studies.

It is logical to suppose that children begin with easy and undifferentiated items and gradually move towards more complex ones when acquiring their first language. Such an approach sees language users' experiences of difficulty as major criteria of complexity, and these experiences are reflected in the order of acquisition. This view is very much open to question, one of the biggest questions being whether it is linguistic or cognitive complexity that affects the difficulty of acquisition and what the relationship between them is (cf. Bowerman 1996). Moreover, the relationship between the different processes involved in acquisition (perception, comprehension, production, etc.) is neglected. Studies on acquisition order give good descriptions of developmental course but do not actually explain the reasons for difficulties or the order itself and thus the comprehension of complexity itself is rather vague.

In acquisition studies, it is common to quantify complexity. Typically this is done by measuring the length of utterances in morpheme, word or clause units. The more units there are in an utterance the more complex it is taken to be. Thus the assessment of complexity is reduced to the quantification of certain units of language production without paying attention to their inner quality. The questions of whether length is a suitable indicator of complexity and what aspect(s) of complexity it probably represents are neither asked nor answered.

As we can see, the notion of complexity is usually passed by without going deeply into it, as if it were self-evident. There are few exceptions to this treatment of complexity in acquisition studies. One of them is an experimental study by Smith and van Kleeck (1986; see also Smith 1988), where they distinguish between surface complexity with reference to sentence length, and interpretive complexity with reference to requirements of interpreting empty categories in sentences. They support a so-called weak complexity hypothesis, which suggests that some complexity factors affect performance more than others, and that types of performance are different in the way complexity affects them.

In the special area of disfluency studies, complexity has been more explicitly the focus of interest, and the consideration of linguistic complexity has been twofold since both competence and performance have been taken into account. When linguistic abilities are considered, it can be concluded that children who stutter are able to produce utterances as complex and as grammatical as fluently speaking children do (Bernstein Ratner 1995: 34), and thus disfluency does not mean lack of linguistic competence. However, another point of view links structural complexity in utterances produced with the occurrence of disfluencies. In several studies it has been stated that disfluencies are most likely to appear in utterances which are structurally more complex than others (see Bernstein Ratner 1995, and Silverman & Bernstein Ratner 1997, for review).

In contrast with traditional acquisition studies, where measure of length has very often been taken to represent measure of complexity, complexity and length are separated from each other in several studies of stuttering, and the relationship between them has been seriously discussed (see Silverman & Bernstein Ratner 1997; Yaruss 1999). Bernstein Ratner and Sih (1987) found that syntactic complexity affected stuttering more than sentence length. In a later study by Logan and LaSalle (1999), the effect of syntactic complexity (measured by counting the number of clausal constituents) on the occurrence of disfluency clusters and disfluencies in general was significant. Logan and LaSalle also separated length (measured in syllables) from complexity, and found that utterances containing disfluency were longer than fluent ones, even though the independent contribution of structural complexity and utterance length to disfluency was not assessed. In addition to distinguishing length as just one individual factor, disfluency researchers have analysed syntactic complexity more thoroughly than has been done in traditional acquisition analysis. For example, Yaruss (1999) examined sentence structure, clause structure, and phrase structure, and their relationship to stuttering in children aged 40–66 months.

To summarize, in traditional studies on language development the notion of complexity has very often been taken for granted without proper identification or theoretical discussion of any kind. Only in some special areas of acquisition studies has the subject been given deeper consideration. For this reason complexity has been taken as the subject of this study, and the concept itself will be discussed and identified for the purpose in hand.

1.2 Special areas of the study

1.2.1 Morphosyntactic development

In this study the term *morphosyntax* refers to the simultaneous presence of morphological and syntactic units in a child's utterance and how these units together build the structural entirety of the utterance. The semantic and syntactic functions of morphological units are set aside and the focus is totally on the structure.

There are two major reasons why morphosyntactic development in children acquiring Finnish is at the core of this study. First, there are few studies focusing on this field in general, and second, the nature of Finnish morphosyntax provides an interesting field when complexity is in question.

Finnish studies of language acquisition in general have a rather short history and, predictably, there are significant gaps existing within the research area of morphosyntactic development in Finnish-speaking children. Toivainen's (1980) study on the acquisition of morphology was the first research in this field to be carried out in a systematic fashion. His so-called Oulu corpus consisted of spoken material from 25 children, and their morphological development was followed from ages 1;0–2;1 to 2;9–4;4. The purpose of the study was to investigate the order of acquisition of Finnish inflectional morphemes. In the field of psychology, Lyytinen (1978) conducted an experimental study (N = 140) of morphological development and learning assessed by means of a Finnish version of Jean Berko's Wug test. In addition to these large scale systematic studies, there have been minor studies on morphological acquisition, based on more or less nonsystematic diary data collected from one or two children (e.g. Räisänen 1975; Niemi & Niemi 1985; Laalo 1994, 1995, 1997, 1998). Although the order of acquisition of Finnish inflectional morphemes has been described in several ways in the literature, other sectors in morphological development have been given much less attention.

In the field of syntax and morphosyntax the gaps are even wider than in morphology. Bowerman's (1973) Finnish-American study was the first to investigate the emergence of syntactic structures in Finnish-speaking children. In her research the language production of two Finnish children temporarily living in the Boston area was analysed syntactically in terms of pivot grammar, transformational grammar, and case grammar. The analysis was restricted to Early and Late Stage I as defined by Brown (1973/1976), in other words, to the initial phase of multi-word utterances. Since Bowerman's study, Kangassalo (1995) has focussed on the development of interrogative expressions in Finnish-speaking children living in Sweden, Lieko (1992) has investigated the emergence of complex sentences, and Kauppinen (1996, 1998) has covered expressions including the conditional mood. In Kangassalo's study, there were 11 subjects aged 1;7–4;1, whereas Lieko's observations were based on the production of one child and so were Kauppinen's, except for occasional references to the productions of 15 children in the Oulu and Helsinki corpuses.

The list of studies focusing on the syntactic acquisition of Finnish reveals several lacunae in the research field. The first of these is the linguistic scope of the research. Between the stage covered by Bowerman's (1973) study on the initial phase of multi-word utterances as produced by two Finnish-speaking children and the stage covered by Lieko's (1992) case study on the production of complex sentences, there is a large unexplored area. This is unfortunate because during the intervening period large steps are taken in the acquisition of all aspects of language, and especially in morphology and syntax. The other studies mentioned above are specialized in one way or another, and a general picture of syntactic development remains incomplete. The second lacuna is due to the small number of subjects covered and their nature. Only Kangassalo (1995) used more than two subjects in her study and only one or two children were involved in the other studies. As for the nature of the subjects, both Kangassalo and Bowerman (1973) studied Finnish-speaking children living in a non-Finnish environment, and this may have had an effect on the results.

From the point of view of this study the most remarkable lacuna is the fact that morphosyntactic development is only touched upon but not focused on. Except for Kauppinen's (1998) research on formulaic utterances embodying the conditional mood, all the studies introduced above are either morphologically or syntactically biased and the interplay of these two major sectors in child language development is not emphasized. This is unfortunate, because the relationship between syntactic structures and inflectional morphemes is a central feature of Finnish structure.

Although it is not reasonable to think that one study could fill all the gaps existing within this research area, this study deals with all the issues discussed above. For one thing, using 40 subjects affords more opportunity to draw general conclusions than a study based on the production of only one or a few children. Secondly, the data was collected at an acquisition phase between the phases described in Bowerman's and Lieko's studies. Thirdly, the focus of this study is on the interplay of morphology and syntax with special attention to development in structural complexity. And finally, the fact that all the subjects are within one month of each other in age confirms that there will be evidence of variation at the age of 30 months.

The other reason why Finnish morphosyntax is an excellent field for studying structural complexity and its variation as part of language acquisition is the richness of Finnish morphology. There are 150–200 derivational affixes and approximately 40 inflectional suffixes in Finnish, including infinitives and participles (Karlsson 1983: 231). It has been estimated that each nominal (that is, noun, pronoun, adjective, and numeral) has 2 200 and each verb 12 000 different inflectional forms. If derivational forms were included the number of word forms would be many times greater. (Ibid. 356–57.) Apart from the considerable number of suffixes, morphophonological variation is a very common phenomenon in Finnish, and it involves both stems and suffixes. Finnish verbs and nominal words may have either one or two stems, within which the morphophonological variation takes place. (ISK 2004: 87–88) The existence of different vowel and consonant stems is of historical origin, a result of historical changes in the language (Hakulinen 1979: 80–84).

In Suomen kielen perussanakirja (1990–1994; a Finnish general dictionary), there are 49 different types of nominal inflection (51 if the two compound types are included) and 27 types of verb inflection (ibid. XIV–XVIII). The different types are identified according to stem type, stem variation and suffix variation (excluding effects of vowel harmony and consonant gradation). Exceptional inflection, like that in the negative auxiliary *ei* 'no, not' and in personal pronouns, is not included in these type groups. (ISK 2004: 94.)

Another property characteristic of Finnish is agreement, which can be divided in two types of phenomenon. First, there is purely morphological agreement, which indicates that the modifier and the head of a noun phrase agree in number and case. Second, there is semantic agreement, when the same thing is referred to more than once in a clause or a phrase. A typical occurrence of semantic agreement in Finnish is personal agreement between a subject and a predicate where a person is redundantly referred to in a clause. The inflectional categories involved in agreement in Finnish are case, number and person. (ISK 2004: 1217.)

To sum up, gaps in former research on the morphosyntactic development of Finnish-speaking children and the richness of Finnish morphology intertwined with syntax are the reasons why morphosyntax is at the core of this study. They provide a suitable basis for investigating how structural complexity manifests itself and varies in children's language production. When the starting point for the study is spontaneous speech and complete utterances produced by children, co-operation between syntactic structures and morphological forms in various developmental phases is most likely to be present in the data. Rich inflectional morphology will in all likelihood make the co-operation intricate. These factors make Finnish a morphosyntactically interesting language in a research field which is dominated by Indo-European languages.

1.2.2 Developmental dyslexia

The second special area in this study is developmental dyslexia. It has already for several decades been identified with reference to linguistic abilities. According to Catts, dyslexia is a "developmental language disorder that involves a specific deficit(s) in the processing of phonological information"

(Catts 1989a: 58). The core reason for dyslexia is present all through the sufferer's life and manifests itself in different ways (Catts 1991).

The most popular explanation for reading difficulties has been poor phonological processing. According to a strong version of the phonological deficit hypothesis, children in dyslexic families are likely to have difficulties with phonological processing which expose them to literacy problems, whereas the development of other language skills is normal (Gallagher, Frith & Snowling 2000). However, speech articulation (see e.g. Catts 1989b) and phonological processing (see e.g. Richardson 1998) have not been the only symptoms associated with dyslexia. There is much evidence that supports a weaker version of the hypothesis which argues for a broader delay in language development among dyslexic children (Gallagher et al. 2000). In her retrospective study of dyslexic and normal readers, Scarborough (1990b) pointed out that both phonological and syntactic production skills at the age of 30 months were strongly predictive of later reading ability. In fact, Scarborough (1991) suggests that the strongly supported phonological hypothesis may be incomplete, because there is also evidence of syntactic and morphological problems in dyslexic children, which cannot be explained in terms of phonological processing problems or as consequences of early reading failure.

The view that dyslexia is more than just a specific reading disability is shared by many researchers (see Locke, Hodgson, Macaruso, Roberts, Lambrecht-Smith & Guttentag 1997: 72–73). Studies that have reported delays in the acquisition of oral language skills by dyslexic children have led to a suggestion that specific language impairment (SLI) and dyslexia represent different manifestations of the same underlying disorder (see Bishop & Snowling 2004, for a review). Snowling, Bishop and Stothard (2000) have rejected this so called severity hypothesis which identifies dyslexia qualitatively as a similar impairment to SLI, differing only in degree. Instead of supporting the hypothesis that dyslexia is a mild form of SLI Snowling et al. claim that the difference between these two conditions is a qualitative one. Although both conditions are associated with literacy problems the underlying mechanisms are different. In dyslexic children poor phonological skills seem to cause the literacy problems whereas in SLI other linguistic problems are the main cause of reading difficulties, although limitations in phonological awareness may be involved in some cases as well. Bishop and Snowling (2004) introduced a two-dimensional model which provides a useful framework for thinking about subtypes of literacy problems and their relationships with other language difficulties. The quadrant model is presented in Figure 1. In this model dyslexia, SLI and poor reading comprehension are placed according to the findings made as to the phonological and morphological skills of children suffering these conditions. This model emphasizes that there is no reason to assume a straight continuum between dyslexia and SLI. Rather it is a question of different combinations of various linguistic skills.

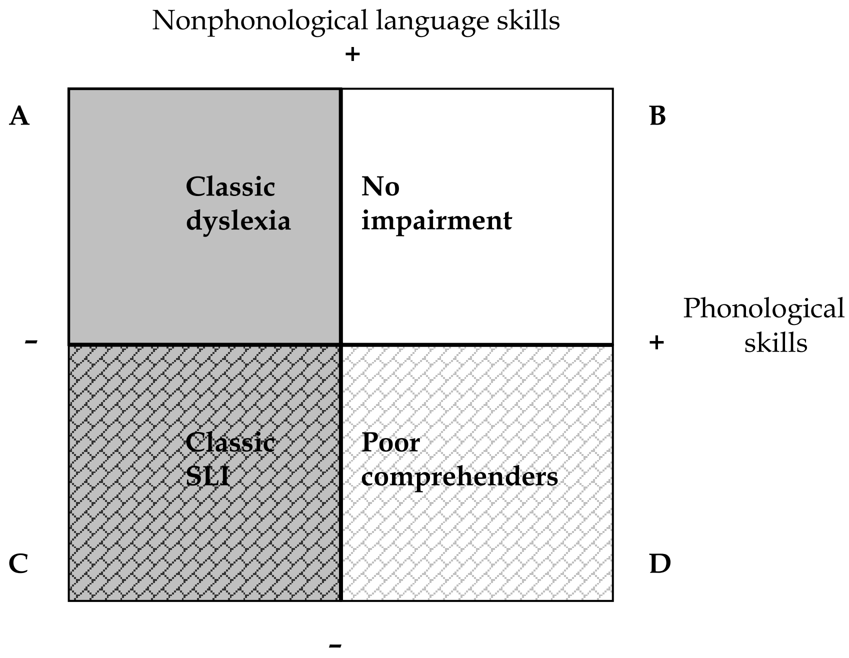


FIGURE 1 A two-dimensional model of the relationship between dyslexia and specific language impairment (Bishop & Snowling 2004: 859). Phonological skills are shown increasing from left to right and nonphonological skills from the bottom upwards.

The model gets support from a study on reading comprehension by Nation, Clarke, Marshall and Durand (2004). They found that a substantial minority of children with problems in reading comprehension actually meet the traditional criteria for SLI although neither parents nor teachers have noticed their condition because of their normal phonological skills. In Bishop and Snowling's model poor comprehenders are placed next to SLI children, the difference being only in phonological skills. Children with a classic form of dyslexia have their own quadrant with good nonphonological language and poor phonological skills. Some support for the model comes from brain structure studies as well. Dyslectic subjects and subjects with SLI deviate from the population mean in anatomical risk factor (ARF7), generated from seven measurements of brain anatomy. Moreover the deviations of dyslectics and subjects with SLI were in opposite directions from the population mean, and this suggests that the two disorders are qualitatively different conditions. (Leonard, Lombardino, Walsh, Eckert, Mockler, Rowe, Williams & DeBose 2002.)

Bishop's and Snowling's two-dimensional model is an oversimplification of the actual condition, and there is a growing consensus that neither SLI nor dyslexia is a homogeneous disorder. In fact, concentration on phonological deficit in poor readers may have caused researchers to neglect other linguistic deficits that might have a causal relationship to impaired reading (Bishop & Snowling 2004). There are also studies reporting evidence of morphological problems in dyslexic people. Leong and Parkinson (1995) suggested that poor readers may use strategies different from those used by good readers in producing derived or base forms of morphology. Joanisse, Manis, Keating and

Seidenberg (2000) found that inflectional morphology was deficient in dyslexic children, especially in a group of children with language impairment but also in a phonologically impaired group. Therefore in this study the focus is on morphosyntactic skills rather than phonological abilities. The study is inspired by the work of Scarborough (1990b, 1991), who found that syntactic problems are evident between 2½ and 4 years of age. She concluded that syntactic complexity in naturalistic language samples distinguished dyslectic children from their controls. However, the concept of complexity was left undefined. Apart from statistical analysis of results in quantitative measurements there is no precise description of what syntactic structures differentiated the groups and how the difference was actually manifested.

There are at least three different views on how syntactic skills are related to problems in phonological processes in dyslectic children. According to one of them, the lack of syntactic skills is a consequence of either delayed development in language skills or of structural deficiencies in the language system. Therefore the basic syntactic skills are defective. A second view argues that children use syntactic rules in the same way as their normally reading peers do and that their speech is grammatically correct. The deviant syntactic skills observed in dyslexic children are only a reflection of limitations in short term memory capacity which are caused by a basic difficulty in generating phonological codes. (For more detailed descriptions of the two views, see Deutsch & Bentin 1996 and Leikin & Assayag-Bouskila 2004.) The third view, formulated by Deutsch and Bentin (1996), suggests that lack of syntactic skills does not reflect absence of basic syntactic knowledge but is a sign of poor ability to use the knowledge in practice. Thus it is claimed that basic syntactic skills are not defective in dyslexic people. However, the different results in syntax are not due to phonological processing deficits, either. It is rather a question of a metalinguistic problem.

By now, it is evident that morphosyntactic skills and complexity are connected with dyslexia. For this reason, the data of this study comes partly from children with a high genetic risk of dyslexia, and a research framework providing for a comparison of two groups in their handling of structural complexity is appropriate.

1.2.3 The quantification of complexity

In the field of acquisition studies it has been important to compare individuals and groups, examine variation, standardize the linguistic skills of different groups defined by age or a given medical condition, and find linguistic matches among subjects participating in research projects. For these purposes there are several different scales of measurement which give numerical results suitable for quick comparison and further statistical analysis. Simultaneous progress on several linguistic levels can thus be presented on a linear scale, where linguistic development is taken to be shown by an increase in points or other units of quantity. Probably the most utilized index of this kind is Mean Length of Utterance (MLU; Brown 1973/1976: 77–78) with its many variations.

A more recently developed scale is the Index of Productive Syntax (IPSyn; Scarborough 1990a). Both scales are used in the present study. As measured on these scales children are put into developmental order: the higher the score, the more advanced they are considered to be in linguistic abilities. At the same time, such scales are taken to evaluate complexity of linguistic productions. In some complexity studies several scales of measurement have been used to confirm each others' results. During the past decade, MLU and Index of Productive Syntax have become one such pair (Scarborough 1990b; Rescorla, Dahlsgaard & Roberts 2000; Rescorla, Bascome, Lampard & Feeny 2001). Even though both scales are more or less sensitive to length, and high correlation has been found between them (Scarborough, Rescorla, Tager-Flusberg, Fowler & Sudhalter 1991) they are not equivalent (see Kemper, Rice & Chen 1995, for review).

The unfortunate fact that the concept of complexity in itself has been given very little if any serious attention within traditional child language research, even though complexity is a focus of many studies, has meant that complexity remains undefined and there is still no theoretical approach to the matter. One of the consequences is that in different studies MLU is said to be "a developmental index of language proficiency" (Johnston 2001: 161), "a general indicator of structural development" (Miller 1981: 25), "a valid predictor of syntactical complexity and diversity" (Rondal, Ghiotto, Bredart & Bachelet 1987: 444), and an "index of clausal complexity" (Blake, Quartaro & Onorati 1993: 139). Language proficiency, structural development, syntactic complexity and diversity, and clausal complexity are not synonyms. If anything, there is a hierarchy of meaning, language proficiency being the widest and clausal complexity the narrowest concept, and each of the characterizations of MLU is included in the preceding one. Both MLU values and IPSyn points are generated quite easily, but very little research is being done to find out what kinds of factors are hidden behind different values obtained and what the relationship is between these factors and complexity, if the latter is discussed and identified properly.

The reduction of complexity to figures is needed when results are analysed statistically but the facts hidden behind the scores are important in interpreting the results. If a scale is said to measure complexity, there should be knowledge about what kind of complexity and what aspects of it the method investigates. Otherwise, the results remain figures without any true value other than as a basis for compiling statistics, arranging them in order of magnitude, or comparing them with other figures of the same kind.

In this study the main reason for using Mean Length of Utterance and the Index of Productive Syntax is not to gain figures but to investigate what the figures actually tell about the structural complexity of the language a child is producing. Thus, this study seeks to gain an insight into variation in structural complexity through a suitable domain of language (Finnish morphosyntax), a phenomenon where structural complexity may play an important role (developmental dyslexia), and commonly used scales of measurement which have been said to focus on complexity (MLU and IPSyn).

1.3 Subjects

The participants in the study were 40 normally developing children from Finnish-speaking families living in the province of Central Finland. All the children were approximately 30 months old with ages ranging from 2;5.20 to 2;6.20. 20 of them had a high genetic risk of dyslexia and formed a risk group (RG; 10 boys and 10 girls), while the remaining 20 children belonged to the control group (CG; 12 boys and 8 girls). The subjects presented in Table 1 were selected from among the participants in the Jyväskylä Longitudinal Study of Dyslexia (JLD) headed by Professor Heikki Lyytinen of the Department of Psychology in the University of Jyväskylä.

TABLE 1 Subjects participating the study.

Control group				**Risk group**			
ID	Name	Gender	Age	ID	Name	Gender	Age
1	Tuija	F	2;6.7	21	Paula	F	2;6.7
2	Riikka	F	2;5.20	22	Anna	F	2;6.4
3	Laura	F	2;6.1	23	Anniina	F	2;6.1
4	Henna	F	2;6.4	24	Sanni	F	2;6.3
5	Liisa	F	2;6.4	25	Siiri	F	2;5.28
6	Ronja	F	2;5.21	26	Tiina	F	2;6.3
7	Taru	F	2;6.5	27	Janna	F	2;6.3
8	Aino	F	2;5.27	28	Jenni	F	2;6.3
9	Lassi	M	2;6.3	29	Elisa	F	2;6.1
10	Santeri	M	2;6.10	30	Elina	F	2;6.4
11	Jaakko	M	2;5.26	31	Aleksi	M	2;6.12
12	Lauri	M	2;6.6	32	Juho	M	2;6.20
13	Joel	M	2;5.23	33	Jukka	M	2;6.12
14	Patrik	M	2;6.3	34	Leo	M	2;6.4
15	Mika	M	2;6.4	35	Seppo	M	2;5.29
16	Kyösti	M	2;5.29	36	Sampo	M	2;6.5
17	Tatu	M	2;6.9	37	Henri	M	2;6.3
18	Saku	M	2;6.1	38	Pekka	M	2;6.10
19	Tuomo	M	2;6.13	39	Juuso	M	2;6.5
20	Tuomas	M	2;5.30	40	Risto	M	2;6.3

For the JLD risk group subjects (total N = 106) were selected through a three-stage procedure for screening families. The first stage was a short questionnaire sent to every family in the city of Jyväskylä and its surrounding communities expecting a baby during the period 1993–1996. The questionnaire was administered at local maternity clinics. The second stage was a more comprehensive questionnaire about parents' school experiences and reading and writing experiences, and it was sent to the families that had given their contact information in the first screening and thus expressed their interest in participating in the research project. At the third stage, parents who had reported serious reading and/or spelling difficulties and had at least one close relative having the same kind of problems, were invited to participate in an interview and a comprehensive assessment of their reading, spelling, and reading-related phonological and orthographical skills. Children included in the risk group were those with a parent who had at least three test scores more than 1 SD unit below the norm. (Lyytinen, Leinonen, Nikula, Aro & Leiwo 1995; Lyytinen, Poikkeus, Laakso, Eklund & Lyytinen 2001.)

The sample of 20 risk group children and 20 control group children in this study was selected by a researcher in the Department of Psychology, and the division into risk and control group was not revealed to the author of this study until the quantitative analyses (MLU and Index of Productive Syntax) were completed. All the children selected for the sample came from families with a similar socioeconomic status.

1.4 Data collection and transcription

The spontaneous speech samples used as data were collected during play sessions with an experimenter. These sessions, which were recorded both on audio and video tapes, took place in the children's homes immediately after a morphological test and a naming test conducted by the experimenter. Thus the children were already acquainted with both the experimenter and the equipment used in the recording (Super uni-directional electric condenser microphone on the table in front of the child and the experimenter, Canon UC8Hi video recorder and Sony Walkman Professional tape recorder) in a familiar environment before the play session. Every child had the same toys to play with and the experimenter divided the play into three phases, in which three different sets of toys were used. Giving the play a standard framework ensured that every child had equal opportunity to show linguistic skills in similar situations.

The first set of toys consisted of little dolls and doll's house furniture (beds and bedclothes, chairs, a table and a telephone), dishes (plates, cups, a jug, a spoon and a fork), a feeding bottle, a syringe, and a bottle of medicine. At the beginning of the first stage of play the experimenter introduced the dolls as the mother, father and children of a family living in the house. After a few moments of play, the second set of toys, consisting of farm animals (ponies, a dog, a hen, a pig and a sheep), some animal food (play dough in a trough) and a pen for the animals, was introduced. The experimenter told the child that these animals belonged to the family introduced earlier, and it was time to take care of them now. After a while the third set of toys, consisting of a tractor and a trailer, tools (a screwdriver and a pair of pliers) and a petrol station, was introduced.

During the first stage the children usually pretended to make some breakfast for the whole family after the experimenter said that it was morning and everybody had woken up. They also took care of a sick doll and pretended to give it some medicine with a syringe. In the second stage the children were very interested in the play dough and started to model snakes, worms etc. Usually they also moved the animals around and gave them some food.

During this stage the experimenter pushed one of the ponies so that it fell down and hurt its leg and then made the child call a vet to find out how to take care of the pony after the accident. During the third stage of the play session the children moved the tractor around and took the dolls or animal food from place to place on the trailer. They also filled up the tank of the tractor and repaired the tyres or the engine of the tractor with the tools available. The play session ended when the child was no longer interested in playing.

The conversations between each child and the experimenter during the play sessions were transcribed from the audio tapes in CHAT-format by using the conventions of the Child Language Data Exchange System (CHILDES; MacWhinney 1991). The purpose of the transcription was to make it possible to analyse the data morphologically, and so a very narrow phonetic transcription procedure was not used. However, in Finnish quantity may distinguish word meanings, like in *tuli* 'fire, *tuuli* 'wind', and *tulli* 'customs'. In the same way, it may be a distinguishing feature between some frequently used cases, like nominative and partitive (*kirja* 'book' vs. *kirjaa* 'book-PARTIT') as well as genitive and illative (*auton* 'car-GEN' vs. *autoon* 'car-ILL'). Thus, special attention was paid to phoneme length, because it affects both semantic and morphological interpretation of language production. Pauses were not measured, but they were categorized in three different groups of relative length. The manner of utterance (whispering, shouting etc.) as well as overlapping of turns and alternative transcriptions of a word were also documented. During the transcription notes were also taken on background noise and the technical quality of the tapes to be taken into account in the interpretation of the language produced.

The video recordings were utilized in three ways. First, the action taking place during the play sessions and the references of deictic expressions, such as demonstrative pronouns, were noted down in the transcription. Second, the video recordings were used in checking the transcriptions first by the transcriber and then by another trained colleague. And third, during the checking of transcriptions, the division of children's language production into utterances was confirmed. In this study, an utterance is defined both phonetically and semantically as a sequence of speech which is separated from other utterances prosodically with falling intonation contour and which conveys a semantic whole interpreted in a situational context. In most cases there is a pause between utterances, but not every pause marks an utterance boundary. (For deciding utterance boundaries see also Klee & Fitzgerald 1985: 255.) In all the data base of this study consisted of 7514 intelligible speech utterances which were coded morphologically. On average there were 187.85 utterances from each child, the range being from 77 to 392.

Transcription is the first step towards data analysis. Selection of the final sample was preceded by a careful morphological coding of the data. Both the coding and selection procedure are described in detail in Chapter 3, together with the methods of the two quantitative analyses (MLU and IPSyn).

1.5 Aims, research questions, and outline of the study

The basic aim of this study is to investigate the manifestation of structural complexity and its variation in spontaneous speech produced by 30-month-old children acquiring Finnish as their first language. Special attention is paid to the morphosyntactic structure of Finnish, high genetic risk of dyslexia and the quantitative methods employing MLU and the Index of Productive Syntax used to evaluate complexity in child language research. The main research questions are as follows:

1. How is morphosyntactic complexity manifested in children's spontaneous speech?
2. What is the relation between risk group children (children with high genetic risk of dyslexia) and control group children (children without high genetic risk of dyslexia) when special attention is paid to the complexity of morphosyntactic structures in their spontaneous speech production?
3. What kind of information about structural complexity in utterances do commonly used measuring scales like MLU and IPSyn give?

Before moving on to seek concrete answers to these research questions, I will discuss the notion of complexity in general and structural complexity in particular in Chapter 2. In the same context, I will make some short comments on the position I take in this study with respect to language and language acquisition.

After specifying the starting points of the study, I will first introduce MLU and IPSYn in detail in Chapter 3 and discuss their relationship to structural complexity. In the first stage complexity analysis will be performed in a traditional way, by measuring complexity using MLU and the Index of Productive Syntax. The results obtained on the two scales will be compared to determine whether they offer the same view of the complexity of children's utterances. This will be reported in Chapter 4.

The second stage will involve more qualitative analysis of the data and the focus will be on Questions 1 and 3. Children's utterances will be investigated using Utterance Analysis, which is designed to describe the morphosyntactic structure of each utterance in detail. The method itself is introduced in Chapter 5, and it is consistent with the notion of structural complexity discussed in Chapter 2. Utterance Analysis will be used to investigate what kind of structural phenomena lie behind the quantitative results. In other words, the manifestation of morphosyntactic complexity in utterances will be discussed and its relation to the quantitative scores derived from MLU and IPSyn will be dealt with. These issues are reported in Chapters 6–8.

The last step in dealing with the research questions will be to return to the issue of Question 2: does the relationship between risk group and control group look different now that structural complexity is dealt with in a more qualitative manner? This will be discussed in Chapter 9. The results of the study as a whole will be summarized and conclusions presented in Chapter 10.

2 COMPLEXITY

In this chapter the theoretical background as well as the basic concepts of the study are discussed. The study has been carried out within the framework of absolute complexity which approaches the issue from an information theoretical point of view. *Utterance*, *structure* and *morphosyntactic complexity* are introduced and defined as essential concepts for the study. Finally, connections to a usage-based model of grammar and language acquisition are explicated.

2.1 Complexity in information theory

Everyone agrees that there are 'simple' objects as well as 'complex' ones, but describing their complexity (or simplicity) may result in very different descriptions. In the 1960s, in the field of information theory Ray Solomonoff, Andrei Kolgomorov and Gregory Chaitin independently introduced a theory of algorithmic information content (AIC) which actually referred to a length of a computer program (Gell-Mann 1994: 35). The major innovation of this theory, also referred to as the theory of Kolgomorov complexity, was that with the help of algorithms it was possible to restrict the arbitrariness of the descriptions of complexity and define complexity as an invariant concept (Gammerman & Vovk 1999).

According to algorithmic information content, the complexity of an object is measured by the length of the shortest possible specification or description of it. The longer the shortest specification is, the more complex the object is. In order to make the shortest possible description one needs to recognize a pattern within the object being measured. A pattern is a way of simplifying the representation of the object, and of generating its shortest possible specification. This is easily illustrated with strings of digits:

a) 286286286286286 → pattern = 286
b) 28654286542865428654 → pattern = 28654

As we can see, both strings have an iterative pattern, and because the pattern of the string b) needs a longer description (5 digits) we can conclude that it is more complex than string a) with only a 3-digit pattern. Another conclusion is that according to this theory the most complex string of digits is one with no internal pattern at all (e.g. 83406713294072519…), since it is not compressible, and therefore the shortest possible description is actually the string itself. (Dahl 2004: 21–22, 24; Gell-Mann 1994: 34–36.)

The theory of algorithmic information content requires compressibility that results in a single, iterative pattern in order to make the shortest description of an object equal the pattern. Thus, complexity comparison between two objects is actually done between the two compressed patterns. Logically, complexity culminates in random, un-patterned strings. However, algorithmic information content does not correspond to what is meant by complexity in many situations. Not all objects can be compressed into a single pattern like algorithmic information content. There is more likely to be a string of different realizations of a given pattern, a set of regularities. This is a perspective which Gell-Mann (1994: 50; see also Dahl 2004: 24) calls *effective complexity*, which refers to a schema of identified regularities abstracted from the available data by complex adaptive systems (Gell-Mann 1994: 55–56).

Complex adaptive systems, like human beings, acquire information about the environment they live in and try to identify regularities in that information. The next step is to condense these regularities into a schema or a model and act in the environment on the basis of the schema. If there are competing schemata the feedback that a complex adaptive system gets from the environment influences the competition among the schemata. One of them may get stronger while the others get weaker or are thrown aside. (Gell-Mann 1994: 17.)

A good example of how complex adaptive systems behave in identifying regularities and condensing them into schemata is a child acquiring a language. To be able to use the language a child must notice the regularities in the ways things are expressed in the particular language. In other words, he or she must abstract the underlying grammar (among other things) and realize how to apply the schema in new linguistic situations. The effective complexity of the grammar would be the length of the schema that the child is able to condense from the regularities. Of course, in the case of children acquiring a language, the schema called grammar is under constant modification and its length of is thus ever changing. (Gell-Mann 1994: 53–54.)

2.2 Different approaches to complexity in language

2.2.1 The absolute approach to complexity

It is possible to examine linguistic complexity from many points of view, and different perspectives lead to different conclusions as to what is and what is not considered to be complex. Two major starting points in defining complexity are language users on the one hand, and language itself on the other. These perspectives are referred to respectively as the relative and the absolute[1] approaches (Miestamo 2006).

Östen Dahl (2004), representing the absolute view, approaches the problem of linguistic complexity from the information theoretical point of view. He agrees with the idea that the complexity of an object is to be measured by the length of the shortest possible specification or description of it. Dahl also argues that complexity should be as objective a notion as possible, distinct from the use and users of the system whose complexity we are evaluating. (Dahl 2004: 21, 39–40).

The maximum objectivity claimed by Dahl (2004: 39) characterises the absolute approach to linguistic complexity, which concentrates on language itself, without paying attention to language users. The focus is on, for example, categorical elaboration, number of units, number of rules, etc. (Miestamo 2006). This approach has recently been introduced in the field of language typology with McWhorter's (2001) criticism of the widely accepted assumption that the grammars of all languages are equal in terms of complexity and that simplicity in one grammatical area is always counterbalanced by complexity in another. In his article, McWhorter introduces a scale of overall grammatical complexity, on the basis of which he claims that creole grammars are simpler than the grammars of other languages. McWhorter uses the following principles in identifying complex grammars: a) a phonemic inventory is more complex to the extent that it has more marked members, b) the syntax of one language is more complex than that of another to the extent that it requires the processing of more rules, c) one grammar is more complex than another to the extent that it gives overt and grammaticalized expression to more fine-grained semantic and/or pragmatic distinctions than another, and d) inflectional morphology renders one grammar more complex than another in most cases. (McWhorter 2001: 135-137.) There is a concordance between McWhorter's ideas of what makes a language more complex and the information theoretical identification of complexity emphasized by Dahl. According to McWhorter's principles, complexity is associated with increased number of marked members, syntactic rules, and fine-grained distinctions, which imply longer descriptions of the linguistic system.

While McWhorter (2001) describes the overall complexity of a language, Dahl (2004) distinguishes *system complexity* and *structural complexity*. System

---

1 Dahl (2004) also refers to absolute and relative complexity, but he uses these concepts in a different way, in connection with the length of description needed for characterizing an object. Absolute complexity includes all possible aspects of an entity. For example, when characterizing a person we should describe him/her as an object with two legs, two arms, five fingers on both hands, etc., whereas a description with relative complexity approach would ignore the information we all share already and concentrate on aspects that are not common to all the entities in questions, that is, human beings in this case. Choice of either absolute or relative complexity will affect the length of description characterizing an object, and this, for its part, will affect how complex the object is defined as being, if complexity is discussed in an information theoretical framework. (Dahl 2004: 25–26.)

complexity refers to both resources and regulations that determine how to use the resources. The distinction between resources and regulations resembles the one between lexicon and grammar but does not coincide with it. System complexity is a matter of an entire language, and thus comparable to McWhorter's ideas, but in the case of structural complexity Dahl refers to the grammatical structure of individual expressions and utterances. He introduces two other ways of considering complexity. One is *phonetic weight*[2], which tells whether an expression is phonetically heavy or light, and the other is *length of derivational history,* which refers to the number of steps necessary to generate an expression in a formal system. There may be trade-offs between the different types of complexity. For example, a more complex structure may be even lighter phonetically than a simpler structure. (Dahl 2004: 40–45.)

It is important to keep the absolute approach to linguistic complexity apart from notions like *difficulty*, *expressive power*, or *efficiency*. If one language is said to have a simpler grammar than another this claim does not involve any value judgement. A simpler grammar is not necessarily an easier one to process, and it is by no means evidence of the language being primitive or inefficient as a means of communication. (Miestamo 2006; McWhorter 2001.) In fact, McWhorter introduces the interesting idea that highly elaborated grammars are actually easier rather than harder to process because they make distinctions more clearly and thus leave less to context (McWhorter 2001: 135). Kusters (2003: 2–3) notes that speaking about the complexity or simplicity of a language has not always been an issue free from value judgements. History has shown that the characterization of a language as simple is only a short step away from characterizing the speakers of the language as simple. The fear that "behind measurement of complexity the measurement of skull is hidden" (Kusters 2003: 2) has resurfaced in connection with studies on the complexity of creole languages. Thus, complexity has been a somewhat dangerous issue within linguistics and this may be one reason why it has so often been dismissed without proper discussion.

Absolute complexity does not equal objective truth, because there is always a subjective aspect within it. When describing an entity in the framework of absolute complexity, there is always someone who has to make choices and decisions while formulating the description. One has to answer questions like 'In what context is the complexity of an entity defined?', 'How finely detailed is the description?', 'What is the observational unit?' and so on. Each choice and decision will interact with the outcome, and absolute complexity is absolute only within the conditions the analyser has created by answering essential and inevitable questions such as those presented above. (Cf. Gell-Mann 1994: 33.) Kusters (2003: 6) has come to similar conclusions, and argues that a truly absolute approach to complexity does not exist. He sees complexity as a relation between two entities, one being a language and the other an evaluator of a language. The evaluator can be a language user (e.g. native speaker, non-native speaker, L1

2 Dahl's concept of phonetic weight originates from Langacker's notion of "signal simplicity", which is defined by Langacker as "economy in regard to the production of the physical speech signal" (Dahl 2004: 43).

learner, or L2 learner) or a language analyser (a linguist applying some linguistic theory). Thus, by Kusters' definition, linguistic complexity can be analysed only from a relative perspective.

To my mind, Kusters is right to argue that an absolute approach to complexity is logically impossible (Kusters 2003: 6), if we interpret 'absolute' literally. However, I find the idea of "absolute" complexity very useful, since it concentrates on evaluating language whereas Kusters focuses on the user or evaluator, and thus tries to answer questions such as what is experienced as difficult. Therefore, instead of speaking about absolute complexity it would be less provocative to call it a language-based approach as opposed to a user-based view to complexity.

2.2.2 The relative approach to complexity

In the relative approach to linguistic complexity the focus is on the language user – native speaker, non-native speaker or language learner – and the main question is "What is complex to whom?" The relative approach examines complexity in relation to linguistic processes, and thus the resulting notion of complexity strongly depends on which linguistic process and which group of language users is in question. Exactly the same grammatical structure may be complex from the L2 learner's point of view while it may not cause any difficulties in the L1 learner's acquisition process and vice versa. (Kusters & Muysken 2001: 184.) It is also possible that grammatical expressions are difficult from the point of view of perception or comprehension but still easy to produce or acquire.

A good example of applying the relative perspective to classifying linguistic elements as either complex or non-complex is Peters' (1997) suggestion as to the complexity of bound morphemes. According to her, both phonological and semantic characteristics may cause morphemes to be easy or difficult to acquire. The properties characterizing easily acquired morphemes are high frequency, easiness in segmenting, fixed position relative to an open class stem, clear function, easily recognized form, and final position. By contrast, infixes, pormanteau morphemes, homophonous morphemes, and morphemes with "bleaching" of semantic function are hard to acquire. (Peters 1997: 181–182.) According to Peters (1997), ease of perception, in other words the saliency of a morpheme, seems to relate to ease in the process of acquisition. However, bound morphemes or other linguistic elements are seldom characterized by only one property, and the balance between properties as well as possible bias towards one of them is likely to be the most important issue in placing a bound morpheme on the scale of complexity.

In the preceding paragraphs, the words *complexity* and *difficulty* occur side by side, and that is not accidental. When there is an agent, whose linguistic processes are at the centre of our interests, difficulties in processing are interpreted as evidence of complexity: what is difficult must be complex as well. Another concept which needs to be considered in connection with complexity in this sense is *cost*. The relationship between *cost* and *complexity* is not as direct as that between *difficulty* and *complexity*. *Cost* refers to the resources needed by an agent to achieve a certain goal. If a task demands a lot of resources or forces an agent to the limits of his or her capacity, the task is experienced as difficult, that is, complex. Linking complexity tightly with difficulty and cost turns it into a subjective concept and raises many questions. Using language involves many kinds of tasks, and agents performing these tasks are different not only in their relationship to the target language but also as individuals. Thus, identifying complexity with difficulty would involve assessing linguistic tasks as well as the subjective experience of different agents. (Dahl 2004: 39–40.)

Relative complexity is dynamic by nature. For example, as language learning or acquisition proceeds, different features of a language may appear difficult. What is difficult at the moment for a language user is probably found easy after a while. However, the direction of change in experienced complexity is not necessarily from complex (difficult) to simple (easy), since structures being acquired as a whole may turn via more sophisticated analysis caused by wider linguistic experience into a more fine-grained system, which produces difficulties for a time. Using undivided units feels easier than acquiring and applying the more abstract system behind the units.

The relative approach to linguistic complexity is a multifaceted issue, since it involves at the least the relation between a language, the language user as an individual, the linguistic proficiency of the language user, a specific domain of linguistic process (e.g. perception, production or comprehension), and changes over time and linguistic experience. Thus, a careful demarcation of variables is always needed, when complexity of language is approached from the relative perspective.

2.3 The concept of structural complexity in this study

2.3.1 Defining the basic concepts: utterance and structure

The view of linguistic complexity applied in this study is the absolute approach in the sense Miestamo (2006) defines it. This view is consistent with Dahl's (2004: 39) insistence that the question of the complexity of a language should be kept as objective as possible, in isolation from the language user. I find an approach which focuses on the language and not on the language user more suitable for a study where the subjects are small children and their experiences and feelings about language and its structure are extremely difficult to investigate. It is also risky to leap to conclusions as to the difficulty or easiness of linguistic structures merely on the basis of occurrences of these structures in conversation during an "everyday" play session. The absolute approach makes it possible to evaluate the complexity of structures in small data samples because evidence of experienced difficulty is not needed. Yet another reason to support the absolute approach in this study is the need to avoid the possibility of circular reasoning. Half of the subjects in this study have a high genetic risk of dyslexia. One of the aims of the study is to find out whether the risk children and control children differ in terms of the structural complexity of their utterances. In a relative approach the complexity of utterances would be identified on the basis of experienced difficulties in language use. Therefore, what is considered complex would be affected by the possibly existing difficulties in language processing of risk children. In the absolute approach the complexity of utterances is not identified in terms of experiences in either of the groups, and thus it is a more neutral way of investigating the issue.

My aim is to investigate the structural complexity of utterances produced by children, and the target subsystem in the linguistic structure is morphosyntax. However, the concepts of *utterance, structure* and *complexity* need more clarification. Talking about the structure of *utterances* already implies that in my view the composition of combinations and wholes is emphasized rather than individual subunits. It is more important to look at what is constructed out of subunits than separate them from their actual context and try to analyse them one at a time. Taking the utterance as the fundamental unit of a linguistic whole is based on several considerations that will be discussed below.

Within the framework of a usage-based model of language acquisition, following cognitive linguistics and functionally oriented theorists, Tomasello (2000c: 63) suggests that the most fundamental psycholinguistic unit is the utterance, and defines it as "a linguistic act in which one person expresses towards another, within a single intonation contour, a relatively coherent communicative intention in a communicative context". Additional support in favour of the utterance comes from the framework of conversation analysis, which takes the utterance as the elementary unit of lexical turns in a conversation (for a review, see Linell 1998: 159–161). What is in common to both of these views is that they are based on actual language use and communication and do not presuppose any innate abstract system behind language as in the generative tradition.

Taking the utterance as the basic unit in analysis has some important consequences. First, there is no need to apply any traditional sentence structure analysis. Thus, no judgements or categorisations are made on the basis of clausal completeness. Secondly, as a corollary of this, all utterances are equally important. In an everyday conversation normatively inaccurate utterances have a major role. In my opinion, a great deal information concerning morphosyntactic complexity is provided by "incomplete" utterances, and thus it would be a serious mistake to exclude them from analysis in the way that, for example, Yaruss (1999) does. It is reasonable to think that in these elliptical utterances focused on conveying essential information within a conversation, both syntactic and morphological elaboration may reach a higher level of complexity than in utterances containing all the constituents needed for a grammatically complete sentence. This so called trade-off effect (increase in complexity in one area resulting in decrease in complexity in another area) is considered in several studies in the literature. Bock (1982) focuses especially on the interaction between lexical and syntactic processing and claims that word retrieval during sentence formulation influences the syntactic form of a sentence: the use of words occurring in earlier sentences gives more scope for syntactic formulation, while correspondingly the retrieval of new words limits the syntactic structure of a sentence. Crystal (1987) considers the trade-off effect from a wider point of view, with no assumption of a hierarchy governing different linguistic subsystems determining which of them is the cause and which the receiver in a trade-off. He uses the metaphor of a bucket with holes in it to describe linguistic processing capacity. When the amount of "linguistic water" reaches the level of the holes in the bucket, each extra "drop" of phonology, syntax, semantics etc. causes an overflow of a "drop" of "linguistic water" through the holes. Thus, new demands in one linguistic area may cause lower processing capacity in another area of language. Crystal applies the leaky bucket theory especially to language disability. However, it is worth considering whether young children with normal linguistic development are subject to extreme conditions in their language processing. Even though both Bock and Crystal are concerned with trade-offs between linguistic areas, it is possible that the same kind of interrelations also play a role within a linguistic area. Elliptical utterances might be examples of such behaviour: when the demands at sentence level are low and there are few syntactic components present, the demands at phrasal level are free to increase in terms of syntactic elaboration, not to mention the morphological possibilities of word modification.

Let us now turn to the concept of structure. Dahl (2004: 46–50) distinguishes *choice structure* and *output structure*. He uses the term choice structure to refer to decisions made by a language user relative to the language system in formulating an expression: for example, which words and word forms are chosen and how free or bound these choices are. Thus, the choice structure reflects the grammar of the language, but in my opinion, it reflects above all the language user's grammar and what choices there are available to a learner. It is particularly difficult to decode the choice structure of utterances especially because of the influence of each language user's individual grammar, which may differ widely from another's in respect of the composition of its units. Langacker (1987: 57) defines grammar as a "structured inventory of conventional linguistic units". The units can be specific expressions of different kinds (inflected word forms or multiword expressions stored as units in a grammar) as well as schemata that can be used to compute an expression or word form needed. Which expressions have the status of a unit varies from speaker to speaker and changes with linguistic experience. (Langacker 1988b: 130.) Altenberg (1990) has estimated that different kinds of formulaic expressions, either fixed or more flexible ones, may encompass as much as 70 % of adult native language (Wray & Perkins 2000). In the case of children acquiring their first language, the proportion of formulas and various redundant expressions is even greater (Tomasello 2000c; Lieven, Behrens, Speares & Tomasello 2003). Access to these expressions and, further, to the choice structure of utterances would require the longitudinal follow-up of a subject. Otherwise recycled linguistic units and usage-based syntactic operations of "cut and paste" (Tomasello 2000c: 74) will not be discovered.

By *output structure* Dahl (2004: 48–49) means the actual realization of an utterance. In analysing choice structure we need to know precisely what the units in a subject's grammar are. Otherwise it would not be possible to identify the decisions and choices made by a speaker when constructing an utterance. But in the analysis of output structure, it is actually the researcher's choice what kind of framework is applied. The aim of analysis is not to describe the construction process and the speaker's actual linguistic units but the final outcome of the process. In this study the focus is on the output structure of utterances.

Taking the utterance as the basic unit of analysis without analysing utterances as clauses is consistent with a usage-based theory of language acquisition. Tomasello (2000a) has stated that young children's productive use of grammatical forms or structures is very limited and only children three years of age and older show evidence of the existence of more abstract linguistic constructions. In other words, Tomasello (2000a: 210–211; 2003: 142–143) opposes the so-called continuity assumption postulated in generative approaches to language acquisition. The continuity assumption states that children and adults have similar cognitive as well as grammatical mechanisms. This means that the grammars of both adults and children represent the same kind of rule and class systems, and thus it is not necessary to assume that children's grammars are essentially different from those of adults. (Pinker 1984: 7.) Tomasello (2000a: 229; 2003: 142–143) argues that only the continuity assumption framework allows a researcher to use the concepts of formal adult grammar because of the underlying belief in innate and unchangeable categories.

Though strongly approving Tomasello's idea that young children do not have syntactic categories similar to those in adults, I now return to the concept of output structure. I stated that it is the researcher's choice which framework is applied in the analysis of output structure. This implies that the language user and language analyser do not need to have identical categories. Since, because my data is cross-sectional, I do not have access to children's actual units of grammar, I use phrasal and morphological categories in the way they are used in the descriptive grammar of adult Finnish in Iso suomen kielioppi (ISK 2004), totally aware of the fact that they may not be identical to the units children actually have in their individual grammars. There are two major benefits in choosing analytical tools on the basis of the descriptive grammar of adults. First, it is a systematic and unbiased way to analyse utterances because the same categories and rules are applied to every child's utterances. Secondly, by using these tools it is possible to create a multidimensional picture of utterances. Separation of phrases illustrates the linearity of the major parts of utterances, and analysis of every phrase into its morphemes shows the inner composition of phrases. In this study multidimensionality is an essential feature of structural complexity and it will be discussed in detail in the following section.

2.3.2 Morphosyntactic complexity

In studies concerned with the structural complexity of language production complexity is very often quantified by counting specific linguistic units, for example morphemes in the case of MLU. Such an approach characterizes language structure and its complexity as uni-dimensional. In my view of structural complexity, however, multidimensionality is an essential property. In utterances there is a linear structure which consists of phrases, for example NP + V, and other components such as interjections. A phrase can be unelaborated, that is, a plain word without any inflection or syntactic extension. Unelaborated phrases are unidimensional constructions in a one single layer. However, many phrases produced in utterances are elaborated morphologically, syntactically or in both ways, and this creates a layered structure and multidimensionality. Each inflection creates a new layer within a phrase and so does each syntactic extension such as a modifier or an auxiliary. Thus, in this context, multidimensionality means that an utterance has a linear structure consisting of phrases and in addition to that each phrase has its own layered structure. It is this layered structure that has been unrecognized in earlier studies concerning structural complexity in child language. However, I find it essential to emphasize the stratification of utterances, since I categorize elaborated phrases as morphosyntactically more complex than unelaborated ones. Morphosyntactic complexity may grow in two ways, in the first place by an increase in the components (e. g. phrases) in an utterance, and secondly by increased elaboration within a syntactic component. Thus the morphosyntactic complexity of an utterance is a combination of the number of syntactic components and the number of elaborations in each of them.

Describing structural complexity as a multidimensional phenomenon and analysing it phrase by phrase seems to have points of resemblance to constituent analysis of sentences. The most extensive similarity between these two is the idea that a verbal entity (a sentence or an utterance) consists of parts (constituents or components) that have an inner grammatical structure of their own. Thus, a sentence or an utterance is not a simple chain of words. In constituent analysis, a sentence (S) is usually first divided into two parts, a subject (NP) and a predicate (VP), and each of these is further analysed into its subconstituents. This creates a picture of a sentence as a hierarchical structure which is very often presented in the form of a tree diagram, where a higher node dominates the lower nodes connected to it (Figure 2).

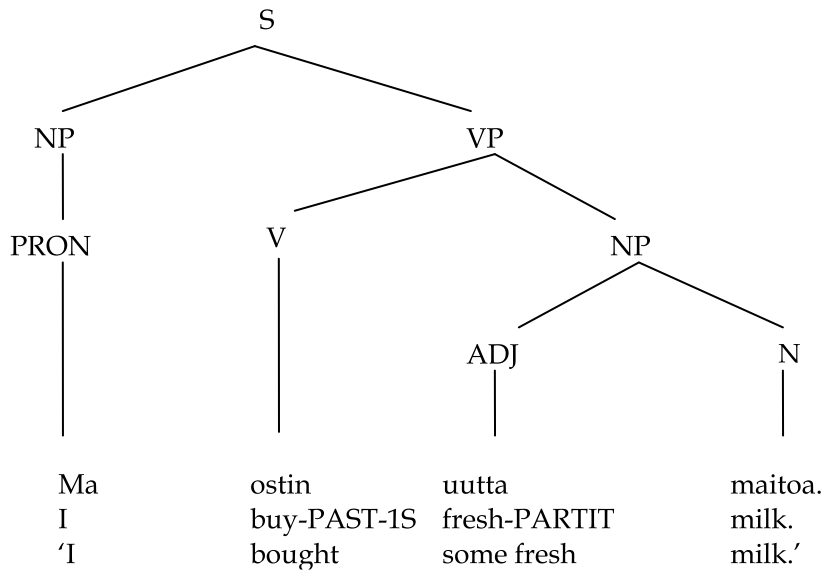


FIGURE 2 A constituent structure of an utterance produced by Kyösti (ID 16).

However, despite the similar basic idea, constituent analysis has several features that do not coincide with my way of describing morphosyntactic complexity. First of all, in constituent analysis the main object of analysis is a sentence whereas in my analysis a complete sentence structure is not required. Consequently, many utterances cannot be divided into subject part (NP) and predicate part (VP), as is done in constituent analysis. It is more important for my purpose to know what kind of phrases there are in an utterance than in what relation they are to each other or which part is dominant and which parts are subordinate. In fact, many of the utterances produced in a conversation and especially by children of this age are difficult to analyse into parts that are in some hierarchical relation to each other, as the following examples from Tatu (ID 17) show:

ADVP	V	V	NP	NP
Jo	*on*	*ampua*	*kaikkia*	*pyssy.*
Already	is	shoot-1INF	everybody-PARTIT	gun

'Already shooting at everybody with a gun.' (A possible interpretation)

NP	NP	NP	V
Juulia	*ponit*	*niitä*	*on.*
Juulia	pony-PL	they-PARTIT	is.

'They are Juulia's ponies.' (A possible interpretation)

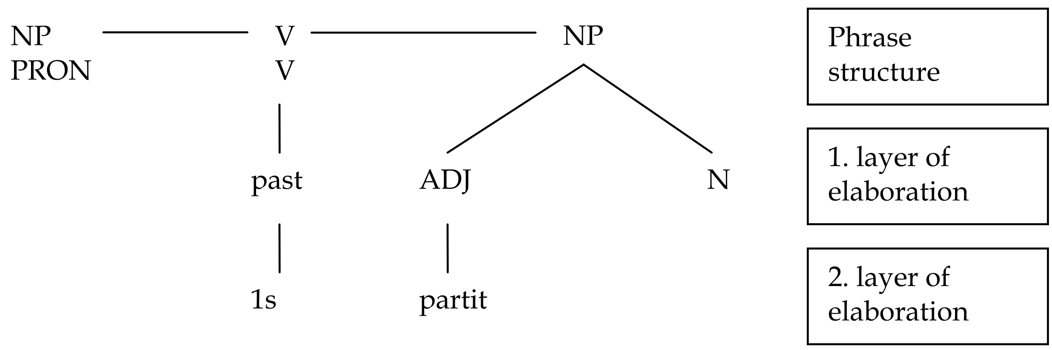
NP	NP	NP
Kahvia	*kaakaota*	*jätskiä.*
Coffee	cacao	ice-cream.

'Coffee, cacao, ice-cream.'

PTL	CONJ	NP
Niin	*ja*	*kampa.*
Yes	and	comb.

'Yes and a comb.'

There is yet another clear difference between constituent analysis and the way utterance structures are understood in this study. Here, morphological structure plays an important role but in constituent analysis the smallest element included is a word. Thus, while constituent analysis focuses on syntax and on hierarchical roles within it structural complexity is seen here as a morphosyntactic composition without an emphasis on hierarchy. The layered structure of utterances is created by elaboration and not by dominance. In Figure 3, the same utterance as in Figure 2 is described in terms of phrasal composition.



Ma	ost-i-n	uut-ta	maitoa.
I	buy-PAST-1S	fresh-PARTIT	milk.
'I	bought	some fresh	milk.'

FIGURE 3 The phrasal composition of an utterance produced by Kyösti (ID 16). The connecting lines downwards express morphological and syntactic elaboration and not dominance.

Replacing the linear interpretation of morphosyntactic complexity in children's utterances with a layered one does not affect the underlying theory in terms of absolute complexity. In my opinion, the information theoretical point of view emphasized by Dahl (2004) does not restrict the description to a linear form. Descriptions become more open to various interpretations because the existence of layers means it is not a straightforward matter to determine the degree of complexity. It is evident that there are clear-cut examples of simple and complex utterances but between them there will be a substantial range of utterances in which advances in one dimension make up for deficiency in another. Thus, it is a matter for debate which is more complex: a three-word utterance without any elaboration or a one-word utterance with two instances of morphological elaboration. Is combining units in a linear fashion more complex or simpler than elaborating a unit? When comparing two elaborated phrases, which is more complex: the one with syntactic elaboration or the one with morphological elaboration? In my opinion, there is no clear need even to answer these questions, since if we accept the multidimensional idea of morphosyntactic complexity we have to accept also the fact that there must be different ways in which the same level of complexity may be reached and we do not have to determine which way is more complex than the others. This leads to the conclusion that my approach does not provide an absolute *measure* of complexity, similar to MLU or IPSyn, but a way to describe varying forms of morphosyntactic complexity which takes equal account of syntactic and morphological elaboration.

To conclude, my concept of morphosyntactic complexity arises from phrases in linear order and especially from the layered elaboration of phrases simultaneously. Complexity can expand in different directions, either through the addition of new phrases or their elaboration them morphologically, syntactically, or in both ways. Syntactic and morphological elaboration are both equally significant. The aim is to describe different forms of complexity without putting them into any particular order, which means that similar levels of morphosyntactic complexity can be reached in various ways.

2.3.3 Connections to a usage-based model of grammar and language acquisition

On several occasions I have already shown that many of the ideas and choices made in this study have points of contact with Langacker's and Tomasello's work within the framework of a usage-based model of grammar and language acquisition. However, because of the cross-sectional nature of my data, it is not possible rigorously to follow the ideas of the model, and therefore I prefer to say that this work is done in the spirit of their ideas, but without following them in every detail. In this chapter, I will sum up the connections the study has with a usage-based model.

Langacker (1988b: 131–133) advocates a maximalist, non-reductive and bottom-up view of linguistic description as opposed to the search for a minimalist, reductive and top-down approach in the generative tradition. On becoming familiar with the currently accepted measuring scales for complexity, for example, Mean Length of Utterance and Index of Productive Syntax, I was not satisfied with them mainly because they reduced a phenomenon which I regarded as multivariable to a quantitative mean value or number of points and at the same time effectively concealed the actual language in question. There was no way to connect a certain value or number of points with a specification of language use. I was not satisfied with the idea that a simple number expresses aspects of morphosyntactic complexity. Such a minimalist and reductive approach does have advantages, for example in statistical analysis, but nevertheless there should be some knowledge of what kind of complexity these figures actually reflect. To find this out I needed first a definition of structural complexity and then a maximalist and non-reductive way to find out how complexity thus understood manifests itself in my data. This is why utterances will be analysed phrase by phrase. Each phrase and unit within the phrase will be registered as well as the relations between the units within each phrase. The morphosyntactic relations are not reduced to a linear chain of abstract observational units. Morphemes are identified one by one and all the units involved in an utterance are placed within a multilayered framework to get an exhaustive picture of how morphological and syntactic devices work together and what kind of structural complexity arises from the interaction. This way of analysing an utterance has common features with what Langacker (1988a: 26–27) calls a "compositional path".

Tomasello (2000c: 63) argues that the utterance is the most fundamental psycholinguistic unit. In my analysis of complexity the starting point for description is always an actual utterance and whatever may occur within it. This implies that syntactical units such as sentence, clause, subject, object or predicate are not defined, and clausal completeness is not evaluated. None of these is relevant for utterances, which are identified as "linguistic act[s] in which one person expresses towards another, within a single intonation contour, a relatively coherent communicative intention in a communicative context" (Tomasello 2000c: 63), and not as linguistic units including functional subunits belonging to certain syntactic categories.

According to a usage-based model, grammatical units emerge through the use of language**.** Grammar is actually an inventory of symbolic resources and it is a language user's task to exploit these resources to create novel expressions (Langacker 1988a: 13). Thus, there is no need to assume the existence of an innate system of grammar which is triggered by the linguistic environment and gradually taken into use. Grammar consists of conventional linguistic units and abstract schemata. Through linguistic experience units first acquired as wholes may be analysed into structural subunits and individual grammar becomes more and more abstract. However, this does not prevent established units from being used together with applications of more abstract schemata as well. The computability of a structure does not mean that it *must* be computed every time it is used (Langacker 1988b: 129). This may seem to be in contradiction with the idea of describing morphosyntactic complexity through the detailed syntactic and morphological description of each phrase in each utterance. However, as long as there is no access to subjects' individual grammars, utterance structures must be described using the available devices. This is why I separate choice structure from output structure and focus on the latter, being aware of the fact that it does not necessarily coincide with the choice structure, that is, the actual grammatical units of individual subjects.

3 QUANTITATIVE METHODS

In this chapter the methods of quantitative analysis in this study and their purpose are discussed in detail. To evaluate morphosyntactic complexity in a traditional way two different quantitative methods are used. Mean Length of Utterance (MLU) generates the average number of morphemes in utterances, while the Index of Productive Syntax (IPSyn) focuses especially on morphosyntactic structures and gives a more analytic picture of utterances than MLU. However, the first stage of analysis is actually a morphological coding of the transcribed data, since the final sample selection as well as the results obtained from both quantitative scales of measurement are entirely based on it.

The other main concern in this chapter is to provide a discussion of whether the methods are in line with the absolute approach to complexity in general.

3.1 Morphological coding

3.1.1 General principles

Before analysis was possible using either MLU or IPSyn the transcribed data had to be coded morphologically. In the process of morphological coding stems and productive suffixes in children's utterances were analysed and tagged following the coding conventions established in CHILDES (MacWhinney 1991). A child's actual language production was transcribed on main tiers utterance by utterance. During the coding procedure, a morphological tier (%mor) was attached to every main tier which contained an intelligible utterance. Each entry on the morphological tier of the transcript contained a code for a word class, stem, and identification of suffixes, as shown in Example 1.

Example 1

*CHI: vaihetaaNko tuolit ?
%mor: v|vaihta-pass-clit:int:ko n|tuoli-pl .
%eng : shall we change the chairs? (vaihta-=change; tuoli=chair)

The basic conventions were specified in terms of the following principles, based on the quality of the data and the requirements of systematic coding in CHILDES taking into account the special properties of Finnish.

Only the final form of an utterance was coded morphologically. In practice this means that repetitions of a word and other trials before the final word form were ignored. In cases when an utterance was broken off and started all over again only the last version was coded.

One of the principles in CHILDES coding is the highly systematic notation in order to make the use of different analytical programs as effective as possible. For this reason morphophonological changes in stems were neglected, and invariant stem notation was selected instead, even though this procedure conflicts with the Finnish inflectional system, which includes morphophonological changes in the stem in terms of consonant gradation and vowel changes. In coding nouns, adjectives, pronouns and numerals, the nominative case was selected as the constant stem, because it is the uninflected basic form in these word classes. In finite verb coding, the so-called strong vowel stem was chosen. This stem type is used in the 3rd person singular (*juokse-e* 'run-s', *istu-u* 'sit-s'), and in monosyllabic verbs it equals the form of the 3rd person singular (*syö* 'eats', *juo* 'drinks'). It is also considered to be the psychological basic form of Finnish verbs (Karlsson 1983: 207–209), rather than the 1st infinitive (*juosta* 'to run', *istua* 'to sit', *syödä* 'to eat', *juoda* 'to drink') which is used as the basic form in Finnish dictionaries.

There were three major exceptions in the notation of stems. In the case of mass nouns, the partitive case was taken as the stem. In Finnish, mass nouns are also used in the partitive in situations where other nouns would occur in the nominative case (e.g. in naming situations). The view that the partitive form is the real basic form of mass nouns is supported by examples from child language such as *vet-tä-ä* 'water-PARTIT- PARTIT' (Laalo 1998: 362) and *voi-a-ssa* 'butter-PARTIT-INESS' (Nieminen 1999: 192), where the partitive form is inflected again in the partitive or in some other case. Such multiple case inflections strongly contradict Finnish inflectional rules, which allow only one case at a time in word forms.

Another exception in the use of stable stems related to verbs in the 3rd infinitive (also referred to as MA-infinitive in ISK), which are inflected in the inessive (*tekemä**ssä*** 'in doing something'), elative (*tekemä**stä*** 'from doing something'), illative (*tekemä**än*** 'to do something'), adessive (*tekemä**llä*** 'by doing something') or abessive (*tekemä**ttä*** 'without doing something'). Without the case inflection, the 3rd infinitive appears only in compound words (***istuma**paikka* 'a place for sitting, sit-3INF-place', ***nukkuma**paikka* 'a place for sleeping, sleep-3INF-place', ***syömä**puikko* 'chopstick', literally 'a stick for eating, eat-3INF-stick'). Thus, the marker of the 3rd infinitive (*-mA*) is likely to be taken by children to be a part of the stem and not a separate suffix. This seems logical also because the cases that are most frequently used together with the 3rd infinitive (inessive, e.g. *syömä**ssä*** 'eat-3INF-INESS', and illative *nukkuma**an*** 'sleep-3INF-ILL') have already become familiar in noun inflection. Therefore in the morphological coding of children's utterances stems of the 3rd infinitive forms include the 3rd infinitive marker *–mA* as well. In Example 2, the three consecutive utterances produced by a Finnish boy (age 3;3.20) give further support to this view. In this example the 3rd infinitive form is treated as if it was a noun referring to a place for swimming (the most likely interpretation according to the context). The illative case ending which is present in the first utterance is eliminated in the second and third utterances to make a "nominative" form, which is ungrammatical (marked with *) in Finnish:

Example 2

Tästä mennään ***uimaan****.*
Here-ELAT go-PASS swim-3INF-ILL.
'We are going swimming from here.'

*Onko tässä *****uima****?*
Is-CLIT:INT here-INESS *swim-3INF?
Is there *a swim-3INF in here?

*Onko äiti etelässä *****uima****?*
Is-CLIT:INTmother south-INESS *swim-3 INF?
Mother, is there *a swim-3 INF in the south?'

The third exception to the stem rule was morphologically unanalysed wholes. A word form was counted as one single morpheme in the following situations: (i) an utterance was an imitation of the preceding utterance by another participant in the conversation, (ii) an utterance was a rhyme or a song, (iii) a word form included a suffix which was not defined as productive in a child's language because of insufficient evidence of productivity (see principles for determining productivity in section 3.1.2 below), and (iv) a word was a frozen form representing an archaic and no longer productive use of a case (e.g. locative adverbs *koto**na*** 'at home' and *ulko**na*** 'outside',with essive suffix *-nA*). The exceptions to the basic principle of the unvaried stems in these three special cases was a way to ensure that children were not credited for unproductive morphemes.

3.1.2 The principles for determining the productivity of morphemes

Since only productive morphemes were coded separately, principles for defining the productivity of morphemes were needed. Most of the solutions in defining productivity of morphemes are by-products of studies using MLU as a method of evaluation (see e.g. Brown 1973/1976: 78; Dromi & Berman 1982: 410–414; Hickey 1991: 560; Thordardottir & Weismer 1998: 15–16; Toivainen 1980: 24). To avoid overestimation of a child's morphological skills, every grammatical morpheme had to meet the criteria concerning the occurrence of a morpheme. Otherwise it was considered to be unproductive and was ignored in coding. In brief, the productivity of morphemes was defined morpheme by morpheme and child by child.

In this study a suffix was defined as productive in a child's language if it occurred in two different contexts. Only clear and intelligible tokens were taken into account. Moreover immediate imitations, songs or rhymes were not accepted as evidence of productivity. Colloquial representations of the morphemes were as acceptable as standard Finnish representations.

The principle that only two distinct occurrences of a morpheme were required was adopted mainly on the basis of the quality of the data. Due to the shortness of the language samples, requiring three occurrences would have underestimated children's skills. However, nor was it intended to take it for granted that every morpheme was productively used by a child, and thus two different occurrences was taken to be a requirement for productivity. Exceptions to and limitations on the basic rule are discussed in detail in the following chapter.

3.1.3 Exceptions and restrictions in determining the productivity of morphemes

In determining the productivity of case suffixes, some clarifications and restrictions should be noted. A case suffix was regarded as productive only if it was used at least twice to inflect nouns, adjectives, pronouns or numerals. This ruled out the use of case suffixes in infinitive and participle forms of verbs and in demonstrative adverbs. Jorma Toivainen (1980: 95–101, 1990: 41) has noted that demonstrative adverbs (e.g. *tässä* 'in here', *tästä* 'from here', *tähän* 'to here', *täällä* 'here', *täältä* 'from here' etc.) are regularly acquired before the corresponding local cases of nouns. At least in the beginning, these adverbs are the only word forms representing the inflection paradigm of local cases. Thus, it is very likely that they are acquired as indivisible wholes. For this reason the demonstrative adverbs were not regarded as evidence of the productivity of local cases. These restrictions on what counted as acceptable occurrences tightened up the basic productivity rule remarkably and made it more valid.

Similarly a restriction was applied in determining the productivity of suffixes marking person in verbs. The verb inflection for person was considered productive only if a child used different verbs inflected for at least two distinct persons other than the 3rd person singular or plural, which are similar forms in colloquial Finnish (Kielinen 1997; Mielikäinen 1981, 1984). The 3rd person forms were excluded because in many verbs they are similar to the strong vowel stem, which is the basic form of a verb. The passive forms were accepted in defining the productivity of person inflection, since in colloquial Finnish the 1st person plural is expressed almost exclusively with passive forms (see Manninen 1997; Mielikäinen 1991: 50–51) instead of the original ending *-mme*, the use of which is nowadays reduced practically to official and literary purposes. However, the different forms of the verb *on* 'to be' were not accepted as evidence of productive person inflection of verbs due to the irregular inflectional paradigm of the verb.

There were some additional rules defining the productivity of the partitive and accusative cases. The partitive case was taken as the basic form of mass nouns, and thus productivity in this case had to be confirmed by occurrences in other nouns, adjectives, pronouns and numerals. The accusative case differs from other Finnish cases, because it is syntactic by nature and occurs only in the object position. The accusative has a suffix of its own only in personal pronouns (*minu**t**, sinu**t**, häne**t**, meidä**t**, teidä**t**, heidä**t***) and the interrogative pronoun *kene**t*** 'who-ACC', but in all other occasions the inflection is similar to that of the nominative or the genitive. (Vilkuna 1996: 118.) In these cases it is position in the sentence alone that distinguishes the nominative or genitive forms from the accusative. Since the purpose of the coding was to tag morphological forms, and not syntactic functions, occurrences of the accusative case were categorized according to their actual form as representing the nominative or genitive case and only in the case of personal pronouns as accusative forms. This is consistent with the decision that the utterance was chosen as the basic unit in this study, rather than the sentence or the clause.

It is not always clear which part of speech a word represents. Earlier, in section 3.1.1, there was an example of the 3rd infinitive form of a verb, which was used by a child as if it was a noun. Not only in a child's language, but also in language in general, the boundaries between parts of speech may be opaque. In Finnish, as in most Uralic languages, negation is expressed with a negative auxiliary (Finnish *ei*) together with the invariant negative form of the main verb. Both person and number are expressed in the negative auxiliary: *en halua* 'I don't want', *et halua* 'you don't want', *ei halua* 'he/she/it doesn't want', *emme halua* 'we don't want' (in colloquial Finnish usually in passive *me ei haluta*) *ette halua* 'you don't want' and *eivät halua* 'they don't want'. However, there is still another way to use this negative auxiliary, not as an inflecting verb but as a particle. (Savijärvi 1977: 11–26.) To distinguish these two different usages presented in Example 3, only those occurrences of *ei* that included clear evidence of person were coded as inflected verbs. On other occasions *ei* was treated as a particle.

Example 3

Ei, **et** saa mennä.
ptl|ei v:neg|ei-2S v|saa&NEG[3] v|mene-1INF
No, you are not allowed to go.

In earlier studies concerned with MLU the problem of defining a morpheme has become an essential topic. This was already briefly touched on when the

3 This is an example of morphological coding where there is no separable bound morpheme attached to a stem but a word form has a particular function. The function is expressed by "&" and a code referring to the function, in this case negation NEG. Tagging the lexical verb with NEG is justified, because in conversation it can also be used to express negation without the negative auxiliary (Savijärvi 1977: 28).

problem of the 3rd infinitive was discussed in section 3.1.1. The conclusion was that the marker of the 3rd infinitive (*-mA*) should not count as an individual morpheme but should be included in the stem instead. In passive forms there are also two suffixes attached to a verb stem, a passive marker (*-tA*, *-lA*, *-nA*, *-rA*) and a vowel lengthening + *n* (*Vn*), historically known as a personal ending (e. g. *istu-ta-an* 'sit-PASS-personal ending') (Hakulinen 1979: 240). These suffixes form a solid monosyllabic unit which can be categorized as one individual morpheme on the basis of both form and meaning. Jorma Toivainen (1980: 24) has come to a similar conclusion in his study on how children acquire Finnish morphology. Thus, the passive marker and personal ending were not separated in the coding, and they were tagged with a passive code only.

Every intelligible utterance produced by a child was coded according to the principles described above. All productive, clearly individual morphemes were separated with a hyphen (e.g. *tuolit* → n|tuoli-PL, 'chair-PL') and the elements of fusional inflection (changes in the stem as the only sign of inflection, or two morphemes fused together) were set apart with an ampersand (e.g. [*on*] *leikattu* → v|leikkaa-PP&PASS, '[is] cut', *-ttu* marking both participle and passive voice). The exact codes for parts of speech and grammatical morphemes used in the morphological coding are presented in Appendix 1.

3.1.4 The final selection of the data sample

The final data selection was made after the morphological coding was complete. The selection observed the requirements set down by the methodology of both MLU and IPSyn (see Brown 1973/1976: 78; Scarborough 1990a) for the number of utterances to be included in the data. The normal procedure in both methods is to analyse 100 intelligible, consecutive utterances. However, the purpose of this study was to investigate the complexity of morphosyntactic structures. Thus, the focus was on the most complex and advanced performance of each child rather than on overall skills, and data selection had to be in line with this purpose. Therefore it was decided to use the longest utterances only, because this procedure would maximize the number of multimorphemic utterances to be included in analysis. For the same reason it was possible to reduce the number of utterances from 100 to 80.

In this study 80 intelligible utterances with the largest number of morphemes were selected from the coded data base of each child. Only those morphemes that were both formally and semantically separable (in other words, separated with a hyphen) were taken into account in selecting the utterances. Imitated utterances, partially unintelligible utterances, rhymes and songs were excluded. The selection was made using the MAXWD program, one of the analytical programs in CHILDES. MAXWD tracks a specified number of the longest utterances in data files in CHILDES, and prints them in the order of length. With this procedure the selection emphasized multimorphemic utterances. The data samples were then analysed quantitatively using MLU and IPSyn. These methods are introduced in detail in the following sections.

3.2 Mean Length of Utterance (MLU)

3.2.1 General background

The idea of calculating the mean length of utterances was first introduced in its present form by Roger Brown (1973/1976). MLU serves as an index of the grammatical development of a child. It is a well known fact that age is not a good indicator of children's grammatical development, especially during the first years of language acquisition. Average utterance length in morphemes seems to reflect the structure of a child's utterances in an informative way. Brown argued that "almost every new kind of knowledge increases length: the number of semantic roles expressed in a sentence, the addition of obligatory morphemes, coding modulations of meaning, the addition of negative forms and auxiliaries used in interrogative and negative modalities, and of course, embedding and coordinating" (ibid., 77). With an index such as MLU the linguistic abilities of two or more children are easier to compare. Children who are matched for MLU are more likely to have the same level of constructional complexity than children of the same chronological age. Even though MLU can be used cross-linguistically, the characteristics of different languages in all likelihood affect the results (Hickey 1991; Schnell de Acedo 1994: 250–252). Therefore MLU is a tool for comparison predominantly within a language and not between languages.

The basic method of counting MLU is a very simple one: 100 intelligible utterances produced by a child are analysed in terms of productivity and then the mean number of morphemes per utterance is calculated. However, MLU is not a solid and uniform method. Basically, the idea of counting units and calculating the mean of these units under certain conditions is the only unchanging factor. In practice, it is possible to change everything else. The basic unit can be the morpheme (e.g. Brown 1973/1976), the word (e.g. Hickey 1991), the content word (Rollins, Snow & Willet 1996), the clause (Cheung & Kemper 1992) or even the syllable (Hickey 1991). The selection of utterances included in the calculation can be limited by the number of utterances (e.g. Brown 1973/1976), the recording time (Rollins et al 1996), certain pragmatic factors (Rescorla, Dahlsgaard & Roberts 2000; Johnson 2001) or utterance length (Klee & Fizgerald 1982). The rationale behind using different ways of measuring average length of utterance is that researchers have been looking for either a more reliable and easier or a more suitable way of calculating MLU than was originally suggested by Brown. They have tried to avoid the occasional arbitrariness and ad hoc decisions almost unavoidable in counting morphemes.

Since Brown's research was published in 1973 Mean Length of Utterance has become a widely used index in studies of child language acquisition. It can be considered a research result in itself, describing the level of grammatical development of children as in the study on late talkers by Rescorla, Dahlsgaard and Roberts (2000). In many studies it is a tool of comparison both between and within groups. For example, in Scarborough's study (1990b) dyslectic and non-dyslectic children are compared by using MLU. It has also been used in matching subjects from different groups, as when SLI children are matched with normally developing children before comparing their linguistic abilities more precisely (e.g. Dromi, Leonard & Shteiman 1993; Grela & Leonard 2000).

Mean Length of Utterance as a method has been investigated in several studies. The purpose of these studies has been to identify the factors influencing MLU (e.g. Rollins et al 1996; Johnson 2001) and to investigate how suitable MLU is in evaluating language in normal and language disordered groups (e.g. Scarborough et al 1991) or what kind of correlation there is between MLU and age (e.g. Scarborough, Wyckoff & Davidson 1986; Klee, Schaffer, May, Membrino & Mogey 1989) or MLU and grammatical complexity (e.g. Klee & Fitzgerald 1985; Scarborough et al 1991). In addition, MLU has been compared in several studies with other scales of measurement evaluating syntactic complexity (e.g. Cheung & Kemper 1992; Blake et al 1993; Kemper et al 1995). It has also been a matter of great interest from the point of view of languages other than English, such as Hebrew (Dromi & Berman 1982), Irish (Hickey 1991), Icelandic (Thordardottir & Weismer 1998), Inuktitut (Allen & Crago 1996), and Finnish (Toivainen 1980), to name but a few.

Even though Mean Length of Utterance is probably the most widely used measure of early grammatical development it has been much criticised from the very beginning. In fact, one result of the criticism on MLU is the great variation of different procedures used in calculating it. By proposing new ways of calculation researchers have tried to avoid or at least diminish the flaws pointed out by critics. One of the most criticised factors in Mean Length of Utterance is the original unit of calculation, in other words the concept of a morpheme. It is quite difficult to decide precisely what is a morpheme in a child's language (e.g. Dromi & Berman 1982; Rollins et al. 1996). The original goal was that only productive morphemes should be calculated, but especially on the basis of short speech samples determination of the productivity of morphemes is fairly arbitrary. It has even been argued that MLU is not sensitive to factors other than the number of morphemes (or other units of measurement). Even though Brown (1973/1976) stated in his frequently quoted observations (p. 77) that "almost every new kind of knowledge increases length", he also admited that beyond Stage V (MLU > 4.0) a child "is able to make constructions of such great variety that *what* he happens to say and the MLU of the sample begin to depend more on the character of the interaction than on what the child knows, and so the index loses its value as an indicator of grammatical knowledge" (ibid., 77).

Dromi and Berman (1982) pointed out that increased complexity does not necessarily require increase in length and that length alone cannot be taken as the criterion of linguistic sophistication. Consistent with Dromi and Berman, also Lahey, Liebergott, Chesnick, Menyuk and Adams (1992) noticed that, despite obvious advances in language development, MLU seemed to rise and fall, and they concluded that certain aspects of learning actually result in shorter utterances and thus in reduced MLU. Several opinions have been expressed as to the range within which Mean Length of Utterance is a valid indicator of grammatical development. Roger Brown (1973/1976: 221) limited the validity of MLU to values under 4.0. After him Klee and Fitzgerald (1985) claimed that MLU cannot be reliably used when it exceeds 2.5. Rondal et al. (1987) as well as Scarborough et al. (1991) argued that the critical point is at MLU 3.0. The highest limit for the reliability of MLU (4.5) is set by Blake et al. (1993). All these limiting values, including the highest one, underline the fact that Mean Length of Utterance is a feasible tool for evaluation of linguistic sophistication only during the earliest phases of expressive language. This narrow range of applicability has been one of the reasons why MLU has been criticised (Rollins et al. 1996).

Mean Length of Utterance is thus a rather problematic method for measuring grammatical development, with both many advantages and clear flaws. Being a widely used measure, it deserves methodological evaluation in comparison with other methods and, within the framework of the present study, a discussion of its relationship to a multidimensional view of the morphosyntactic complexity of language.

3.2.2 The calculation of Mean Length of Utterance in this study

In this study, Mean Length of Utterance in productive morphemes is calculated on the basis of the 80 longest intelligible utterances produced by each child. The procedures of morphological coding and final data selection were described in detail in section 3.1.4. The data was not limited by excluding certain types of utterances, such as elliptical answers to questions, since this type of interaction is not necessarily less complex linguistically than a child's grammatically complete utterances.

In the present study MLU is used for two purposes, both to analyse children's language production and in order to evaluate it as a method. First of all it is utilized in a traditional way, to generate results which can be used to compare different children. A traditional quantitative comparison is made between the risk group and control group with the help of MLU and IPSyn to find out if there are any differences between the groups. However, no further interpretations are made concerning the morphosyntactic abilities of the subjects on the basis of the MLU results. Then the MLU results are compared with IPSyn points in order to investigate the relationship between these two methods, thus partly analysing the actual data itself and partly analysing MLU as a method. Finally, the MLU results are evaluated by comparing them with a multidimensional analysis of utterances (see Chapter 5).

3.3 The Index of Productive Syntax (IPSyn)

3.3.1 General background

The other quantitative method utilized in this study, the Index of Productive Syntax (IPSyn), was originally created by Hollis S. Scarborough (1990a). The basic idea of the index is to analyse which essential syntactic structures and morphological forms are realized in 100 intelligible utterances produced by a child. Points are given as follows: zero points for structures and forms not produced, one point for one occurrence of a structure or a form and two points for at least two different occurrences of a structure or a form. Through this procedure emerging productivity is investigated and the grammatical complexity of a child's expressive language is assessed. (Scarborough 1990a.) Thus, IPSyn is a measure of morphological and syntactic types rather than tokens.

By using IPSyn it is possible to achieve in practice three different kinds of results. Firstly, the main result is given by the number of total points. However, the original version of IPSyn is divided into four sections: noun phrases, verb phrases, questions and negation, and sentence structures. This division enables a second type of result, namely sub-scores for each section separately, and children may be compared in these more specific terms. The third possible type of result is of a more qualitative nature. The realizations of each structure and form up to two occurrences are written down on the notation sheet during the analysis and this enables further qualitative analysis on the basis of a child's utterances. (Scarborough 1990a.) IPSyn was first used in the field of dyslexia studies, but lately it has also been applied in evaluating the linguistic production of late talkers (e.g. Rescorla et al 2000; Rescorla et al 2001), children with cochlear implant and hearing aids (Tomblin, Spencer, Flock, Tyler & Gantz 1999), and SLI children (Hadley 1998; Oetting, Cantrell & Horohov 1999).

The selection of relevant and essential forms and structures accepted in IPSyn is language specific. Moreover, the number of items may vary a lot from language to language. For these reasons a separate version of IPSyn is needed for each language and the results are not cross-linguistically comparable. As a consequence of language specificity, it is possible that different versions of IPSyn may behave in a slightly different ways. For example, sensitivity to the length of utterances may vary together with the length of structures accepted in IPSyn. In addition, emphasis on morphology, syntax, and morphosyntax may vary according to the properties of the language in question.

The Index of Productive Syntax has not become as popular and widely used a method as MLU. This is understandable, because utilizing it requires analysing data utterance by utterance and the procedure is not automated in the way that MLU is, for example, in CHILDES. Unlike MLU, there have been only a few studies considering the methodological characteristics of IPSyn. In studies by Kemper et al (1995) and Cheung and Kemper (1992), IPSyn was one of the measures of complexity used in assessing data from children aged 5–10 and adults aged 60–90. It is not surprising that in neither of these studies was IPSyn found able to detect differences between the age groups, since it was designed to analyse language acquisition data collected from preschool children only. However, there is evidence of IPSyn being suitable for analysing structural complexity in language disordered groups representing a wider age range (Scarborough et al 1991).

So far, there has not been any strong criticism of IPSyn in the literature. However, the definition of complexity as either morphological or syntactic in nature is somewhat restricting. Utterances produced by, for example, a 30-month-old child can consist of several forms and structures, but each of them is evaluated separately and thus, the interplay between morphology and syntax and the combination of several structures within an utterance are ignored (Nieminen & Torvelainen 2003: 128). Despite this defect, Index of Productive Syntax gives a fairly broad insight into the morphosyntactic skills of a child and functions as a complement to Mean Length of Utterance for research purposes.

3.3.2 The Finnish version of IPSyn

For the purposes of the present study, a Finnish version of IPSyn was created by Nieminen and Torvelainen (2003). Since IPSyn is by nature a very language specific method, only the basic idea of it was adopted while the specific contents and division into different sections were created on the basis of the Finnish language. To select the essential structures and forms relevant to Finnish child language, spontaneous speech samples produced in a play session by 22 children aged 24 and 30 months were analysed syntactically and morphologically in detail. 12 of the children were categorized as beginners in expressive language, since their MLU values were below 2.000. The remaining 10 children were linguistically more advanced performers (MLU 3.416–5.181). In addition to this data analysis, the results in Bowerman's (1973), Lieko's (1992) and Toivainen's (1980) studies were also consulted in making the choice of structural items. (Nieminen & Torvelainen 2003: 122.)

The selection of structural items was based mainly on two principles: differentiability and stepwise development. The selection of core items thus consists of structures which were found to differentiate between beginners and advanced performers as well as between individuals within each group. The developmental aspect was addressed by creating steps from a simple item to a more complex one. A good example of the realization of these principles is the 3rd infinitive. The advanced children already used structures where the 3rd infinitive form occurred together with a finite verb, but the beginners did not. To track individual development, three distinctive possibilities were noted in IPSyn. The first level was the 3rd infinitive alone, the second was realized by the 3rd infinitive together with a finite verb, and the third level was reached with a structure including three different verbs: a finite verb form, the 1st infinitive form, and the 3rd infinitive form. (See V12–V14 in Appendix 2.)

3rd INF	*Syömään*	'Eat-3INF-ILL'
V + 3rd INF	*Tule syömään*	'Come and eat'
V + 1st INF + 3rd INF	*Haluan tulla syömään*	'I want to come and eat'

In the Finnish version of IPSyn (Appendix 2), there are three sections consisting of 9 nominal phrases (made up of nouns, pronouns, adjectives and numerals), 16 verb constructions and 21 sentence structures. In addition to these specified structures, each IPSyn section includes an entry for two unspecified structures. This enables the evaluator to give points for structures that he finds worth mentioning, for example occurrences of rare cases such as translative or complex verb chains including the negative auxiliary. The maximum score in the Finnish version of IPSyn is 98 points, with 20, 34 and 44 points for the nominal phrase, verb construction and sentence structure sections respectively. Since Finnish has a rich inflectional morphology it is natural that the Finnish IPSyn stresses morphological forms more than Scarborough's original English IPSyn. Cases in nominal phrases as well as inflections for person, tense and mood in verb constructions are well represented, and they clearly turn IPSyn into a morphosyntactic scale of measurement.

The original version of IPSyn devised by Scarborough was divided into four sections: noun phrases, verb phrases, sentence structures and interrogative and negative structures. In the Finnish version the first three sections are retained but that for questions and negations is omitted. The omission of this section does not, however, mean the omission of questions and negative expressions from the analysis, since instead of having a section to themselves these utterances are placed within the other categories in IPSyn. There were two main reasons for excluding a specific section for questions and negations. First, asking questions is not always an appropriate way of communicating in every situation, and it does not seem fair to require a child to formulate questions. The number and quality of questions is more dependent on context than on the existence of morphosyntactically defined structures in a child's repertoire. Therefore questions are analysed as if they were declarative sentences, ignoring their pragmatic and functional aspect. In the case of negation, its omission is based more on grammatical reasons. In Finnish, as in most Uralic languages, negation is expressed with an inflected negative auxiliary plus the negative form of the lexical verb (Savijärvi 1977: 11). Thus, it is appropriate to place such expressions in the same section as other verb constructions, rather than separating them into a different section. (For more details on adapting IPSyn in Finnish, see Nieminen & Torvelainen 2003.)

In the present study the IPSyn results are based on the 80 longest utterances from each child. The IPSyn results are used in two ways, comparable to the utilization of MLU values described in 3.2.2. First, they are used to compare children in the risk and control groups to see whether there are any differences. Since IPSyn gives two kinds of results (total points and separate points for each section) the comparison between the groups is made using both. Another way the IPSyn results are used in this study is to compare them with

MLU values in order to analyse the relationship between these two scales of measurement. Like MLU, IPSyn will also be subject to evaluation as a method.

3.4 The view of complexity in MLU and IPSyn

3.4.1 Effective complexity

Mean Length of Utterance and the Index of Productive Syntax are both described as measures of complexity. MLU is connected with linguistic complexity in general whereas IPSyn is defined as a measure of syntactic complexity in particular (or morphosyntactic complexity in the case of Finnish). Both methods evaluate spontaneous speech production and they represent an absolute approach since the object of analysis is the structural complexity of the actual language produced and not the complexity or difficulty experienced by a language user [4]. In this section, and the following, the two scales of measurement are first evaluated in terms of a theoretical view of complexity and then in terms of the ideas about morphosyntactic complexity presented in Chapter 2. This discussion will give some additional perspective on methodology and therefore help to answer the following research question:

> What kinds of information about structural complexity in utterances do commonly used measuring scales like MLU and IPSyn give?

In addition to the general properties in common mentioned above, MLU and IPSyn have a shared philosophy, that of effective complexity (Gell-Mann 1994: 50, 55–56; see also Dahl 2004: 24), which has its roots in an information theoretical approach to complexity. Effective complexity refers to sets of regularities that are first acquired from environmental information and then condensed into a schema by complex adaptive systems like human beings. The longer the condensed schema, the more complex is the phenomenon in question.

Both MLU and IPSyn identify sets of regularities in utterances. MLU evaluates a realisation of a linguistic system, which is seen as a repetition of one single scheme, a morpheme. The task of MLU is to count these repetitive occurrences within the limits of utterances and then calculate the average number of morphemes. Thus the complexity caught by MLU refers to how many repetitions of the schema on average are combined in an utterance. In the case of IPSyn the goal is to investigate how many schemata a child's linguistic system consists of and how many of these schemata have varying realizations in language production. The schemata themselves are pre-specified

4 The use of language production as data does not automatically mean that the approach to complexity is an absolute one. Language production can be analyzed from a relative perspective as well, for example when what is evaluated is the kind of difficulties the structures produced may cause to different types of language users.

morphosyntactic structures which are chosen to represent the whole linguistic system taking into account their role in the language acquisition process.

MLU and IPSyn do not represent effective complexity in its purest form even though they both look for patterns (morphemes and given structures) and result in figures that refer to the occurrence of those patterns in an object. Both indices are developed somewhat further. In MLU the final result is generated from “raw” results by calculating the average value, and in the case of IPSyn the occurrence of each structure is confirmed by another occurrence whenever possible, in order to suggest achieved or emerging productivity of the structure in question. The ways of evaluating effective complexity in MLU and IPSyn are illustrated in Figures 4 and 5.

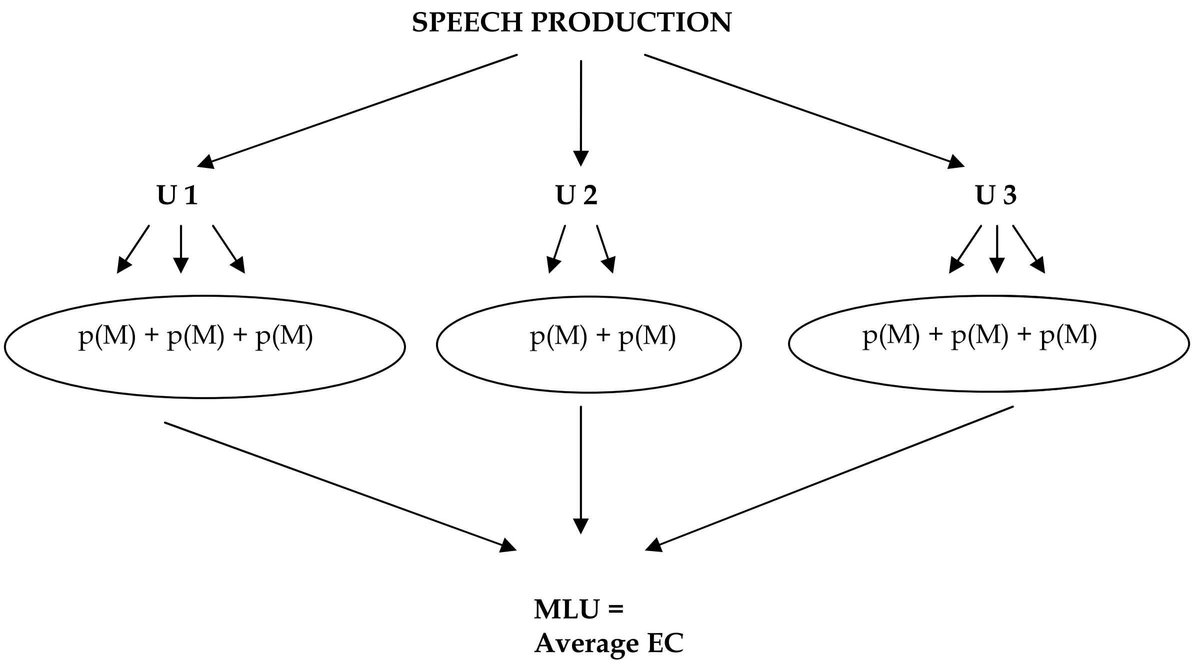


FIGURE 4 The evaluation of effective complexity (EC) by MLU. Speech production is divided into utterances and each utterance (U1, U2…) is analysed in terms of a pattern (p), which is specified as a morpheme (M); the effective complexity of each utterance is the number of manifestations of this pattern (morpheme). MLU compresses EC into an average value.

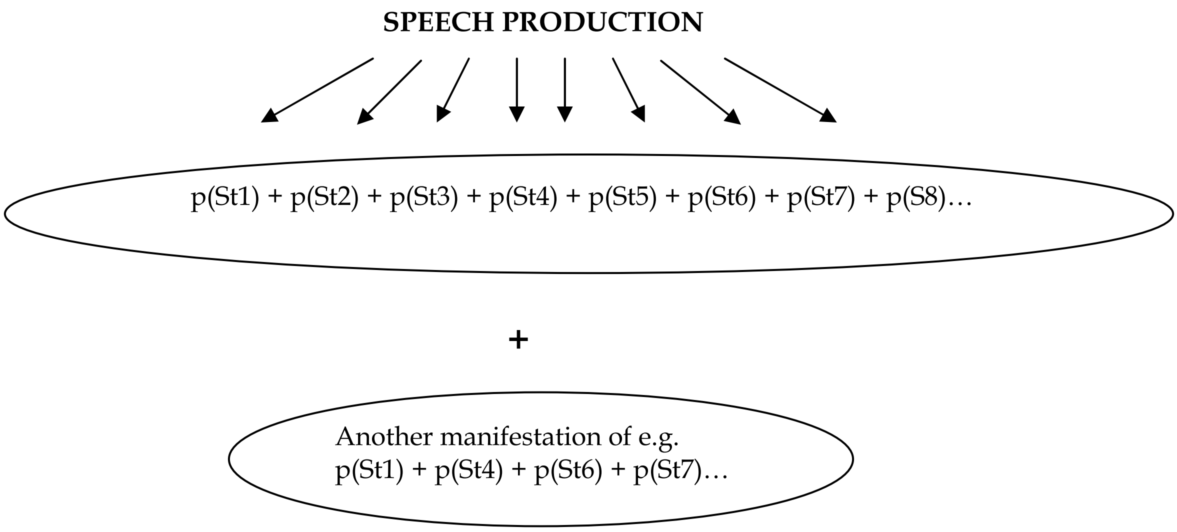


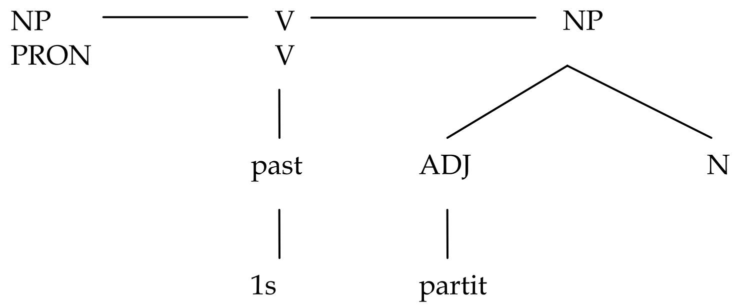
FIGURE 5 The evaluation of effective complexity (EC) by IPSyn. A child's entire speech production in a given situation is analysed in the search for manifestations of patterns (p) which are pre-specified structural items (St1, St2…). The effective complexity of a child's structural system is actually the number of patterns found in his/her speech. However, in IPSyn the result is calculated according to how often structural items occur at lest twice and in a different realization.

3.4.2 MLU, IPSyn and the structural complexity

In Chapter 2, the multidimensional and multilayered morphosyntactic structure of verbal productions was discussed and structural complexity was defined against this background. It was claimed that it is possible to describe syntactic complexity on two levels. First, there is the composition of different phrases within an utterance, and second, the composition of the phrases themselves. Another dimension to this already complex picture is added by morphology, which is used as a means of elaborating phrases as well.

When a complexity view of MLU is compared with a more multidimensional model, there are several aspects that differentiate them from each other. First, MLU focuses on morphemes only, which means that the syntactic properties of productions are ignored. Secondly, morphemes are not divided into subcategories (e.g. bound vs. unbound morphemes). Instead, they are treated as equal representatives of the basic pattern. Thus, morphological and syntactic relations are ignored, any layered structure vanishes, and each utterance is described as merely a linear string of patterns (Figure 6). Furthermore, MLU calculation makes flattening cumulative, because the final value is an average of already flattened strings. However, a compensating factor is that MLU aims to give a general view of each utterance as a whole.

IPSyn approaches linguistic productions in a more analytic way than MLU. It concentrates on morphosyntactic structures, which automatically means that it is dealing with relations between linguistic elements.



Ma	ost-i-n	uut-ta	maitoa.
I	buy-PAST-1S	fresh-PARTIT	milk.
‘I	bought	some fresh	milk.’

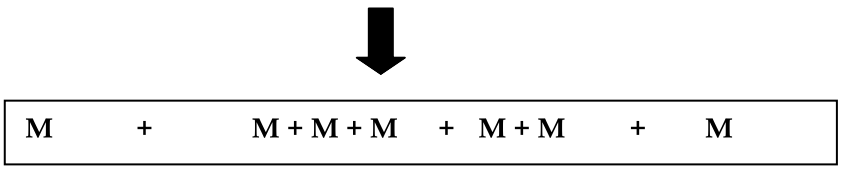


FIGURE 6 Structural complexity of an utterance as treated by MLU. A multilayered structure becomes a linear string of morphemes (M). (*Maitoa* ‘milk’is formally a inflected form maito-PARTIT, but since it is a mass noun, the partitive form is considered to be the psychological basic form of the word, so it is not counted as having an inflection.) (16 Kyösti)

Those are found on several levels: between a stem and a suffix, between a head word and a modifier, between a preposition or postposition and a main word, and between main syntactic elements, etc. IPSyn registers a two-word nominal phrase, a verb in the past tense, and the syntactic structure S + V + O, but it does not record that the two-word NP, is actually in object position in the S + V + O structure, right after the verb in the past tense (Figure 7). Thus, IPSyn is a good index to the building blocks in a linguistic production but it does not put them together to draw a coherent picture of the whole utterance Instead, it forms a compilation of the grammatical resources that can be found in the output structure of utterances.

The differences between MLU and IPSyn revealed so far strongly suggest that they do not measure the same properties of linguistic production. This is not in line with the ideas put forward in some earlier studies (e.g. Scarborough et al 1991) where MLU and IPSyn have been used to prove one another’s validity, thus supposing them to be measures of the same properties of language. However, the differences can be turned into a strength if MLU and IPSyn are made to work in close collaboration. It is actually a manifestation of *divide et impera,* the ancient method of administration, IPSyn representing the dividing and MLU the controlling part. Comparisons between the two indices must be made child by child and not group by group, as has been done in earlier studies. It can be hypothesised that with the resources from IPSyn and the frames provided by MLU within which the resources are used, it is possible to get a more accurate picture of the structural complexity of utterances than by either of the methods alone. Such an analysis would also better conform to the model of multilayered structural complexity in utterances.

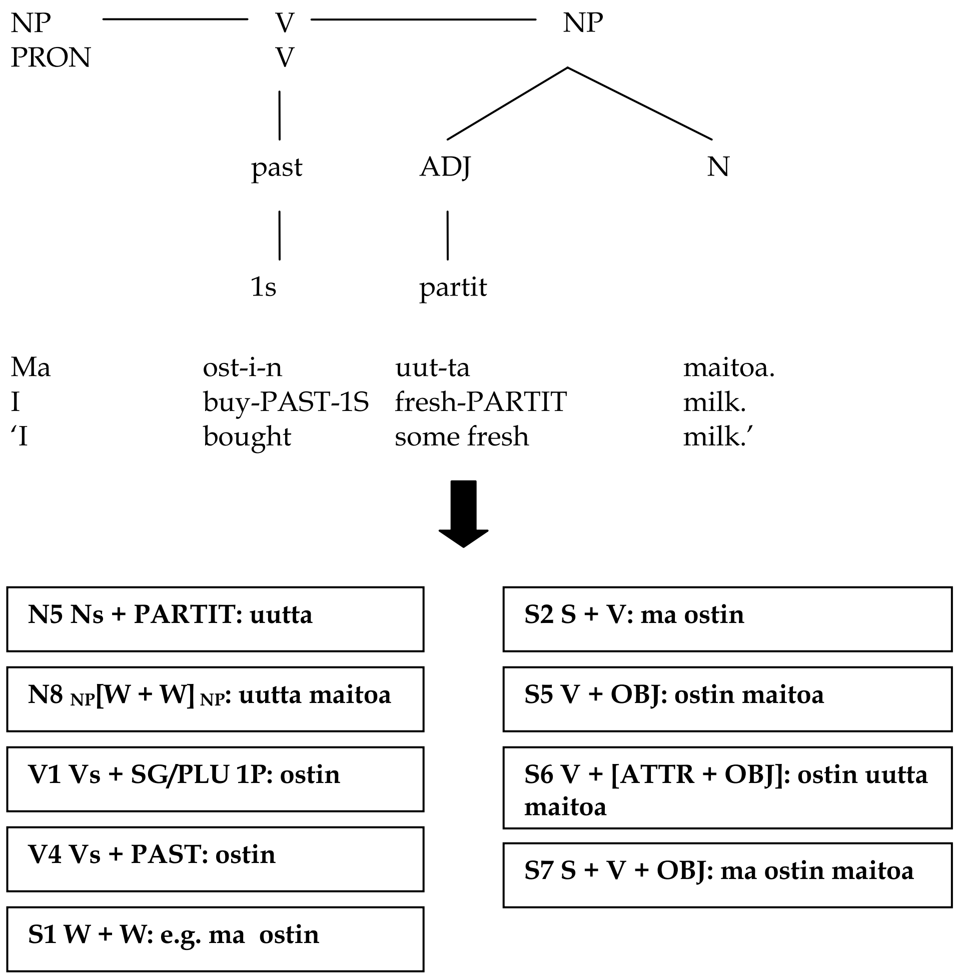


FIGURE 7 The structural complexity of an utterance as treated by IPSyn. An utterance is split into structural building blocks. N5, N8, V1, V4, S1, S2, S5, S6, and S7 refer to structural items in IPSyn (see Appendix 2).

4 EVALUATION OF STRUCTURAL COMPLEXITY USING TRADITIONAL METHODS

In this chapter, the results of the quantitative analysis are presented. Especially Mean Length of Utterance (MLU) but also the Index of Productive Syntax (IPSyn) are the methods of evaluation traditionally used in studies concerning complexity in child language. In the present study they are used to compare the risk group and the control group, but in addition to that, the MLU and IPSyn results are compared with each other. This is done to get a broader insight into the ongoing changes in complexity in the data. A foundation is laid for the later qualitative analysis on the basis of the interpretation of the quantitative results.

4.1 Research questions and the use of the quantitative methods

In Section 1.5, the three main research questions of the present study were presented. In this chapter, questions 2 and 3 are investigated more closely. The first focuses on potential differences between the risk and control groups:

> What is the relation between risk group children (children with high genetic risk of dyslexia) and control group children (children without high genetic risk of dyslexia) when special attention is paid to the complexity of morphosyntactic structures in their spontaneous speech production?

In this chapter answers to this question are sought by way of quantitative analyses. The groups are compared first on the basis of MLU values and total IPSyn points and then by analyzing the nominal phrase, verb construction, and sentence structure sections distinguished by IPSyn. The results are a tuned version of effective complexity of utterances (MLU, an average effective complexity) and morphosyntactic resources (IPSyn, effective complexity augmented with an estimation of the productivity of the structures). However, neither of the methods alone will analyze the multilayered structure described in Chapter 2, but they do function as a necessary step towards the comparison between the methods.

The other research question which will be discussed in this chapter reflects more the view of complexity taken in the present study. The purpose of comparing the results of the quantitative methods is to get more information about what aspects of structural complexity are probably undergoing changes on the basis of the combined body of results. This is the starting point in answering the following research question:

> What kind of information about structural complexity in utterances do commonly used measures like MLU and IPSyn give?

Theoretical discussion of this issue was already started in Section 3.4, and now the assumptions made there are tested in practice. MLU and IPSyn are considered to be two sides of the same phenomenon, and their complementarity may result in a new view of these methods as well as of the phenomena characterizing structural complexity. MLU and IPSyn results are compared with each other child by child. A possible discrepancy in results will make the comparison especially interesting, because each time one measure indicates growth of complexity and the other does not, we have clear evidence of the different focuses of the methods. At the same time, potential discrepancies indicate heterogeneous growth in structural complexity which is neither increase in utterance length nor additions to structural resources alone. There are other factors involved as well, such as structural combining in utterances. Finally, at this point some hypotheses are formed concerning the manifestation of morphosyntactic complexity on the basis of a comparison between MLU and IPSyn.

4.2 Group comparison using MLU (morphemes)

Mean length of utterance was calculated on the basis of the 80 longest utterances from each child. Exceptions to this were Liisa (ID 5) with only 60 available utterances and Lauri (ID 12) and Risto (ID 40) with 70 available utterances. The MLU scores arranged by subject, group, and gender are presented in Table 2. The scores range from 1.233 to 7.162 in the control group (CG) and from 1.587 to 7.862 in the risk group (RG) indicating great individual differences in both groups. However, the wide ranges are partially caused by the calculation procedure which ignores the shortest utterances. Children with a large number of utterances are those who benefit the most from this because, for example, all uni-morphemic utterances can be ignored since there are enough multimorphemic examples in their data. For the same reason most of the scores are higher than usual in MLU studies in general and at the age of 2;6 in particular. It is also worth noticing that the data represents a language with a rich bound morphology, and this language specific feature can be expected to affect the MLU results.

TABLE 2 The mean length of the 80 longest utterances from each subject both in the control group (CG) and the risk group (RG). * 70 utterances available. ** 60 utterances available. The purpose of the gray lines will be explained in connection with Table 9.

Control group					**Risk group**			
ID	**Name**	**Gender**	**MLU**		**ID**	**Name**	**Gender**	**MLU**
5**	Liisa	F	1.233		26	Tiina	F	1.587
12*	Lauri	M	2.614		25	Siiri	F	1.650
9	Lassi	M	3.175		40*	Risto	M	2.329
18	Saku	M	3.962		32	Juho	M	2.450
17	Tatu	M	4.063		33	Jukka	M	2.775
6	Ronja	F	4.100		37	Henri	M	3.638
13	Joel	M	4.713		35	Seppo	M	3.688
8	Aino	F	4.775		22	Anna	F	3.888
15	Mika	M	4.838		36	Sampo	M	3.987
14	Patrik	M	5.500		31	Aleksi	M	4.200
10	Santeri	M	5.713		34	Leo	M	4.263
20	Tuomas	M	5.862		29	Elisa	F	4.388
11	Jaakko	M	6.112		24	Sanni	F	4.963
19	Tuomo	M	6.237		38	Pekka	M	5.275
1	Tuija	F	6.250		21	Paula	F	5.400
3	Laura	F	6.250		23	Anniina	F	6.050
2	Riikka	F	6.287		27	Janna	F	6.300
7	Taru	F	6.350		30	Elina	F	6.325
16	Kyösti	M	6.700		39	Juuso	M	6.963
4	Henna	F	7.162		28	Jenni	F	7.862

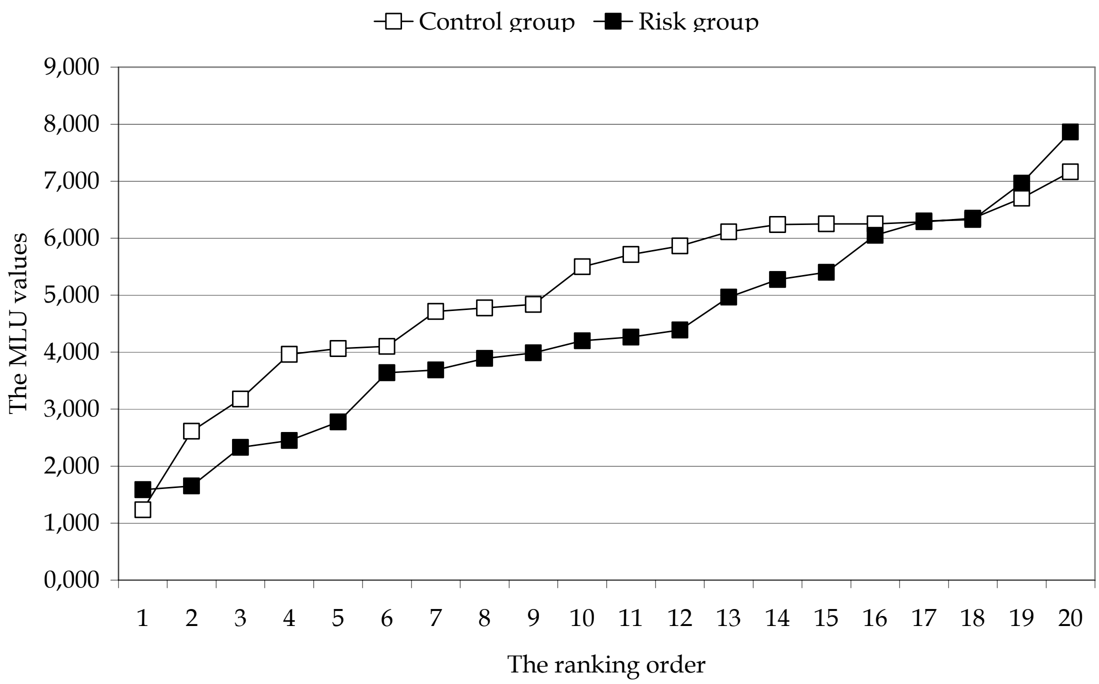


FIGURE 8 The MLU values of the control and risk children ranked in ascending order.

The highest MLU score (7.862) was obtained by a subject in the risk group (28 Jenni) but the overall performance of the group was at a lower level than that of the control children. Only when the MLU scores exceed 6.0 are both groups more or less equal. This is shown in Figure 8, where the results are set out in rank order to illustrate the differences between the overall profiles of the two groups. The curves show that the controls are some steps ahead up to the 17th position in the ranking order. At that point the curves overlap. The better performance of the controls is also shown in Table 3, where the average MLU values are listed at group level and gender level. The difference between the average group values is 0.69. In both groups the male subjects obtained lower values. However, the difference between the genders is much greater in the risk group (0.88; in the control group the difference is 0.34). The majority of the control children tend towards the higher values whereas the majority of the risk children tend towards the lower values. The emphasis at either end of the scale is supported in the way children place themselves in relation to group average values of MLU. Eleven control children exceed the group average while only eight risk children reach the same in relation to their group average.

TABLE 3 The means (M) and standard deviations (SD) of MLU both in the risk and control group.

	Control group			**Risk group**	
	M	**SD**		**M**	**SD**
All subjects	5.09	1.53		4.40	1.75
Female subjects	5.30	1.91		4.84	2.03
Male subjects	4.69	1.29		3.96	1.39

Despite the differences in method, these results are consistent with those of Scarborough (1990b), who reported that at the age of 30 months children from control families produced longer utterances (mean MLU 2.87) than dyslectic children from risk families (mean MLU 2.35). In addition, Scarborough separated a group of later normal readers coming from dyslectic families, and found that they had on average the highest MLU of all (2.97). This indicates that, as with the risk group of the present study, in Scarborough's research too there were both poorly and extremely well performing children coming from dyslectic families, even though the differences between the group averages were smaller than in the present study. Scarborough concluded that syntactic skills (measured by both MLU and IPSyn) at the age of 30 months were closely related to the later outcomes in reading ability of these children, and that children who are later diagnosed as dyslectics are likely to produce shorter utterances with less complex syntactic structures than children from control families.

The tendency of control children to produce utterances consisting of a greater number of morphemes is seen also in the study by Lyytinen, Poikkeus, Laakso, Eklund and Lyytinen (2001), where 200 Finnish children, 106 if them from dyslectic families, were investigated in terms of language development and symbolic play. Among other measurements and tests the three longest utterances produced by each child around the age of 24 months were collected in order to calculate the maximum sentence length (MSL). Utterances were reported by parents in the Finnish toddler version of the MacArthur Communicative Development inventories. MSL was found to reliably differentiate the groups [$F(1, 195)=4.72$, $p<.05$], whereas differences in vocabulary comprehension at 14 months and vocabulary production combined with the Bailey expressive score at 24 months were found not to be significant. (Lyytinen et al 2001: 878–79.)

To sum up, the MLU results both for the risk group and the control group indicate great individual variation in speech production. The MLU values for both groups have a similar range and level but when the group profiles and the average values are considered, a tendency toward better performance in the control group becomes obvious. The MLU values achieved in the control group tend towards the higher end of the scale, and in the risk group towards the lower end. Effective complexity measured in terms of utterance length in morphemes does differentiate the groups but the differences are not very remarkable.

4.3 Group comparison using the Finnish version of IPSyn

4.3.1 The total IPSyn points

Except in the case of three children (5 Liisa, 12 Lauri and 40 Risto) the 80 longest utterances were analyzed using the Finnish version of the Index of Productive Syntax (Nieminen & Torvelainen 2003). The fact that Liisa had only 60 utterances, and Lauri and Risto 70 utterances each analyzed did not strongly affect the results. Each child's utterances were analyzed in order of length, starting with the longest. As shown in Figure 9, most of the points came from the 10 longest utterances and the number of points diminishes dramatically after that, the average being less than 1 point in the utterance ranges 61–70 and 71–80. Thus, starting the analysis from the longest utterances equalized the situation between the subjects, and the lack of 10 or 20 utterances is not crucial. Moreover, such a development in the accumulation of points indicates that short utterances very seldom have morphosyntactic properties that have not already occurred in the longer utterances.

The total IPSyn points arranged by subject, group, and gender are presented in Table 4. In both groups, there is great individual variation and the total points range from 10 to 86 in the control group and from 26 to 86 in the risk group. The highest possible score in the Finnish version of IPSyn is 98 points.

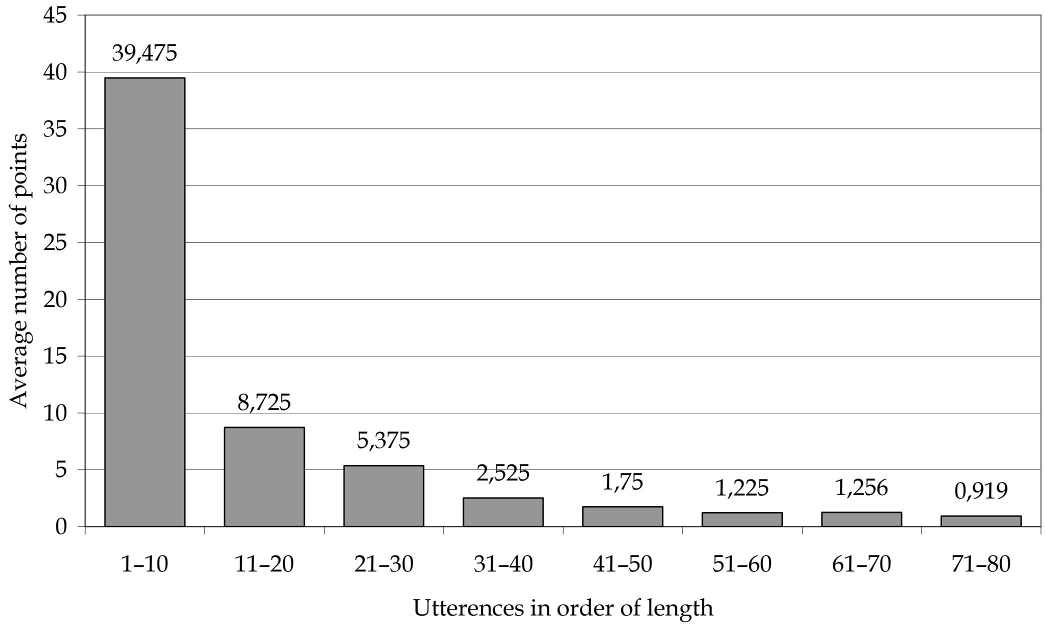


FIGURE 9 The average accumulation of IPSyn points at 10-utterance intervals, starting from the 10 longest utterances.

TABLE 4 The total IPSyn points based on the 80 longest utterances from each subject in the risk (RG) and control (CG) groups. * 70 utterances available. ** 60 utterances available. The purpose of the gray lines will be explained in connection with Table 9.

Control group					**Risk group**			
ID	**Name**	**Gender**	**IPSyn**		**ID**	**Name**	**Gender**	**IPSyn**
5**	Liisa	F	10		26	Tiina	F	26
12*	Lauri	M	47		25	Siiri	F	30
6	Ronja	F	51		40*	Risto	M	30
9	Lassi	M	51		37	Henri	M	39
18	Saku	M	53		32	Juho	M	41
17	Tatu	M	56		22	Anna	F	42
8	Aino	F	61		33	Jukka	M	42
11	Jaakko	M	62		29	Elisa	F	57
13	Joel	M	64		36	Sampo	M	57
10	Santeri	M	65		35	Seppo	M	60
2	Riikka	F	68		38	Pekka	M	65
7	Taru	F	70		24	Sanni	F	66
15	Mika	M	70		31	Aleksi	M	66
16	Kyösti	M	73		34	Leo	M	68
19	Tuomo	M	75		27	Janna	F	69
3	Laura	F	76		23	Anniina	F	72
20	Tuomas	M	76		30	Elina	F	74
14	Patrik	M	78		21	Paula	F	77
1	Tuija	F	84		39	Juuso	M	83
4	Henna	F	86		28	Jenni	F	86

TABLE 5 The means (M) and standard deviations (SD) of total IPSyn points both in the risk and control group.

	Control group			**Risk group**	
	M	**SD**		**M**	**SD**
All subjects	63.80	16.86		57.5	18.32
Female subjects	63.25	24.41		59.90	20.58
Male subjects	64.17	10.56		55.10	16.50

The group means and standard deviations of the total IPSyn points for both groups presented in Table 5 show that there are clear differences both on the group and gender levels. The control children perform better than risk children and the difference between the group means is 6.3 points. The male subjects in the control group show clearly better performance when compared with the male subjects in the risk group, the difference being as much as 9.07 points. The female subjects also have higher points in the control group than in the risk group, but the difference is only 3.35 points. The group differences would be larger if Liisa (ID 5) was excluded from the data. Her performance (10 points) diverges remarkably from that of the other control children, being 37 points lower than the next control child (Lauri, ID 12) and 41 points lower than the next female control (Ronja, ID 6) in the ranking order. The exclusion of Liisa would raise the average points in the control group up to 66.63 (SD = 11.43), and the average of the female controls as high as 70.85 points (SD = 12.44). Consequently the group differences would then increase to 9.13 and 10.95 points respectively. Liisa's deviant performance is easy to see in Figure 10, where her total IPSyn points come first in the ascending order of the control children.

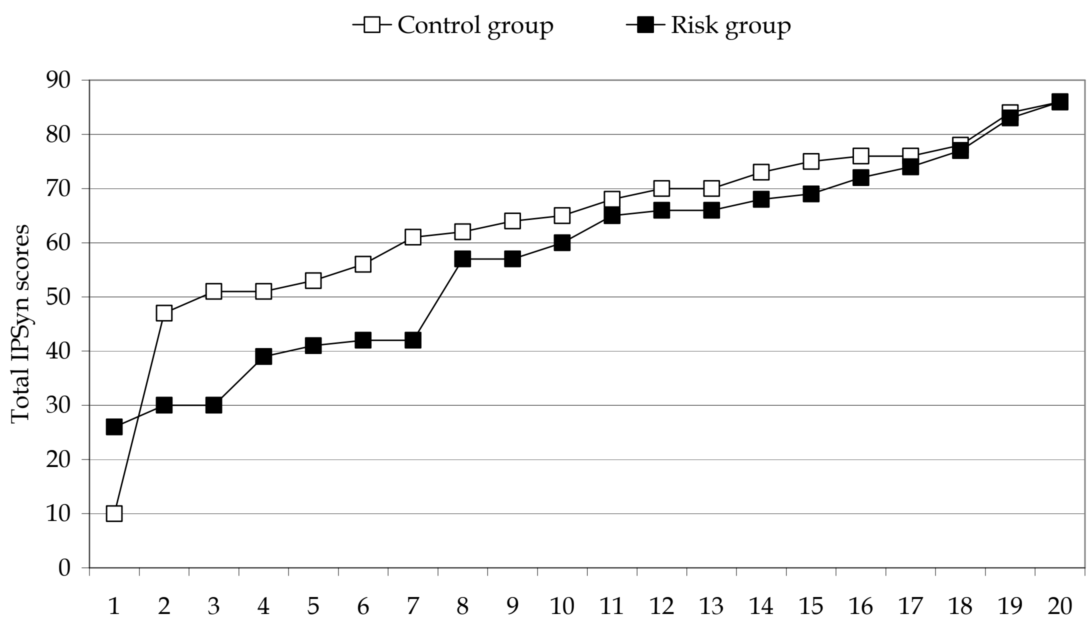


FIGURE 10 The total IPSyn scores of the risk and control children in ascending order.

In Scarborough's study (1990b) poorer performance in children from dyslectic families was also evident. Both the normal and the disabled readers from dyslectic families had a lower IPSyn group average (58.8 and 46.6 points respectively) than normal readers (M = 61.7) from control families.

In Figure 10, the IPSyn results of the risk and control children are arranged in ascending order, and the profiles of both groups can be seen. The difference between the group average values is mostly explained by the difference at the lower end of the scale. Seven children in the risk group form a subgroup with IPSyn points which are remarkably lower than the scores of the other risk children. From the 8th position onwards, the groups have fairly similar profiles in terms of IPSyn scores.

Like the MLU values, the IPSyn scores indicate great individual variation in both groups. Despite the rather similar range in points, the average scores presented in Table 5 as well as the group profiles illustrated in Figure 10 provide evidence of the poorer performance of the risk group. However, this is mostly due to the risk children's bias towards the low end of the IPSyn scale.

4.3.2 IPSyn points in the separate sections

4.3.2.1 Nominal phrases

MLU is a simple figure which cannot be divided into any components[5] but the IPSyn scores consist of the results from three distinct sections (nominal phrases, verb constructions, and sentence structures) and this enables further analysis of their composition. It is quite possible that one or two of the sections are more responsible for the tendency toward poorer performance in the risk group while the sections counterbalance each other, causing the difference in total score not to become significant. To find out if there are any trade-offs between the sections and differences in the make-up of the scores, the groups were compared section by section and the basis of the scores was investigated.

The section for nominal phrases (NP) in the Index of Productive Syntax concentrates on inflected nouns, pronouns, adjectives and numerals as well as on multiword noun phrases. The maximum score in the NP section is 20 points, and it is obtained as a result of two different occurrences of each of the following forms or structures: the plural inflection, inessive/adessive case, elative/ablative case, illative/allative case, partitive case, and genitive case, a clitic particle attached to any nominal word or adverb, a two-word and a three-word nominal phrase. In addition to these nine specified items, there is an unspecified item for, for example, the rare productive cases such as translative or essive.

The individual scores in the NP section together with the average values calculated on the group and gender level are presented in Table 6. The table

---

5 It is possible to use MLU calculation of selected structures (e.g. noun phrases) as well. However, this requires a separate calculation procedure for each structure type. In IPSyn the separate results from nominal phrases, verb constructions and sentence structures are a by-product of the index and require no additional analysis of data.

shows that the groups are very similar in their average scores. In both groups the male subjects have a slightly better average score and less variation than the female subjects.

TABLE 6 The individual scores together with the group means (M) and standard deviations (SD) in the NP section of IPSyn arranged by subject, group and gender. The purpose of the gray lines will be explained in connection with Table 9.

Control group					Risk group			
ID	Name	Gender	NP		ID	Name	Gender	NP
5**	Liisa	F	2		22	Anna	F	8
8	Aino	F	11		25	Siiri	F	8
12*	Lauri	M	11		40*	Risto	M	8
10	Santeri	M	12		26	Tiina	F	9
7	Taru	F	13		32	Juho	M	11
9	Lassi	M	13		37	Henri	M	12
17	Tatu	M	13		33	Jukka	M	12
6	Ronja	F	14		29	Elisa	F	13
13	Joel	M	14		24	Sanni	F	14
18	Saku	M	14		34	Leo	M	14
20	Tuomas	M	14		23	Anniina	F	15
2	Riikka	F	15		31	Aleksi	M	15
11	Jaakko	M	15		35	Seppo	M	15
15	Mika	M	15		36	Sampo	M	15
16	Kyösti	M	15		27	Janna	F	16
3	Laura	F	16		38	Pekka	M	16
1	Tuija	F	17		21	Paula	F	17
19	Tuomo	M	17		28	Jenni	F	17
4	Henna	F	18		30	Elina	F	17
14	Patrik	M	19		39	Juuso	M	19

	Control group			Risk group	
	M	SD		M	SD
All subjects	13.90	3.52		13.55	3.33
Female subjects	13.52	5.06		13.40	3.75
Male subjects	14.33	2.15		13.70	3.06

The control children's tendency to perform slightly better than the risk children in nominal inflection and multi-word nominal phrases is illustrated in Figure 11. The children in both groups are arranged in rank order by their scores in the NP section. The figure suggests that the small difference between the groups is due to the performance of children at the lower end of the scale. Otherwise the groups have almost identical results. Thus the NP section makes no exception to the results reported so far in the present study.

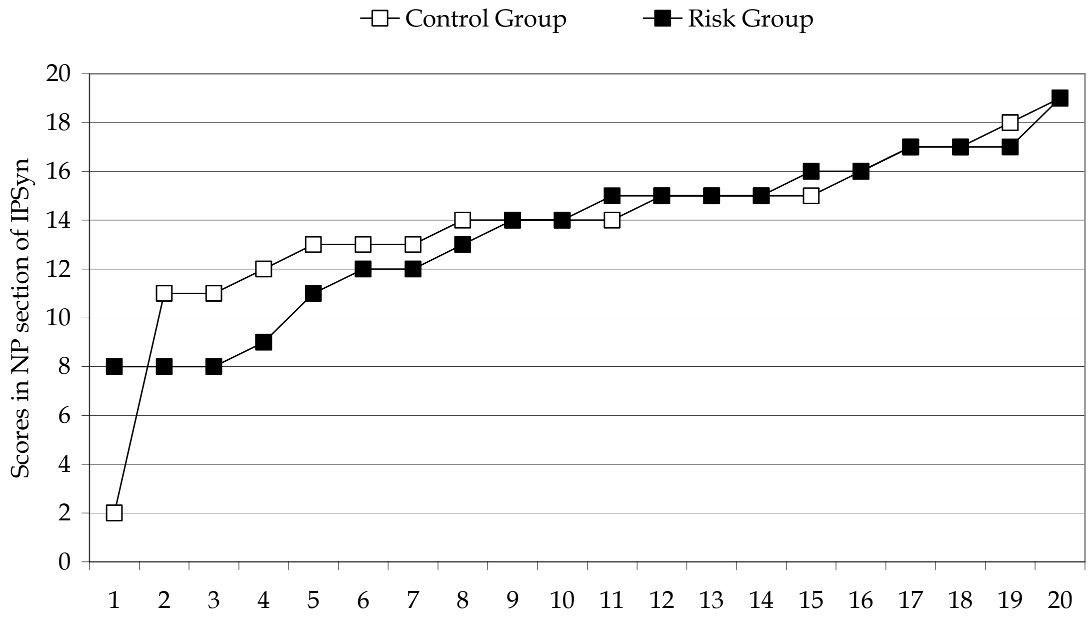


FIGURE 11 The NP section scores of the risk group and the control group in ascending order.

Another noteworthy point emerging from the individual scores in Table 6 and especially in the illustration of them in Figure 11 is that the children attain very high scores in the NP section. Ten risk children and nine control children scored 15 or more points, in other words, at least 75% of the maximum score. There are several possible reasons for such high scoring. First, the inflection of nouns, pronouns, adjectives, and numerals is a central part in Finnish language, and thus it is not surprising that Finnish children produce inflected word forms in a very early phase of language acquisition (for acquisition age and the order of acquisition of Finnish bound suffixes, see Toivainen 1980, 1990). This essential property of Finnish may have been emphasised by collecting the data in a situation where a child takes toys from and out of something, puts them into, onto and under something, gives something to somebody, et cetera. Therefore the action in the play session may have provoked the children to use more inflected forms. Another possible reason for the high scores in the NP section could be a methodological one. In the Finnish version of IPSyn the rich bound morphology of the nominal words is compressed into 7–8 items (N1–N7, possibly also N10), and in addition to this, three of these items include two cases with no requirement of examples of both to get the maximum of 2 points. Thus, the final score in the NP section may be high without sufficient examples from an especially versatile nominal morphology. Methodologically speaking, this is a point where the Finnish version of IPSyn probably needs some adjustments.

The main reason for examining the separate sections in a more detailed fashion was to find out if any of the sections is more responsible for the observed tendency toward better performance by the control children.

According to the analysis reported in this chapter the NP section is not what differentiates the groups remarkably. Therefore the importance of the verb construction and sentence structure sections may be greater.

4.3.2.2 Verb constructions

The section for verb constructions (VC) in the Index of Productive Syntax consists of 17 distinct verb structures, one of them (V17) being an unspecified item for rarely occurring complex forms such as the pluperfect, combinations of at least four verbs et cetera. Nine of the VC structures are inflectional, and the rest of the 16 specified items deal with multiword verb expressions. The inflectional structures include examples of personal inflection (V1–V2), the passive voice (V3), tenses (V4–V5), mood (V6–V7), and the 3rd infinitive (V8) as well as a clitic particle attached to a verb (V9). In multiword structures there are verb chains (V12–V14), utterances including the negative auxiliary (V15–V16), and examples of the coexistence of a verb and adverb(s) (V10–V11). The maximum score in the VC section is 34. (See Appendix 2 for a detailed presentation of the IPSyn items.)

Table 7 shows that the VC scores range from 4 to 30 points in the control group and from 8 to 29 points in the risk group. The average scores differ by 2.80 points in favour of the control group. When the scores from both groups are arranged in rank order, as in Figure 12, an overall profile somewhat different from those illustrated in Figures 9–11 is revealed. Both curves do climb in similar fashions, but, except at the lower end, the risk group curve is constantly 2–7 points lower than that of the controls. In Figures 9–11 it was shown that despite the differences at the level of the group top scores in both groups were similar. However, in this case the curves do not overlap at any point and therefore the difference this time is a consequence of the performance of all the risk children and not only those at the lower end of the scale.

In the NP section almost half of the children (10 risk subjects and 9 control subjects) had at least 75 % of the maximum score. In the VP section the situation is different. Even though the majority of the subjects do reach the 50% boundary (= 17 points), only 2 in the risk group and 6 in the control group exceed 75 % (= 25.5 points). The easiest way of interpreting these results would be to argue that the children are concentrating more on noun phrases and are more advanced in noun inflection than in the verb inflection. However, both the recording situation and the structure of verb section itself must be considered carefully. In Section 4.3.2.1 it was stated that the play session during which the data was collected emphasised the use of inflected nominal phrases. At the same time, the play session did not require expressions of time other than the present tense. Therefore children were able to handle the situation very well with only a few inflectional verb forms. This may have partly affected the results, even though the VC section does not deal with inflection only. In other words, the situation does not force a child to show all her/his skills.

TABLE 7 The individual scores together with group means (M) and standard deviations (SD) in the VC section arranged by subject, group, and gender. The purpose of the gray lines will be explained in connection with Table 9.

Control group					Risk group			
ID	Name	Gender	VP		ID	Name	Gender	VP
5**	Liisa	F	4		26	Tiina	F	8
12*	Lauri	M	14		40*	Risto	M	11
9	Lassi	M	15		25	Siiri	F	12
18	Saku	M	18		37	Henri	M	13
6	Ronja	F	19		32	Juho	M	14
10	Santeri	M	21		22	Anna	F	15
11	Jaakko	M	21		33	Jukka	M	17
15	Mika	M	22		23	Anniina	F	20
17	Tatu	M	22		24	Sanni	F	20
2	Riikka	F	23		27	Janna	F	20
8	Aino	F	23		29	Elisa	F	20
20	Tuomas	M	23		35	Seppo	M	20
7	Taru	F	24		34	Leo	M	21
16	Kyösti	M	25		38	Pekka	M	21
3	Laura	F	26		21	Paula	F	23
13	Joel	M	26		30	Elina	F	23
14	Patrik	M	26		36	Sampo	M	23
19	Tuomo	M	26		31	Aleksi	M	24
1	Tuija	F	30		39	Juuso	M	28
4	Henna	F	30		28	Jenni	F	29

	Control group			Risk group	
	M	SD		M	SD
All subjects	21.90	5.95		19.10	5.53
Female subjects	22.38	8.30		19.00	5.98
Male subjects	21.58	4.12		19.20	5.37

Another reason for not reaching as high proportion of the maximum score as in the NP section may be that nominal inflection and multi-word nominal phrases are compressed into rather few items, whereas the VC section is more fine-grained by nature. It is likely that this leads to bigger differences between the individuals and groups, and reaching the highest possible score is more challenging.

In seeking out the sources of the differences between the risk and control children the investigation of the verb constructions in IPSyn suggests that the tendency towards better performance by the control children is at least partly due to their more versatile use of verb constructions. In terms of effective complexity, the results suggest that on average verb structure resources are smaller for risk children than control children. This leads to the suggestion that the risk children have a less complex grammar of verb structures.

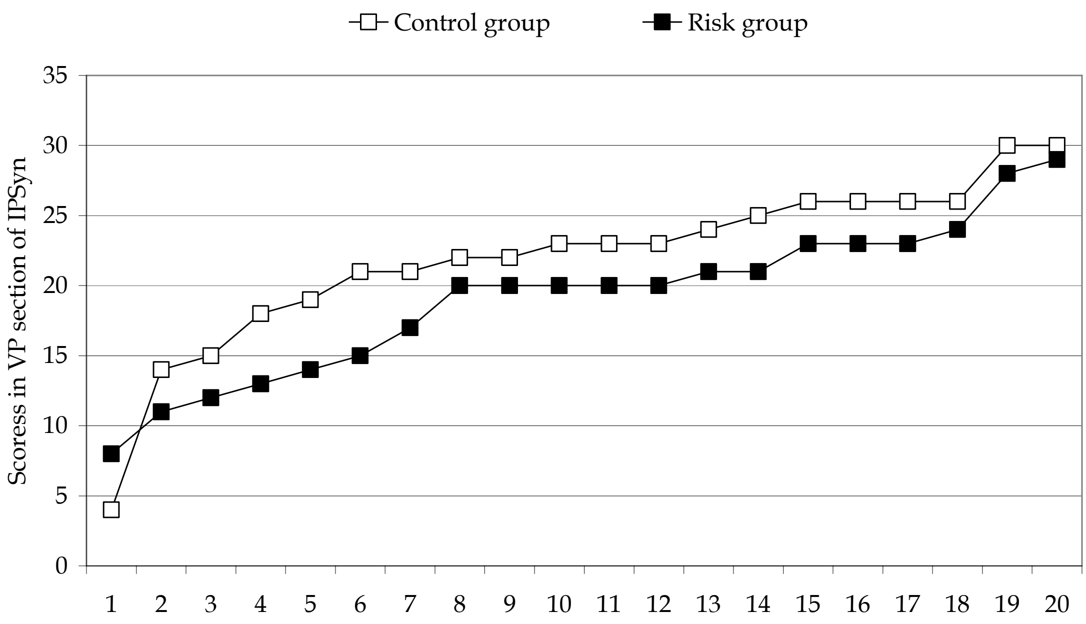


FIGURE 12 The VC section scores of the risk group and the control group in ascending order.

However, since IPSyn is measured by repetitive occurrences of the structures the results are biased and the relationship between single and double occurrences remains unexplored.

4.3.2.3 Sentence structures

The Sentence Structure section (SS) in IPSyn focuses on sentences and other multi-word combinations. In the SS section there are 21 specified structures plus 1 unspecified item for rarely occurring or otherwise less essential syntactic combinations in children's utterances. Thus the maximum score in the SS section is 44 points. The majority of the items consist of various combinations of subject, verb, object and adverbial (S2–S14). Another area of focus is the coordination and subordination of clauses (S17–S21). In addition to these, IPSyn registers any combinations of two words (S1), infinitive phrases in object position (S15) and adpositional phrases (S16).

The individual scores together with group means in the SS section are presented in Table 8 arranged by group and gender. The scores range from 4 to 39 points in the control group and from 9 to 40 points in the risk group. The large variation within groups is reflected also by the high standard deviation figures. The group means differ from each other by 3.15 points. The difference is mostly due to the poorer performance of the male subjects in the risk group (M=22.20) while all the control subjects (female subjects, M=27.63; male subjects, M=28.25) and the female subjects in the risk group (M=27.50) are similar in their average scores.

TABLE 8 The individual scores together with group means (M) and standard deviations (SD) in the SS section of IPSyn arranged by the subject group and gender. The purpose of the gray lines will be explained in connection with Table 9.

Control group					**Risk group**			
ID	**Name**	**Gender**	**IPSyn**		**ID**	**Name**	**Gender**	**IPSyn**
5**	Liisa	F	4		26	Tiina	F	9
6	Ronja	F	18		25	Siiri	F	10
17	Tatu	M	21		40*	Risto	M	11
18	Saku	M	21		33	Jukka	M	13
12*	Lauri	M	22		37	Henri	M	14
9	Lassi	M	23		32	Juho	M	16
13	Joel	M	24		36	Sampo	M	19
11	Jaakko	M	26		22	Anna	F	19
8	Aino	F	27		29	Elisa	F	24
2	Riikka	F	30		35	Seppo	M	25
10	Santeri	M	32		31	Aleksi	M	27
19	Tuomo	M	32		38	Pekka	M	28
7	Taru	F	33		24	Sanni	F	32
14	Patrik	M	33		27	Janna	F	33
15	Mika	M	33		34	Leo	M	33
16	Kyösti	M	33		30	Elina	F	34
3	Laura	F	34		39	Juuso	M	36
1	Tuija	F	37		21	Paula	F	37
4	Henna	F	38		23	Anniina	F	37
20	Tuomas	M	39		28	Jenni	F	40

	Control group			**Risk group**	
	M	**SD**		**M**	**SD**
All subjects	28.00	8.37		24.85	10.26
Female subjects	27.63	11.48		27.50	11.37
Male subjects	28.25	6.08		22.20	8.80

When the results are arranged in rank order (Figure 13), the curves show group profiles similar to those detected already in the previously presented analyses of MLU and total IPSyn as well as the scores in the NP section. The low end of the scale is remarkably lower in the risk group than it is in controls, but when it comes to the highest scores, both groups are equal, and in fact, the highest SS score of all (40 points) wass reached by a risk subject. Another noteworthy point illustrated in Figure 13 and Table 8 is that high scores are very common in both groups. There are as many as 8 control children and 7 risk children who have at least 75% (= 33 points) of the maximum score. At the same time, 4 control children and 8 risk children have less than 50% of the highest possible score.

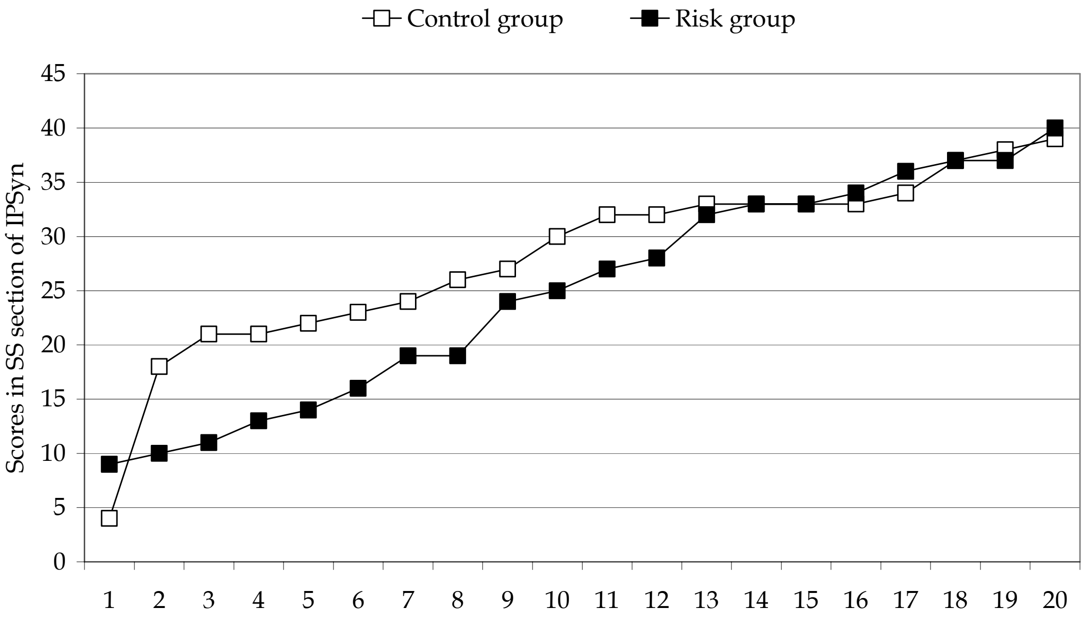


FIGURE 13 The SS section scores of the control group and the risk group in ascending order.

According to the results presented here, it is again the lower end of the scale that is more important when the differences between the controls and risks are considered. The poor performers in the risk group tend to have lower scores than the poor performers in the control group, and the good performance of the risk children at the high end of the scale is not enough to counterbalance the group difference. In terms of effective complexity, these results suggest that as a group the control children have more syntactic resources and thus more complex syntax than risk children. However, the relationship between the actual number of different structures and the number of points is not clear and straightforward because IPSyn gives credit for repetitive occurrences of each structural item.

4.3.3 The effect of tuning: the number of structural items in the IPSyn results

It has already been stated on several occasions that IPSyn is a tuned measure of effective complexity. On the one hand it searches for and counts the distinct structural patterns in a child's utterances. The total number of these patterns could serve as a simple index of effective complexity. On the other hand it also searches for evidence of productivity in the form of repeated occurrence of each specified structure. It adds value for each such repetition so that the structure occurring once is credited with one point, whereas a repeated occurrence is credited with two.

In a way, the idea of crediting productivity is in line with complexity, since in two different occurrences the same basic structure may be elaborated in different ways, one being more complex than another. For example, V + O (item S5 in the SS section) may have, among others, the following manifestations:

ostaa auton	v|osta-3S n|auto-gen	'buys a car'
ostin auton	v|osta-PAST-1S n|auto-GEN	'I bought a car'
en osta autoa	v:neg|ei-1S v|osta n|auto-PARTIT	'I won't buy a car'
ostettiin auto	v|osta-PASS&PAST n|auto	'we bought a car'
ostettaisiinko auto	v|osta-PASS-COND-CLIT:INT n|auto	'should we buy a car'

In each of these, the verb *ostaa* 'buy' is elaborated in a distinct way and, moreover, the noun phrase in the object position has three different forms. Structurally, the last example *ostettaisiinko auto* is more complex with three suffixes attached to the verb than the one preceding it, *ostettiin auto*. Thus, these two variants in the V + O entry in the sentence structure section would definitely indicate differences of complexity within the item. However, there is no requirement that the second occurrence should be more complex than the first. The only rule is that they are *distinct from each other*, since that is adequate evidence of emerging productivity of the structure.

The idea of effective complexity is to look for patterns and then judge complexity on the basis of their frequency. In IPSyn this is done by specifying the patterns beforehand and then looking for realisations of them in language production. The example above, V + O structure, is one of the specified patterns in IPSyn. In terms of effective complexity, repeated occurrence of a pattern does not make it indicative of complexity, even though the specification of the pattern is so broad that it allows substantial variation. All the examples of V + O manifestations listed above are counted as representatives of one single structure, and the morphological elaboration of the components in each manifestation is investigated separately in terms of other specified patterns in the verb constructions and the nominal phrase sections respectively. Thus, the present method of giving credit for repeated occurrences of a pattern has nothing to do with effective complexity, but it does have a strong influence on the final IPSyn scores and thus on the interpretation of these results.

In the foregoing sections the IPSyn results have been analysed in their original form, that is, as points scored. However, to see how strongly the credit given to emerging productivity actually affects the results and our understanding of the original number of patterns found in children's utterances, the analysis moves now to the basis of the IPSyn points. In Table 9, the total IPSyn scores and the number of structural items are presented side by side. In addition to this, the coefficient for generating points is calculated as well. In both groups the correlation between the number of items and the total IPSyn scores is very high (.99 in the control group and .98 in the risk group respectively) indicating that the increase in items is congruent with the increase in total scores. The coefficient for generating points grows alongside the number of items. In terms of effective complexity such a development indicates that counting double occurrences of the structural items actually emphasizes differences in effective complexity: the more items there are, the higher the coefficient for generating points and thus the higher the total score. In other words, looking for evidence of productivity of the items does not confuse the evaluation of effective complexity in the grammatical system. On the contrary, it makes differences even clearer.

TABLE 9 The total IPSyn scores, the number of structural items found in each child's utterances and the coefficient for generating the total scores. The table shows the average score and average number of items as well as the average coefficient for both groups and both genders. The correlation between the score and the number of items is calculated as well. The gray lines show the children who have less than 30 structural items in their IPSyn results.

Control group					**Risk group**				
ID	Name	Score	Items	Coefficient	ID	Name	Score	Items	Coefficient
5**	Liisa	10	7	1,43	26	Tiina	26	17	1,53
12*	Lauri	47	27	1,74	40*	Risto	30	17	1,76
6	Ronja	51	27	1,89	25	Siiri	30	21	1,43
9	Lassi	51	29	1,76	37	Henri	39	22	1,77
18	Saku	53	32	1,66	32	Juho	41	29	1,41
17	Tatu	56	33	1,70	22	Anna	42	25	1,68
8	Aino	61	34	1,79	33	Jukka	42	26	1,61
11	Jaakko	62	35	1,77	29	Elisa	57	33	1,73
13	Joel	64	36	1,78	36	Sampo	57	33	1,73
10	Santeri	65	37	1,76	35	Seppo	60	34	1,76
2	Riikka	68	40	1,70	38	Pekka	65	35	1,86
7	Taru	70	41	1,71	24	Sanni	66	37	1,78
15	Mika	70	39	1,79	31	Aleksi	66	40	1,65
16	Kyösti	73	41	1,78	34	Leo	68	37	1,84
19	Tuomo	75	42	1,79	27	Janna	69	41	1,68
3	Laura	76	43	1,77	23	Anniina	72	42	1,71
20	Tuomas	76	41	1,85	30	Elina	74	41	1,80
14	Patrik	78	42	1,86	21	Paula	77	44	1,75
1	Tuija	84	45	1,87	39	Juuso	83	44	1,89
4	Henna	86	46	1,87	28	Jenni	86	46	1,87

	Control group				**Risk group**		
	Score	**Items**	**Coefficient**		**Score**	**Items**	**Coefficient**
All subjects	63.80	35.85	1.76		57.50	33.20	1.71
Female subjects	63.25	35.38	1.75		59.90	34.70	1.70
Male subjects	64.17	36.17	1.77		55.10	31.70	1.73
Correlation			0.99				0.98

Averaging the index figures shown in Table 9, the control children perform better than the risk children. Moreover only a subgroup of children seems to be responsible for the clear difference between the groups, namely the children with less than 30 items in their IPSyn results (coloured grey in Table 9). Seven risk children and four control children form the subgroup. If they are eliminated from the results, the groups turn out to be astoundingly similar. The average number of structures changes to 39.19 in the control group and 39.00 in the risk group, and the corresponding average coefficients for generating total

IPSyn scores change to 1.78 (CG) and 1.77 (RG). Similarly, the difference between the average total scores disappears and the mean scores rise to 69.81 (CG) and 69.23 (RG) points.

To conclude, the risk children do not perform any worse than the control children if they are matched with controls according to the number of structural items they produce. Risk children producing the same number of items as controls also earn the same number of points from them. In this respect, they are performing similarly, as Figure 14 shows. The explanation for the existing group differences shown in Table 9 is simply the fact that in the risk group there are more children with fewer items in their structural resources, resulting in lower total scores. The performance of the subgroup of seven risk children substantially lowers the performance of the whole group. These results are consistent with the foregoing discussions of the differences between the risk group and the control group.

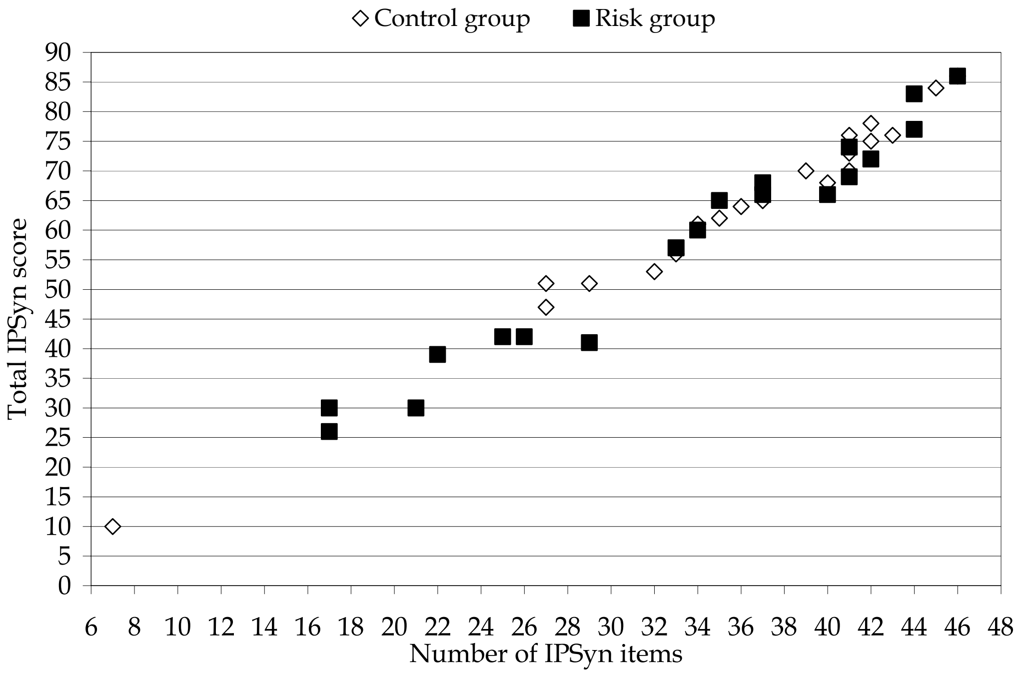


FIGURE 14 The relationship between the number of structural items found in the IPSyn analysis and the total IPSyn score in control and risk children.

4.4. The relationship between MLU and IPSyn scores

4.4.1 The correlation between MLU values and IPSyn scores

In this section, the results for both groups are combined into a single sample to enable analysis of the relationship between Mean Length of Utterance and the Index of Productive Syntax. In Section 3.4 it was suggested that pooling information from both these methods would give more relevant knowledge about structural complexity than either measure alone or a comparison of them. It is reasonable to think that MLU and IPSyn evaluate different aspects of language production rather than the same aspect from two different points of view. Therefore, combining the two aspects (length and resources) may give a more detailed picture of the language produced.

The first task is to find out how well the two indices correlate with each other, since this is an easy way to start assessing what the interaction of the two scales of measurement may reveal. The correlation between MLU values and IPSyn scores is very high (r = .92). The relationship between the two indices is illustrated in Figure 15. A positive correlation was expected, but it is surprising that the correlation is so high because of the exceptionally wide MLU range (1.233–7.862). In several earlier studies (Blake et al. 1993; Brown 1973/1976; Klee & Fitzgerald 1985; Rondal et al. 1987; Scarborough et al. 1991) it has been stated that MLU correlates with other scales used for the evaluation of grammatical complexity only when the MLU scores are relatively low, between 1.0 and 2.5–4.5, the upper limit varying in different studies. This has been interpreted as evidence of MLU being a valid measure of grammatical complexity within this range of values. To find out if the same tendency can be seen also in the present study, the correlation between the two scales was recalculated for different ranges of MLU. The analysis was bisected at the point of MLU 4.5, which was the reliability limit for MLU in Blake et al. (1993). At the same time it is also the approximate midpoint of MLU values, ranging from 1.223 to 7.862, in the present study. There are 18 children in the lower (<4.5) and 22 children in the higher (>4.5) MLU group. For the lower group the correlation between MLU values and IPSyn scores is high (r = .88), but for the higher group it is somewhat weaker (r = .69). Thus the results confirm the tendency already noted in other studies for MLU to be less informative above 4.5.

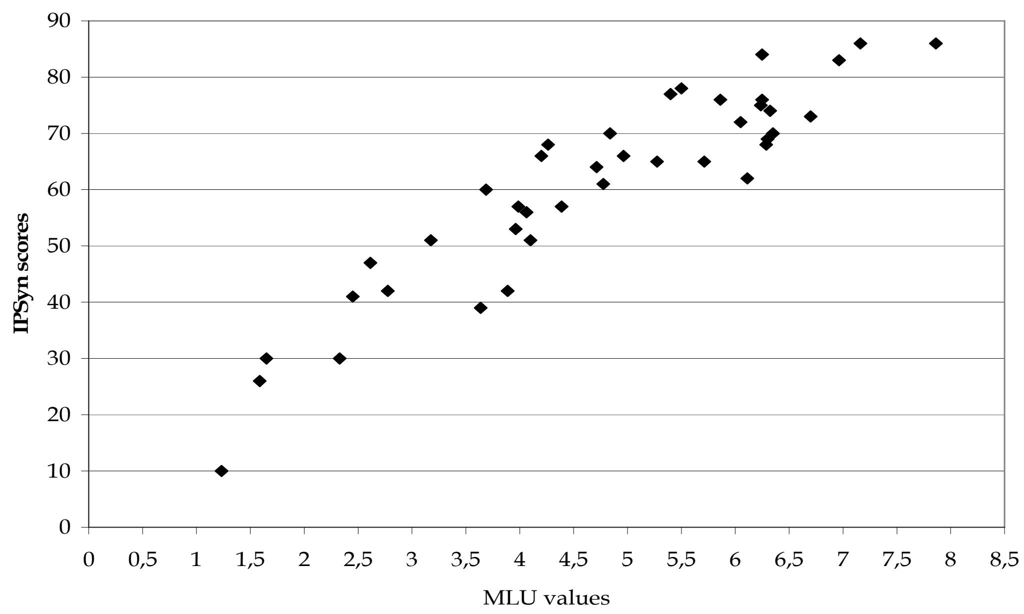


FIGURE 15 The relationship between MLU values and IPSyn scores.

When the correlation between MLU values and IPSyn scores is interpreted it is a question of matching length and resources. In early verbal behaviour and speech production it is obvious that a child constantly needs new words and new means to express herself/himself and to produce more adult-like and expressive utterances. In practice this means that as utterances grow longer, the lexical and grammatical resources to produce the utterances must also grow larger and more versatile. This fact has been known for a long time and Brown (1973/1976: 77) brings it out very clearly when introducing MLU: "– – almost every new kind of knowledge increases length: the number of semantic roles expressed in a sentence, the addition of obligatory morphemes, coding modulations of meaning, the addition of negative forms and auxiliaries used in interrogative and negative modalities, and of course, embedding and coordinating". This is the phase when it is logical to find a very high correlation between MLU and measuring scales like IPSyn.

Before long, introducing new grammatical resources is no longer as crucial to speech production as it was in the beginning, since it is possible to do a lot of things with the structures that are already part of the repertoire. For example, a single inflectional suffix may convey several different meanings in different contexts. In Section 4.3.2.1, it was noted that children tend to get relatively high scores in the NP section. This does not mean that they have mastered the use of noun inflection and noun phrases. Rather, it indicates that the children have a certain variety of means but the adult-like use of these means is not evaluated. In other words, IPSyn does not catch the different functions of an item because it concentrates only on the form. When the functional scope of morphological items is extended and their coexistence with each other becomes more common, MLU grows while IPSyn does not necessarily change at all. Exactly the same thing happens with the widening use of the syntactic items specified in IPSyn, as well as in combining the syntactic and morphological items together in utterances. Thus, more and more versatile use of each item included in IPSyn is a reasonable explanation for the weakening correlation between MLU and IPSyn when MLU exceeds 4.5. However, this does not yet explain the fact that when the whole range of MLU is taken into account the correlation is extremely high, $r = .92$.

4.4.2 Dissonances between MLU and IPSyn

Despite the high overall correlation between MLU values and IPSyn scores, there are some clear contradictions in the results. In this context, a contradiction means an unexpectedly large variation of one measure while the other remains within a narrow range. In other words, there are children, whose MLU values are similar but whose IPSyn scores differ substantially and likewise children whose IPSyn scores are similar but whose average length of utterance varies widely.

Within the combined group of risk and control children, three subgroups were separated out on the basis of an uneven relationship between the two indices. In all, 24 of the 40 children[6] belong to subgroups A, B and C as shown in Table 10 and Figure 16. Subgroup A consists of 9 children whose MLU varies from 3.688 to 4.388. However, for these children within a range of 0.7 morphemes in MLU terms, IPSyn scores vary from 42 to 68 points, a range which in fact covers as much as 26.5% of the whole IPSyn scale. Moreover, it is noteworthy that variation of this extent occurs within the limits (MLU < 4.5) that have been presented in some previous studies (e.g. Blake et al 1993) as valid for drawing reliable conclusions concerning complexity on the basis of MLU.

Subgroup C is another example of quite similar MLU values with widely varying IPSyn scores. This group of 9 children attains high MLU values, ranging from 6.050 to 6.350, and IPSyn scores of 62–84 points. The dissonance between MLU and IPSyn is not as surprising as in case of subgroup A, since these results are consistent with previous studies where it has been argued that when the MLU value is high the connection with other complexity indices is more random and unpredictable.

Between groups A and C, there is another subgroup representing an opposite type of contradiction. Subgroup B consists of 13 children, whose MLU varies from 3.688 to 6.350. The MLU range covers as much as 40.2% of the whole range in this study (1.233–7.862). However, the IPSyn scores for these children stay within the narrow range of 60–70 points. The three subgroups overlap: three of the children (31 Aleksi, 34 Leo, and 35 Seppo) belong to groups A and B, and four of the children (2 Riikka, 7 Taru, 11 Jaakko, and 27 Janna) to groups B and C. Therefore the group with the highly varying MLU (B) forms a bridge between the two other groups with the highly varying IPSyn results as illustrated in Table 10 and Figure16.

Two kinds of conclusions can be drawn on the basis of the contrasts presented above. First, the high correlation (r = .92) between MLU and IPSyn can easily be explained by the two kinds of contradictions that counterbalance each other hiding the discrepancy between the results stability and the other a strong phase of development for the same children range from 3.688 to 6.350 in MLU and from 42 to 84 points in IPSyn.

---

6 The 16 children excluded for not meeting the criteria. These children together with their MLU values and IPSyn scores are listed in the following table.

Control group					**Risk group**			
Name	**Gender**	**IPSyn**	**MLU**		**Name**	**Gender**	**IPSyn**	**MLU**
4 Henna	F	86	7,162		21 Paula	F	77	5,4
5 Liisa	F	10	1,233		25 Siiri	F	30	1,65
9 Lassi	M	51	3,175		26 Tiina	F	26	1,587
12 Lauri	M	47	2,614		28 Jenni	F	86	7,862
14 Patrik	M	78	5,5		32 Juho	M	41	2,45
16 Kyösti	M	73	6,7		33 Jukka	M	42	2,775
20Tuomas	M	76	5,862		37 Henri	M	39	3,638
					39 Juuso	M	83	6,963
					40 Risto	M	30	2,329

TABLE 10 The members of subgroups A, B and C, and their scores in MLU and IPSyn. The C/R column indicates whether the children originally belong to the control or the risk group. *Subjects belonging to groups A and B. **Subjects belonging to groups B and C.

Group A					Group B					Group C				
ID	**Name**	**C/R**	**MLU**	**IPSyn**	**ID**	**Name**	**C/R**	**MLU**	**IPSyn**	**ID**	**Name**	**C/R**	**MLU**	**IPSYN**
35	Seppo*	R	3.688	60	35	Seppo*	R	3.688	60	23	Anniina	R	6.050	72
22	Anna	R	3.888	42	31	Aleksi*	R	4.200	66	11	Jaakko**	C	6.112	62
18	Saku	C	3.962	53	34	Leo*	R	4.263	68	19	Tuomo	C	6.237	75
36	Sampo	R	3.987	57	13	Joel	C	4.713	64	3	Laura	C	6.250	76
17	Tatu	C	4.063	56	8	Aino	C	4.775	61	1	Tuija	C	6.250	84
6	Ronja	C	4.100	51	15	Mika	C	4.838	70	2	Riikka**	C	6.287	68
31	Aleksi*	R	4.200	66	24	Sanni	R	4.963	66	27	Janna**	R	6.300	69
34	Leo*	R	4.263	68	38	Pekka	R	5.275	65	30	Elina	R	6.325	74
29	Elisa	R	4.388	57	10	Santeri	C	5.713	65	7	Taru**	C	6.350	70
					11	Jaakko**	C	6.112	62					
					2	Riikka**	C	6.287	68					
					27	Janna**	R	6.300	69					
					7	Taru**	C	6.350	70					

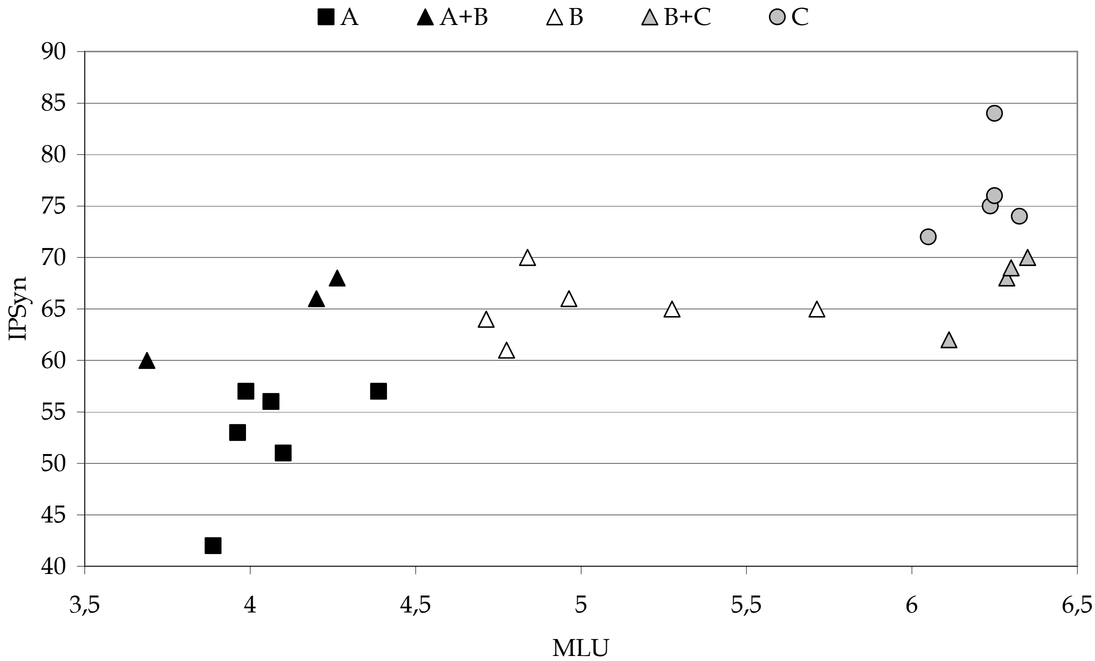


FIGURE 16 The three subgroups representing an unexpected relationship between MLU and IPSyn. The black marks refer to group A (MLU 3.688–4.388; IPSyn 42–68), the triangular marks to group B (MLU 3.688–6.350; IPSyn 60–70), and the grey marks to group C (MLU 6.050–6.350; IPSyn 62–84). The overlap of groups A and B is shown by the black triangular marks, and that of groups B and C by the grey triangular marks respectively.

These ranges leave out only the very lowest and highest results to form the body of the correlation. Secondly, the existence of these contradictory groups indicates that similarities in average utterance length may hide different kinds of morphosyntactic resources, and with resources of similar scope one can build utterances of various lengths. Neither of these excludes the possibility of growth in structural complexity, since MLU and IPSyn measure different aspects of structural complexity (length of production vs. richness of resources), and both aspects are needed for evaluating utterances in terms of multidimensional complexity.

These results are a fruitful basis for speculating about the structural properties of the children's utterances because they shed light on language from two distinct directions. The following Hypotheses 1–5 concerning the subgroups are based on these results:

Hypothesis 1
A short utterance length combined with a small number of resources (part of the children in group A) results in utterances where the multilayered structure is immature. The utterances include only very little elaboration, and the elaboration re-uses the same devices in distinct utterances. The number of syntactic components per utterance is small.

Hypothesis 2
A short utterance length combined with a substantially higher number of resources (some of the children in group A) results in utterances where the multilayered structure is still immature. Utterances have only a little elaboration, but now the elaboration has substantial variation from utterance to utterance. The structural

resources are not yet combined and the number of syntactic components remains small.

Hypothesis 3
When a high utterance length is combined with a large number of structural resources (children in group C), utterance length does not restrict elaboration of the utterances. The multidimensional structural complexity has grown both horizontally and vertically by adding new syntactic components and new elaborations. The structural resources are now combined in utterances. Instead of totally new grammatical details, this growth in resources mostly results in more alternatives for expressing different ideas.

Hypothesis 4
A widely varying length of utterance combined with stability of resources (children in group B) indicates that growth in structural complexity of utterances originates from combining more and more resources in utterances.

Hypothesis 5
Structural complexity grows in a stepwise fashion, where length and structural resources alternate as the focus of development. First the basic elaboration devices to create utterances are acquired. This is shown by short utterance length together with a large variation in resources. The second task is to learn to use the resources more effectively. In this step, the focus is on combining resources and deepening multilayered morphosyntactic structure, which simultaneously makes utterances grow in length. The next step concentrates again on acquiring new structural resources, since the possibilities of the already acquired ones are mastered quite well. At this point utterance length does not restrict elaboration or the number of syntactic components in utterances, and complexity may grow in both dimensions.

In Chapters 6–8, these hypotheses will be tested to investigate the relationship between the MLU and IPSyn scales and structural complexity.

4.5 Summary

In this chapter, two of the main research questions of the present study were given closer consideration. The first of these focused on potential differences between the risk and control groups:

> What is the relation between risk group children (children with high genetic risk of dyslexia) and control group children (children without high genetic risk of dyslexia) when special attention is paid to the complexity of morphosyntactic structures in their spontaneous speech production?

The other question was more of a methodological kind:

> What kind of information about structural complexity in utterances do commonly used scales of measurement like MLU and IPSyn give?

The means used to seek answers were Mean Length of Utterance and the Index of Productive Syntax as methods. Both of them are commonly used quantitative scales of measurement focusing on complexity in acquisition studies.

There was large variation in the control and risk groups in the results from both methods. This was expected, since 30-month-old children are in the middle of an important developmental phase in expressive language. Another factor which further widened the variation was the data selection method. Since the emphasis was on the longest utterances, short utterances which are rather common in conversation did not counterbalance them. This especially benefited the more advanced and talkative children because their final data sample consisted almost exclusively of multi-morphemic utterances.

Children in the control group performed better on both scales on average. They had longer utterances and their IPSyn scores were higher than those of the risk children when the average values were compared. These results were confirmed with further analysis of the NP, VC and SS sections in IPSyn.

Despite the clear tendency for poorer performance by the risk children, each analysis of MLU and IPSyn showed that there was actually a subgroup of seven children which was mostly responsible for the key ratios remaining on a lower level than in the control group. In other words, most of the risk children did not differ from the control children in average utterance length, IPSyn score, or extent of morphosyntactic resources. Therefore these results suggest that high genetic risk of dyslexia does not show in most of the children's verbal performance when their utterances are analysed quantitatively in terms of effective complexity with special emphasis on morphosyntactic structures. However, there is a subgroup of children within the risk group which is more or less behind the others in both utterance length and the extent of their structural resources. Whether this is caused by the high genetic risk of dyslexia or simply normal individual variation which accidentally happens to show up mostly in risk children, remains an unsolved problem.

Next, the MLU values and IPSyn scores of the risk and control children were combined preparatory to examining thr methods. The focus was on the pooling of the results from MLU and IPSyn. Despite a high correlation between the scales, three subgroups with contradictory results were found: in two of the subgroups (A and C) MLU indicated no or very little difference between the children whereas their IPSyn scores suggested substantial variation in complexity, while in the third subgroup (B) the situation was reversed: MLU showed substantial variation while the IPSyn scores remained within a narrow range. These contradictory results counterbalance each other so that the overall correlation becomes very high. Moreover, these results confirm the hypothesis that MLU and IPSyn do not measure the same kind of complexity, even though they both represent measures of effective complexity. Furthermore, these results give reason to believe that growth in morphosyntactic complexity proceeds in a stepwise fashion and that growth is shown in length of utterances or in extension of resources depending on the nature or the quality of the growth. The results suggest the existence of different acquisition phases and processes which cannot be discovered using only one scale of measurement. On the basis of these contradictory results, five hypotheses concerning distinct acquisition phases and an overall model of growth in morphosyntactic complexity were generated. These will be tested in Chapters 6 and 7.

5 UTTERANCE ANALYSIS

In this chapter a new method for evaluating morphosyntactic complexity is introduced. Utterance Analysis (UA) is a device for breaking down utterances into syntactic components, elaboration layers and actual means of elaboration. Utterance Analysis has actually three functions in the present study. First, it analyses utterances in a way that is fully consistent with the view of structural complexity adopted here. Therefore its results provide a comprehensive insight into the aspects of utterance structure which were speculated upon in Hypotheses 1–5. Secondly, and as a consequence of the first function, the results of Utterance Analysis enable re-evaluation of MLU and IPSyn as methods. Finally, Utterance Analysis is a means of re-evaluating the differences between the risk and control children in terms of multidimensional views of structural complexity.

5.1 The principles of Utterance Analysis

5.1.1 The foundation of the analysis

Utterance Analysis is a concrete application of the multidimensional view of structural complexity introduced in Chapter 2. The analysis concentrates on the output structure (Dahl 2004: 48–49) of utterances, and its foundation lies in views that any linguistic production consists of subparts, termed, for example, constituents or phrases in the linguistic literature. The utterance is taken to be the object of analysis, because it has been suggested to be the most fundamental psycholinguistic unit (Tomasello 2000c: 63), especially in the framework of the usage-based model of language acquisition. This ensures that all intelligible verbal expressions are equally included, since the definition of an utterance is based only on semantic content, intonation, and, to some extent, pausing, with no syntactic criteria (Tomasello 2000c: 63; Linell 1998: 159–161; for the definition of the utterance adopted in this study, see Chapter 1.4). If the usage-based model was followed more precisely, the sub-parts of utterances would have been identified individually for each child and they would have been called, for example, frames, frame structures, formulaic utterances, or amalgams. However, since my data is cross-sectional and does not provide any possibility for such individual identification of structures, the sub-parts analysed in Utterance Analysis are phrases (noun phrase, adjective phrase, infinitive phrase, and adpositional phrase) supplemented with certain other separable structures like verb constructions, interjections, particles and conjunctions. Since not all of the separated sub-parts are phrases and since the subparts are analysed down to morpheme level, I prefer to identify them as syntactic components rather than phrases or constituents. The identification of phrases is in accordance with the descriptions provided in Iso suomen kielioppi (ISK, 2004).

Utterance Analysis divides each utterance into components and analyses the morphosyntactic structure of each of them separately. One thing it does not pay attention to is the syntactic position of each component in an utterance because it does not make sense to identify subjects, objects, predicates, and adverbials when complete clauses or sentences are not required by the research context. Utterance Analysis is a description of the composition of each utterance, but dominating relations between the components are not relevant.

The Utterance Analysis carried out is closely connected to the 5 hypotheses based on the contradictory results from MLU and IPSyn. Three factors deserving special attention are mentioned in the hypotheses. These factors are the syntactic components, the morphological and syntactic elaboration and the layered structure of the utterances. The purpose of the Utterance Analysis is to provide new information on these points. This is done by placing each utterance on a chart (Figure 17) where the syntactic components are segmented and each elaboration of a syntactic component is carefully placed on layers of the chart and defined. Layers describe sequential elaborations within a component. Each word has its own column and each layer its own row in the chart. In the following chapters, the analyses of syntactic components, elaboration devices and layered structure are discussed in detail.

1.							Components
2.							Elaborations
3.						1.-4.	Layers
4.							
Utterance:	Word	word	word	word.			

FIGURE 17 Functional areas within the analysis chart of the Utterance Analysis. Components of an utterance are separated on the first layer. If any of the components include elaboration of any kind, description of this is registered on the chart above the word form it concerns.

5.1.2 The syntactic components

The first task in the analysis of a given utterance is to separate it into syntactic components and place them on the top row in the chart. Each component is labelled with an abbreviation (see Appendix 3 for the codes used in the Utterance Analysis chart) and, if the component consists of only one word, its word class is placed on the same row, under the component identification. If a syntactic component is a combination of two or more words they are placed on the next row and connected to the component label with a line. The placing of syntactic components and their constituents is illustrated in Figure 18.

1.	NP PRO:INTER		V	NP N
2.		aux	V	pl
3.			part	

Kuka *on* *ostanup* *point?*
Kuka on osta-nut poni-t?
Who has buy-PART pony-PL?
'Who has bought the ponies?'

FIGURE 18 Syntactic components placed in the Utterance Analysis chart. (17 Tatu; 5/336) [7]

The basic syntactic components separated in this analysis are mostly phrases of different kinds. The phrase types in Finnish are: the noun phrase (NP), adjective phrase (AP), adposition phrase (PP), adverb phrase (AdvP), infinitive phrase (InfP), and participle phrase (PartP). The name of each phrase refers to the main word in all types except one, namely the noun phrase. Other than a noun the main word of a NP can be also a pronoun or a numeral. (ISK 2004: 428–432.)

A syntactic component is not always a phrase. In this study finite verbs are not considered to be main words of a verb phrase because, in the first place, there is no clear evidence for the existence for a verb phrase in Finnish (ISK 2004: 433). Obviously finite verbs are important parts of utterances whenever they appear, and therefore they are treated as individual syntactic components alongside phrases. The various kinds of particles and interjections form another group of syntactic components which are not defined as phrases. However, these words and constructions are treated as individual syntactic components on the same level as phrases and finite verb constructions. Any hierarchy between the individual components in an utterance is ignored. This is a conscious choice based on the quality of the data. If all the utterances had been complete clauses it would also have been possible to investigate the hierarchy

[7] Each utterance used as an example is entered on four lines. A phonetic transcription of a child's utterance is entered on the top line, and an adult language 'translation' is placed on the next line. The third line is a glossed and tagged version of the utterance in English and the bottom line is a translation of the utterance. Each example also has information for identification attached. The information consists of the subject's identification number and name, the utterance number based on the order of length in the child's data sample and, finally, the number of the row in the original transcription, where the utterance can be found in its context.

between clausal components. However, one of the basic principles in this analysis is to study every utterance, no matter how incomplete in terms of clausal syntax, and therefore it is often difficult to determine the relations between components.

5.1.3 The layered structure and elaboration

Besides the syntactic components themselves, the layered structure and the elaboration of components are aspects that are immediately apparent in the chart description. They have a very close connection to one another, since it is the elaboration that causes the layered structure to emerge.

After determining the syntactic components of an utterance each component is analysed to reveal its inner structure. The basic rule in the marking of elaboration in the chart is that each elaboration, syntactic or morphological, moves the marking down to the next layer. If there is no elaboration of any kind the entry on the top layer is enough to define the syntactic component and its part of speech. If the word is inflected the inflectional code is marked on the next layer. An example of this appears in Figure 18 (Section 5.1.2) where the first component is a NP consisting of an uninflected interrogative pronoun *kuka* 'who'. In Figure 18, the last syntactic component, NP *ponit* (poni-PL) 'ponies', includes morphological elaboration, and the plural inflection has an entry on the second layer. If there was yet another inflection in the same word, for example *pone-i-lle* (*pony*-PL-ALL) 'to the ponies', it would be placed on the third layer in the chart (Figure 19).

1.	NP N
2.	pl
3.	all

Poneille.
Pone-i-lle.
Pony-PL-ALL.
'To the ponies.'

FIGURE 19 An example of a syntactic component including two inflectional suffixes.

What is considered to be morphological elaboration is a matter of general principles on one hand and each child's individual skills on the other hand. The marking of morphological elaboration in Utterance Analysis is solely based on the earlier morphological coding of the data (see Section 3.1.1), and only those word forms which were earlier coded as inflected receive an entry showing morphological elaboration in Utterance Analysis.

In coding syntactic elaboration in the chart, the basic principles are similar to those for morphological elaboration: each new kind of elaboration moves the marking to the next layer. If a syntactic component consists of one word only there is no need for a further entry to mark syntactic elaboration, but when a component consists of two or more words each word has a separate entry on the next layer. It is common for syntactic and morphological elaboration to appear together within the same phrase. In these situations the syntactic elaboration has the first entry, and the morphological elaboration of the members of the syntactic structure is marked in the following layers of the chart. In Figure 20 a case of simultaneous morphological and syntactic elaboration of phrases is illustrated. The simultaneous use of both devices enables the phrases to reach the third layer in the chart.

1.		V		NP
2.	AUX:NEG	V	PRO:INDEF	N
3.		neg	partit	partit

Eei *kuulum* *mitää* *ääntä.*
Ei kuulu mi-tään ään-tä.
Not be heard-NEG any-PARTIT sound-PARTIT.
'There is not any sound.'

FIGURE 20 Simultaneous elaboration with syntactic and morphological devices within phrases. (35 Seppo; 21/167)

The morphological elaboration in Utterance Analysis is consistent with the earlier morphological coding of the data, whereas the syntactic elaboration requires a more detailed description of the procedure, since it was not given attention in the earlier phases of the coding. Next, I will briefly demonstrate how multiword phrases and other structures such as verb constructions are defined for treatment in Utterance Analysis.

5.2 Multiword structures

5.2.1 Noun phrases (NP)

Noun phrases are probably the most frequently occurring phrases in children's utterances in the data. In particular, the frequent use of pronouns in conversations multiplies the number of noun phrases. However, if the focus is on multiword noun phrases, cases with a pronoun as the head word are very rare, since pronouns tend to occur alone without extensions or qualifiers of any kind (ISK 2004: 563). When it comes to phrases with a noun as the main word, complements and qualifiers are more common and constitute multiword phrases together with a noun. A prototypical multiword noun phrase consists of a noun and a qualifying adjective which formally agrees in case and number with the head word. Another common type of multiword NP is a combination of two nouns where the first expresses a possessor and the second what is possessed. Figure 21 provides an example of a noun phrase consisting of an adjective and a noun.

1.		NP	V	ADVP ADV:DEM		PP
2.	ADJ	N			N	AD
3.		clit				ill
4.						

Pikku poniki menee tuone ihää viejee.
Pikku poni-kin menee tuonne isä* viere-en.
Little pony-CLIT goes over there father beside-ILL.
'Also the little pony goes over there, beside the father.'

FIGURE 21 The first NP represents a prototypical type of noun phrase: a noun with a qualifying adjective. In the adposition phrase, the * after the word *isä* 'father', signals a missing genitive ending. (24 Sanni; 2/865)

In spoken language in general and in conversations in particular, demonstrative pronouns are very frequently used in noun phrases as pronominal determiners. The data was recorded in a situation which especially supports pronouns such as *tämä* 'this' and *tuo* 'that'. The situational use of these pronouns refers to things that can be perceived by seeing, hearing or touching and they are either close (*tämä*) or remote (*tuo*) from the speaker's point of view (Larjavaara 2001a: 17–19). In the play session both of these pronouns are used frequently when children describe what they are doing with the toys and when they point at the toys around them. This use of pronouns is illustrated in Figure 22. It has been argued elsewhere that the use of the demonstrative pronouns *se* 'it', *ne* 'they', *tämä* 'this', and *tuo* 'that' in spoken Finnish have gradually come to resemble definite articles in many Indo-European languages (ISK 2004: 564). Whether they are actually articles or pronouns with some article-like features is a question of an ongoing debate, especially in the case of the pronoun *se*. For example, Laury (1996, 1997) strongly proposes that *se* 'it' is an article in spoken Finnish although it is not yet fully grammaticized, whereas Larjavaara (2001b) takes a more conservative stand in emphasizing the differences between the use of the Finnish *se* and, for example, the English definite article *the*. Nevertheless, these conversational uses of demonstrative pronouns as pronominal determiners or article-like parts of a structure create a considerable number of syntactically elaborated noun phrases in the data.

1.		NP	V	ADVP ADV:DEM
2.	PRO:DEM	N		ill
3.				

Tää tyttö menee tähän.
Tämä tyttö menee tä-hän.
This girl goes here-ILL.
'This girl goes in here.'

FIGURE 22 The demonstrative pronoun *tämä* 'this' as a pronominal determiner in a noun phrase. (31 Aleksi; 18/285)

Another possible way to form a noun phrase is to co-ordinate two or more equal words with a conjunction. Sometimes children use this to express plurality instead of using plural inflections, creating rather long phrases with a repetitious structure similar to the one presented in Figure 23.

ADVP ADV	V					NP				
		N	CONJ	N	CONJ	P:INDEF	CONJ	P:INDEF	CONJ	P:INDEF
Tiin+	*o*	*isi*	*jä*	*äiti*	*ja*	*toine*	*ja*	*toine`*	*ja*	*toine.*
Siinä	on	isi	ja	äiti	ja	toinen	ja	toinen	ja	toinen.
Siinä	on	isi	ja	äiti	ja	toinen	ja	toinen	ja	toinen.
'There	is	daddy	and	mother	and	another	and	another	and	another.'

FIGURE 23 A noun phrase with a repetitive structure. (23 Anniina; 1/603)

There is yet another structure rather typical of spoken language that needs special attention. This is the use of locative proadverbs in connection with noun phrases, adverb phrases or adposition phrases which also express location. The so-called extensive locative proadverbs *täällä* 'in here', *täältä* 'from here', *tänne* 'to here', *siellä* 'in there', *sieltä* 'from there', *sinne* 'to there', *tuolla* '(in) over there', *tuolta* '(from) over there', and *tuonne* '(to) over there' resemble qualifiers which are in agreement with the head word. However, the proadverbs do not modify the following phrase because they actually have the same referent. Instead of being a part of the following phrase, the locative proadverbs form an appositive structure with it. In Figure 21, there is an example of an appositive structure consisting of an adverb phrase *tuonne* ('(to) over there'; a proadverb) followed an adposition phrase *isä* viereen* ('beside the daddy'). Note that the postposition *viereen* requires its modifiers to be in the genitive case, but here the genitive ending is missing, and thus the adposition phrase is morphologically incorrect. The appositional relationship involves two or more consecutive phrases; thus a phrase and a preceding proadverb are analysed as separate syntactic components. (ISK 2004: 1009–1010, 1248; Vilkuna 1996: 209–214.) The appositive structure is therefore not an example of syntactic elaboration.

To sum up, noun phrases are frequently occurring syntactic components in the data. They exist in several kinds of multiword constructions. The situational factor especially promotes the use of pronominal determiners in noun phrases as well as the use of demonstrative pronouns alone.

5.2.2 Adjective phrases (AP)

Adjective phrases are much less frequent than noun phrases. A prototypical occurrence consists of an adjective alone, multiword adjective phrases being rather rare in the data. The few occurrences of multiword adjective phrases in the children's utterances represent three different types, two being frame-like structures or idioms.

The frame-like constructions are used in connection with only a few adjectives. The first special case is a phrase with an adjective as the head word and a qualifying infinitive phrase. The infinitive expresses the action being characterized while the adjective itself expresses the characterization (ISK 2004: 613). Not all adjectives are suitable for describing actions. In my data, only the adjective *vaikea* 'difficult' occurs in this structure (Figure 24).

1.	NP PRO:DEM	V	AP	
2.			ADJ	V
3.				1inf

Se+ *on:* *vaikee* *tehdä*
Se *on* *vaikea* *teh-dä.*
It is difficult do-1INF.
'It is difficult to do.'

FIGURE 24 An adjective phrase consisting of an adjective and an infinitive verb qualifying the adjective. (1 Tuija; 65/1125)

The other special idiomatic structure is used to characterize an object by means of observations through the speaker's senses. This particular requirement limits possible adjectives to very few examples, such as *makuinen* '-flavoured', *näköinen* '-looking, looks like', *kuuloinen* 'sounds like' and *tuntuinen* 'feels like'. In this data, only *makuinen* and *näköinen* are used (Figure 25).

1.	NP PRO:DEM	V		AP
2.			ADJ	ADJ
3.				partit

Ta+ *om* *paha* *makutta*
Tämä on paha* makuis-ta.
This is bad flavoured-PARTIT.
'This tastes bad.'

FIGURE 25 An adjective phrase expressing observations through the speaker's senses. The asterisk (*) after the word *paha* 'bad' signals a missing genitive ending. (38 Pekka; 47/1728)

The only multiword adjective phrase in the data with more productive possibilities than the frame-like structures described above is one constructed of an adjective qualified by an adverb. In the example in Figure 26, this kind of adjective phrase is a part of a larger noun phrase. However, there are not many examples of such combinations present in the data.

1.	V			NP
2.	pass		AP	N
3.		ADV	ADJ	

Otetaa: *oik`e* *p6ieni* *puhelin .*
Ote-taan oikein pieni puhelin.
Take+PASS really small telephone.
'Let's take a really small telephone.'

FIGURE 26 An adjective phrase situated within a noun phrase. (23 Anniina; 57/1156)

To sum up, adjective phrases are not very common in the data in general, and it is even rarer to have a syntactically elaborated, multiword adjective phrase. Moreover, two of the multiword phrase-types in the data represent idiomatic structures with semantically restricted possibilities for productivity.

5.2.3 Adposition phrases (PP)

In Finnish, there are both pre- and postpositions, as well as some particles that can be used in both adpositional positions. However, the change of a postposition into a preposition may cause stylistic changes in the meaning of the phrase (for example, into a more poetic style). Sometimes the borderline between adposition phrase and other phrases is not very clear, since words used as adpositions are ambiguous as to what word class they actually represent (ISK 2004: 687).

Basically, an adposition phrase is always a syntactically elaborated structure because an adposition requires a noun phrase as a complement. Thus in this analysis, a clear syntactically elaborated multiword structure is the only possible manifestation of an adposition phrase. All other instances – ambiguous multiword structures or occurrences of single words identical to the words used as adpositions – are excluded from this category and analysed on other bases, mostly as adverbs. The complementing noun phrase within an adposition phrase can of course be a multiword structure itself. Therefore an adposition phrase may grow into a rather complex structure of several layers and multiple elaborations since adpositions require their modifiers in specific inflected forms, usually the genitive or partitive case. Figure 27 shows an example of a multilayered adposition phrase.

1.	NP PRO:DEM	V				PP
2.					NP	AD
3.			PRO:DEM	PRO:INDEF	N	
4.			gen	gen	gen	
	Tää	*polkee*	*ton*	*yhen*	*tytön*	*luo.*
	Tämä	polkee	tuo-n	yhde-n	tytö-n	luo.
	This	cycles	that-GEN	one-GEN	girl-GEN	to.

'This one cycles over to that (one) girl.'

FIGURE 27 An example of the multilayered structure of an adposition phrase. (31 Aleksi; 2/151)

Adposition phrases rarely appear in the current data. However, when such a phrase occurs, it is always a syntactically elaborated structure including specified morphological requirements, although children of this age do not always fulfil them accurately yet.

5.2.4. Adverb phrases (AdvP)

Alongside noun phrases, adverb phrases are the most frequently occurring structures in the data. In one utterance there may be several adverbs but multiword adverb phrases are not very common. In fact the few types found in the data are examples of special idiomatic expressions.

The first multiword adverb structure worth noting is *pois jostakin* 'away from something, off something,' where the adverb *pois* forms a phrase together with a nominal word (*pois auto-sta*, 'out of a car') or proadverb (*pois täältä* 'away from here') inflected with local cases of motion away (elative or ablative case). *Pois* 'away' may occur also with an adposition phrase (*pois pöydä-n alta* 'away from under a table'). Whether the extension is a complement or a qualifier is not always clear but nevertheless, the combination is a phrase.

Another type of multiword structure with an adverb as the head word is a combination where the particles *noin* ('in that way, so'), *näin* ('in this way'), and sometimes even *niin* ('so') are used to emphasize the meaning of a proadverb in the head word position. These particles sometimes occur in the same function with demonstrative pronouns as well, in which case the structure is a noun phrase. Such use of these particles is typical of spoken language in connection with deictic and demonstrative words. Such a phrase functions as a verbal pointer instead of – or in addition to – a physical pointing with a finger or nod of the head in the direction in question. An example of such a multiword adverb phrase is presented in Figure 28.

1.		NP	V	ADVP	
2.	PRO:DEM	N		ADV:DEM	PTL

He	*poika*	*menee*	*tuonnen*	*noi.*
Se	poika	menee	tuonne	noin.
The	boy	goes	there	+ an emphasizing particle

'The boy goes over there.'

FIGURE 28 A multiword adverbphrase including an emphasizing particle. (11 Jaakko; 46/529)

5.2.5 Verb constructions (V)

There is no clear evidence for the existence of a verb phrase in Finnish (ISK 2004: 433). Therefore finite verbs are considered to be individual components of an utterance. Syntactical elaboration of verbs can be shown in verb components in more versatile ways than in the phrases discussed so far. The verb constructions relevant to the present study are verb chains, verb alliances, verb idioms and phrasal verbs[8]. There are also other constructions which are treated in the Utterance Analysis as syntactically elaborated verb constructions although they are very often included under the inflectional paradigm of verbs. These inflectional forms are the perfect and past perfect tense and negative structures, all of which are verbal forms consisting of an auxiliary (*olla* 'to be' and *ei* 'no, not') and a head verb.

A verb chain consists of a modal or some other abstract verb and an infinitive. The following verbs occurring in the data form verb chains with infinitives: *pitää* 'have to, must', *tarvita* 'need', *täytyä* 'have to, must', *alkaa* 'begin', *ehtiä* 'have time to do sth', *meinata* 'mean to do sth', *saada* 'may', *saattaa* 'may', *tahtoa* 'want', *taitaa* 'be likely to, may', and *voida* 'can, be able to'. Such

8 The terminology used in analyzing verb constructions originates from Iso suomen kielioppi (2004), and it may differ from the terminology used in general linguistics.

verbs as *aikoa* 'intend', *haluta* 'want', *yrittää* 'try', and *tottua* 'get used to' may also form constructions which can be interpreted as verb chains. However, according to another interpretation, any infinitive phrase following these verbs functions as a complementary object or adverb. In the Utterance Analysis the latter parsing is adopted and therefore infinitive verbs together with possible extensions following the verbs *aikoa*, *haluta*, *yrittää,* or *tottua* are interpreted as examples of infinitive phrases. (ISK 2004: 493–495.)

A verb alliance ('verbiliitto') differs from a verb chain in several aspects, even though it is also a combination of two verbs with the second one being an infinitive. Verb alliances are idioms, with an initial auxiliary (*olla* 'to be' or *tulla* 'to come'); together the verbs create a meaning which cannot be inferred on the basis of each verb alone. For example, the verb alliance *tulla tekemään* (literarily 'come to do sth') has a future meaning 'will do', although neither verb has such a meaning by itself. (ISK 2004: 443–446.) The most common type of verb alliance in the data is *olla tekemässä* (to be + do-3INF-INESS), which refers to an ongoing action and corresponds approximately to the English present tense form *is doing.*

A verb idiom is a combination of a verb with a specified extension. Some general verbs such as *olla* 'to be', *ottaa* 'to take', *mennä* 'to go', *tulla* 'to come', and *antaa* 'to give' are prone to idiom formation. For example *ottaa* 'to take' participates at least in the following idioms *ottaa huomioon* 'show consideration', *ottaa lukuun* 'take into account', *ottaa opikseen* 'learn from sth', and *ottaa selvää* 'find out'. Verb idioms differ from other combinations containing verbs in that the specified extension is usually attached to only one particular verb, or the meaning of the combination is separate from the use of each individual component of an idiom. (ISK 2004: 447.)

Phrasal verbs are combinations of verbs and particles. These combinations are so tight that very often they form compound words, especially in the 1st infinitive (*antaa ilmi* ~ *ilmiantaa* 'denounce, turn in'). Moreover, compound nouns can also be derived from phrasal verbs (*ilmianto* 'denunciation').

It is not very important to differentiate between phrasal verbs and verb idioms, since both are constructed of two parts, and thus behave similarly in the Utterance Analysis. Moreover, the border between a co-occurrence of a single verb and a particle and either a verb idiom or a phrasal verb is not always that clear. How tight must the bond between a verb and a particle be to qualify individual words occurring in the same utterance as a phrasal verb? Are the verb idioms and phrasal verbs in a child's language similar to those of an adult? Do children use some combinations that in adult language are separate units in an idiomatic way or always in a tight connection with each other? In order to avoid such speculation only the most typical cases of the combination types are accepted here as representatives of either verb idioms or phrasal verbs. This is not a very open or creative way to do the analysis, but in the case of cross-sectional data it is a safe choice.

The perfect and past perfect tense as well as negative verb expressions are part of the inflectional paradigm. However, here they are analysed the same way as syntactic elaborations consisting of at least two components. These verb structures are combinations of an auxiliary verb and a main verb. In the perfect and past perfect tense, the auxiliary is *olla* 'to be', inflected with the appropriate person and tense. In the negative forms, which are relatively frequent in the data, there is the negative auxiliary *ei* 'no, not' inflected in the appropriate person plus the main verb which occurs in a special negative form in the present tense and in infinitive forms in the other tenses. Negation is a powerful device of elaboration in several ways. It turns any verb into a multilayered structure and at the same time opens up possibilities for many kinds of inflection. If the negative auxiliary is attached to a verb chain or other multi-verb construction the result is a syntactically and often morphologically multilayered complex structure such as shown in Figure 29**.**

1.			V	CONJ	V
2.	AUX:NEG		V		pass
3.	1s	AUX	V		
4.		neg	part		

en	*oo*	*keittänyk*	*ku*	*orotetaa*
E-n	ole&NEG	keittä-nyt	kun	odote-taan.
No-1S	be&NEG	cook-PART	because	wait-PASS.

'I haven't cooked because we are waiting.'

FIGURE 29 The use of negation together with a multi-verb construction results in a complex combination where there are several slots for agglutinative inflection as well. (1 Tuija; 9/1008)

To sum up, verbs are a central place for both morphological and syntactic elaboration in utterances. It is possible for verbs to expand onto several layers and form complex components of an utterance.

5.2.6 Infinitive phrases (InfP)

An infinitive phrase is composed of an infinitive verb and its complements or qualifiers. However, the occurrence of an infinitive form does not automatically mean that there is an infinitive phrase in an utterance. When an infinitive is a component of a verb chain, the perfect or past tense forms, or a verb alliance, it is not analysed as an infinitive phrase (ISK 2004: 488, 836).

In a way, infinitive phrases resemble clauses more than other phrases do. This is due to the complements and qualifiers which form a similar relationship to the infinitive form as subject, object, and adverbial to a verb in a complete clause. The clause-like structure also makes it possible for an infinitive phrase often to grow larger than other phrases. An example of a large infinitive phrase is illustrated in Figure 30.

1.	NP PRO:DEM		V		NP N	INFP		
2.		V	V	ADV:DEM		3INF		NP
3.		clit	1inf			ill	N	N
4.								ill

Tämä poiki mennät tänne tyttö ukkumaa poika sääkyy
Tämä voi-kin men-nä tänne tyttö nukku-ma-an poika* sänky-yn.
This may-CLIT go-1INF here girl sleep-3INF-ILL boy* bed-ILL.
'This girl may indeed go to sleep here in a boy's bed.'

FIGURE 30 A large infinitive phrase with clause-like properties. (7 Taru; 1/320)

The infinitive forms mostly occurring in the data of the children in the present study are the 1st (A-) and 3rd (MA-) infinitives. An extended infinitive structure is not yet a part of every child's language but the 3rd infinitive forms such as *syömään* 'eat-3INF-ILL' and *nukkumaan* 'sleep-3INF-ILL' recurrently occur in every day conversation in families. It is not always clear whether these forms of these particular verbs are productively inflected or frozen units. They are usually the first infinitive forms in a child's repertoire, and often very frequently used before the occurrence of any other verbs in the 3rd infinitive form.

5.3 The view of complexity in Utterance Analysis

Utterance Analysis (UA) clearly represents an absolute approach to linguistic complexity, since utterance structure is evaluated without any reference to a language user's processes or experiences. Utterances or structures found in them are not used to evaluate whether the production process was difficult or not. UA is in some respect comparable to with MLU, since it takes each utterance into consideration and analyses it down to morphemic level. However, despite similarities in the theoretical background and the objective of the analysis, there are major differences between these methods. UA carefully specifies each morpheme and component, and shows the relationships between the morphemes within each component. The method draws a concrete picture of how the units within the structure of each utterance take their place on layers and how large structures are built out of smaller units and structures. Thus, UA sees an utterance as a multidimensional structure with units of different sizes and categories, possibly including some subunits as well. It does not force units of just one category (morphemes) into a linear order in the way that MLU does, and neither does it ignore connections between the units.

UA also has some similarities with IPSyn. Several items in IPSyn are identified as phrases in UA, that is, both IPSyn and UA do note that the major units in an utterance may consist of several subunits. The principle of cumulativity [9] in IPSyn and the layered structure in UA go well together.

[9] The principle of cumulativity in IPSyn refers to the overlap of the items specified: a structurally less complex item is also included in a more complex item. The overlap

However, IPSyn does not necessarily show which subunits a larger structural unit consists of, whereas UA ties them together in each utterance. This difference arises from the distinct objectives of these methods. IPSyn concentrates on diversity of resources whereas UA focuses on utterance diversity, and therefore the views of complexity are different. In IPSyn structural complexity is defined as the resources of distinct building blocks for building utterances but in UA it arises from the way the building blocks are combined to form a construction.

UA exhaustively analyses utterance structure, and each and every part of an utterance contributes to structural complexity. Dahl (2004: 48) differentiates two kinds of structure, namely choice structure and output structure. When defining choice structure he argues that "structural representations of utterances in linguistics can, by and large, be seen as specifications of structured sets of choices made by speakers relative to the language system". It is possible to think of UA as showing each moment of choice and specifying each choice made by a speaker. However, assembling each utterance from individual morphemes is not in conformity with Tomasello's (e.g. 2000a, 2000b, 2000c, 2003), Langacker's (e.g. 1987, 1988a, 1988b) or Peters' (1983) thoughts about how a human being processes language and the kinds of elements in which language is stored in brain. In their view it is not psychologically plausible that linguistic expressions are produced simply be using abstract rules and the smallest possible meaningful units. It is more likely that human beings build their utterances from various kinds of language elements: rote learned units, partially analysed frame structures, and abstract schemes. Tomasello (2000c: 77) describes the processes leading to full range utterances as "cut and paste" syntactic strategies. According to usage-based models and cognitive grammar, the choices made by a speaker are not choices of individual words and morphemes but of various kinds of frozen or semi-frozen structures of different sizes, and only the occasional application of abstract rules. Therefore it is clear that, rather than choice structure, UA describes output structure in terms of abstract linguistics.

Dahl (2004: 39) strongly suggests that complexity should be kept as an information-theoretic notion, unrelated to a user or an agent. UA observes this principle. It concentrates on the multidimensionality of each utterance and emphasizes its output structure. Structural complexity arises from co-operation between the forms and phrases related to each other on several layers. Thus, it is the breadth of description that mostly manifests the complexity. In a way, UA puts together the results obtained from MLU and IPSyn because it focuses on the whole structure of utterances (like MLU) and at the same time specifies the smaller and larger linguistic units included in each utterance. UA is not a picture of an individual's actual linguistic units or choices of structures but a description of the result in general linguistic terms.

---

of the items ensures that several stages of development can be traced. In the case of, for example, the 3rd infinitive the least complex item is the verb form specified alone (*syömään* 'eat+3rd inf.'; item V8 in IPSyn), a more complex item including the foregoing is a structure with two verbs, the second in the 3rd infinitive (*mennä syömään* 'to go to eat+3rd inf.'; V13) and finally, the most complex item including both of the previous items, is one with three verbs (*haluan mennä syömään* 'I want to got to eat+3rd inf.'; V14).

6 INDEPENDENT SYNTACTIC COMPONENTS

It has been argued that structural complexity grows both horizontally and vertically. Horizontal growth arises from an increase in the number of syntactic components, and vertical growth from an increase in elaboration on more layers within the components. In this chapter the focus is on the top layer of the utterances: the number and quality of the components. The data comes from children in subgroups A, B and C, and consists of 720, 1040 and 720 utterances respectively.

6.1 The number of syntactic components

Hypotheses 1, 2 and 3 (Section 4.4.2) refer to the number of syntactic components. It is hypothesized that in group A the number of components remains small whereas in group C there are no restrictions on how many components there may be in an utterance. The syntactic components are specified on the top layer in the UA chart, and they represent different phrases (NP, AP, INFP, ADVP, and PP), verb constructions (V), conjunctions (CONJ), particles (PTL), or interjections (INTERJ).To investigate the number of syntactic components in their utterances, all the utterances produced by the children in subgroups A, B, and C were classified according to the number of syntactic components they contain. In addition, the average number of utterances, the range of occurrences, and the median value were calculated within each of the categories. This procedure built up a general view of the composition of the utterances. The results of this analysis are presented in Tables 11–13**.**

The range in the number of utterances in each category varies substantially, especially in subgroups A and B, though in group C there is less variation. For example, in group B the number of three-component utterances ranges from 24 utterances produced by Mika (ID 15) to 58 in Pekka’s (ID 38) data. Similarly, in group A the range of three-component utterances is from 19 (Saku, ID 18) to 52 (Anna, ID 22).

TABLE 11 The utterances (N= 720) produced by the children in subgroup A in terms of number of syntactic components.

	Number of syntactic components in utterances									
Subjects	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
6 Ronja	7	33	31	9	0	0	0	0	0	80
17 Tatu	8	22	41	8	1	0	0	0	0	80
18 Saku	19	38	19	3	1	0	0	0	0	80
22 Anna	0	21	52	6	1	0	0	0	0	80
29 Elisa	2	27	41	8	0	1	0	0	1	80
31 Aleksi	16	19	35	9	1	0	0	0	0	80
34 Leo	4	26	32	12	5	1	0	0	0	80
35 Seppo	21	26	26	6	1	0	0	0	0	80
36 Sampo	13	34	28	4	1	0	0	0	0	80
Total	90	246	305	65	11	2	0	0	1	720
%	12.5	34.2	42.4	9.0	1.5	0.3	0.0	0.0	0.1	100.0
Range	0–21	19–38	19–52	3–12	0–5	0–1	0	0	0–1	
Median	8	26	32	8	1	0	0	0	0	
Mean	10.0	27.3	33.9	7.2	1.2	0.2	0.0	0.0	0.1	

TABLE 12 The utterances (N= 1040) produced by the children in subgroup B in terms of number of the syntactic components.

	Number of components in utterances									
Subjects	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
2 Riikka	1	12	45	19	3	0	0	0	0	80
7 Taru	0	7	41	25	3	2	2	0	0	80
8 Aino	3	23	45	8	1	0	0	0	0	80
10 Santeri	1	8	47	17	4	1	1	0	1	80
11 Jaakko	0	2	34	33	10	1	0	0	0	80
13 Joel	8	23	42	5	2	0	0	0	0	80
15 Mika	14	22	24	14	2	3	0	0	1	80
24 Sanni	3	13	46	15	3	0	0	0	0	80
27 Janna	1	6	37	24	5	2	2	3	0	80
31 Aleksi	16	19	35	9	1	0	0	0	0	80
34 Leo	4	26	32	12	5	1	0	0	0	80
35 Seppo	21	26	26	6	1	0	0	0	0	80
38 Pekka	0	3	58	15	4	0	0	0	0	80
Total	72	190	512	202	44	10	5	3	2	1040
%	6.9	18.3	49.2	19.4	4.2	1.0	0.5	0.3	0.2	100
Range	0–21	2–26	24–58	5–33	1–10	0–3	0–2	0–3	0–1	
Median	3	13	41	15	3	0	0	0	0	
Mean	5.5	14.6	39.4	15.5	3.4	0.8	0.4	0.2	0.2	

TABLE 13 The utterances (N= 720) produced by the children in subgroup C categorized in terms of syntactic components.

	Number of syntactic components in utterances										
Subjects	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**11**	**Total**
1 Tuija	1	12	36	18	10	2	1	0	0	0	80
2 Riikka	1	12	45	19	3	0	0	0	0	0	80
3 Laura	0	8	46	22	4	0	0	0	0	0	80
7 Taru	0	7	41	25	3	2	2	0	0	0	80
11 Jaakko	0	2	34	33	10	1	0	0	0	0	80
19 Tuomo	1	6	47	22	4	0	0	0	0	0	80
23Anniina	1	7	40	22	3	2	2	1	1	1	80
27 Janna	1	6	37	24	5	2	2	3	0	0	80
30 Elina	1	7	36	27	7	1	1	0	0	0	80
Total	6	67	362	212	49	10	8	4	1	1	720
%	0.8	9.3	50.3	29.4	6.8	1.4	1.1	0.6	0.1	0.1	99.9
Range	0–1	2–12	34–47	18–33	3–10	0–2	0–2	0–3	0–1	0–1	
Median	1	7	40	22	4	1	1	0	0	0	
Mean	0.7	7.4	40.2	23.6	5.4	1.1	0.9	0.4	0.1	0.1	

Another factor shared by subgroups A and B is that utterances consisting of five or more syntactic components are marginal. Only 1.9 % of all utterances (14 out of 720) in group A and 6.2% (64 out of 1040) in group B have five or more components. In subgroup C the situation is somewhat different, since the marginal categories are those of six or more components and the single-component category. Altogether, 4.1 % of the utterances (30 out of 720) belong to these categories in group C. However, the fact that single-component utterances are so marginal does not mean that children in group C do not produce such utterances. This is a concomitant of the data selection. When only the 80 longest utterances from each child are considered, single-component utterances[10] are not selected, since these children have enough multicomponent utterances to make up the sample of the 80 longest utterances.

A major feature in common between the groups is the dominance of three-component utterances. Of all the utterances, 42.4% in group A, 49.2% in group B, and 50.3% in group C consist of three syntactic components. However, some noteworthy differences emerge when the other major categories are considered. The distribution of utterances in component categories presented in Figure 31 shows that 34.2% of utterances in group A consist of two syntactic components, whereas in group C the second most significant category is four-component utterances with a proportion of 23.6%.

10 The single-component utterances are not necessarily mono-morphemic. A single component may be elaborated in different ways, resulting in a multi-morphemic utterance. However, the dividing line between the 80 longest utterances and shorter productions is at such a high level in group C that even a multifaceted elaboration does not guarantee that a single-component utterance will be selected.

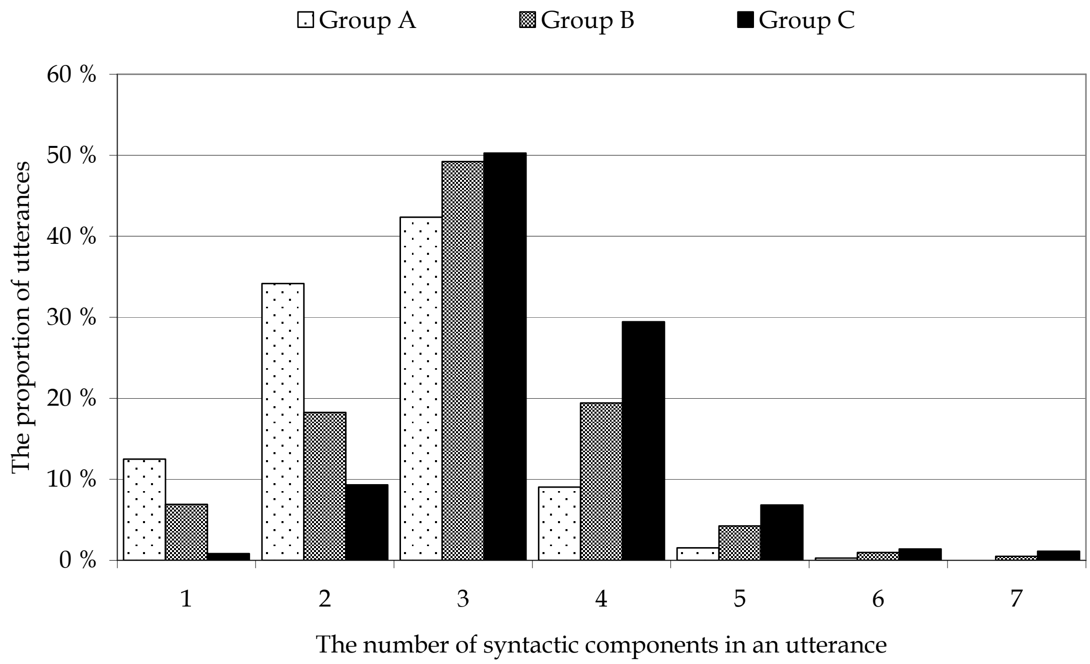


FIGURE 31 The proportional distribution of utterances in subgroups A, B, and C on the basis of number of components. The most marginal categories, utterances consisting of eight to eleven components, are omitted.

In group B, which forms a bridge between the others, apart from three-component utterances two categories are equally represented, the two- and four-component categories with 18.3% and 19.4% respectively.

The distribution of utterances into distinct categories on the basis of the number of their syntactic components forms a picture of a gradual shift from utterances mostly containing two or three syntactic components in group A, to utterances mostly made up of three or four components in group C. The children in group B, producing equally two- and four-component utterances together with the predominant three-component ones, form a bridge between groups A and C just as they did in respect of MLU values and IPSyn scores.

The growth in the number of syntactic components partly explains the growth in MLU values, but obviously there must be other factors affecting this outcome too. Otherwise the group differences would be much more dramatic than the smooth transition detected in the data. It was hypothesized that in group A the number of components remains low (Hypothesis 1 and 2) whereas in group C new components are added. The quantitative results of MLU and IPSyn show no restricting factors that would affect either the horizontal or vertical development in complexity (Hypothesis 3 and 5). The results presented here are consistent with the hypotheses: they do show an increase in the number of components when groups A and C are compared. However, the magnitude of the increase turned out to be much smaller than expected: group C is not far ahead the other groups. These results suggest that growth incomplexity concentrates on the elaboration of components more than increasing their numbers. How this happens will be examined later. Before that, however, more attention must be given to the component types the utterances consist of.

6.2 The quality of the components

6.2.1 The combinations of components in utterances

After analysing the number of syntactic components in utterances, more detailed attention is now given to the type of components and the combinations of them. The aim is to investigate whether component quality varies between the groups and affects the observed variation in structural complexity. The utterances were categorized according to their component combinations, and comparisons were then made between groups A, B, and C. The results for the utterances consisting of one to five syntactic components are presented in Table 14. Utterances with six or more components are excluded from this comparison because of their marginal proportions in the data (0.42%, 1.92% and 3.33% in the groups A, B and C respectively).

Table 14 shows that verb structures (V), noun phrases (NP), and adverb phrases (AdvP) are the most common components in the utterances of all three groups and in all combination lengths. Other components (adposition phrase PP, adjective phrase AP, and conjunction CONJ) play only a minor role in the most common combination types. Only the infinitive phrase (INFP) has a more significant role in groups B and C. Similarity is evident in the most frequent component combinations in all the groups.

In all three groups the three most common combination types account for the vast majority of all utterances in each category as defined by number of components. This is especially striking in the case of one- to three-component utterances. For example, in group A, the most frequently occurring combinations (V + NP, V + AdvP, and NP + AdvP) account for 79.7% of all two-component utterances (N = 246). In groups B and C the combined proportion of the same combinations is 79.0 % out of 190 tokens, and 80.6 % out of 67 tokens respectively. In the case of three-component utterances, the predominance of the given combination types is even more evident in groups B and C. It is obvious that when there are only a few components in the utterances the possibilities for different combinations are substantially smaller than when there are more. With an increase in alternatives for combination it is logical that there should be fewer favourite combination types and therefore the utterances may be distributed more equally across combination types. The proportions of the three most common combination types in four- and five-component utterances, especially in groups B and C, reflect this strongly. The combined proportion of the three favourite types has decreased in group B from 83.9% in three-component utterances to 57.4% in four-component utterances, and in group C from 82.6% to 54.7% respectively. A parallel development continues for five-component utterances with proportions decreasing to 47.7% (B) and 38.8% (C).

TABLE 14 The three most common combination types in utterances containing one to five components in groups A, B and C. The table shows the number and proportion (%) of tokens of each combination type for each of the groups. The number of single-component utterances in group C and five-component utterances in group A is so low that the proportions are not calculated (marked with * in the percentage column).

Combination types	**Tokens**	%	**Tokens**	%	**Tokens**	%
Single-component utterances	**90**		**72**		**6**	
V	33	36.7%	21	29.2%	1	*
NP	31	34.4%	24	33.3%	3	*
ADV	17	18.9%	15	20.8%		
PP					1	*
INFP					1	*
Combined proportion of the		**90.0%**		**83.3%**		*
Three favourite types						
Two-component utterances	**246**		**190**		**67**	
V + NP	132	53.7%	106	55.8%	40	59.7%
V + ADVP	32	13.0%	23	12.1%	9	13.4%
NP + ADVP	32	13.0%	21	11.1%	5	7.5%
Combined proportion of the		**79.7%**		**79.0%**		**80.6%**
Three favourite combinations						
Three-component utterances	**305**		**512**		**362**	
V + NP + ADVP	120	39.3%	165	32.2%	145	40.1%
V + NP + NP	95	31.1%	203	39.6%	108	29.8%
V + NP + AP	22	7.2%				
V + NP + INFP			62	12.1%	46	12.7%
Combined proportion of the		77.7%		83.9%		82.6%
Three favourite combinations						
Four-component utterances	**65**		**202**		**212**	
V + NP + NP + ADVP	21	32.3%	52	25.7%	52	24.5%
V + NP + ADVP + ADVP	12	18.5%	44	21.8%	43	20.3%
V + NP + NP + NP	9	13.8%	20	9.9%	21	9.9%
Combined proportion of the						
Three favourite combinations						
Five-component utterances	**11**		**44**		**49**	
V + NP + NP + ADVP + ADVP	4	*	11	25.0%	8	16.3%
V + NP + NP + NP + ADVP	2	*	6	13.6%	7	14.3%
V + V + NP + NP + ADVP	1	*				
V + V + NP + ADVP + ADVP	1	*				
V + NP + NP + NP + NP	1	*				
V + NP + NP + ADVP + CONJ	2	*				
V + V + NP + NP + NP			4	9.1%	4	8.2%
Combined proportion of the		*		**47.7%**		**38.8%**
Three favourite combinations						

The investigation of favourite combination types of syntactic components did not reveal any striking differences between the three groups. The majority of utterances follow quite similar patterns in all groups. It seems that members of each group make similar use of opportunities to build their utterances. The basic combination patterns are present from very early on. The differences arise in the case of four- to five-component utterances, because preferences for given combination types are no longer as clear as in the utterances with fewer components, and the composition of combinations may vary substantially.

6.2.2 The distinct syntactic components and structural complexity

So far, groups A, B, and C appear to behave in a fairly similar way when the composition of utterances is considered. In number of syntactic components, there is a smooth transition from a preference for two- to three-component utterances in group A to a preference for three- to four-component utterances in group C, while group B has a bridging role with properties similar to both the other two groups. The groups showed even more similarity in combining components into utterances when the most preferred combination types were considered. Therefore it appears that syntactic components alone are responsible for only a small proportion of the differences in the structural complexity of utterances. However, there is still one aspect of components worth considering, namely the different potential of each component to be elaborated.

When elaboration potential is considered, the component types specified in the Utterance Analysis can be divided into three categories. First, there are components which are not elaborated, or for which elaboration is very rare. Particles, interjections and conjunctions belong to this category. Particles and interjections form a heterogeneous group of pragmatic components and routine expressions (for example, *kiitos* 'thank you', *hei-hei* 'bye-bye'). They do not require the existence of other components beside them, whereas conjunctions function by themselves as connectives between clauses and phrases. Another important function especially of the conjunction *ja* 'and' is to connect consecutive turns to each other and connect speech to other ongoing activities. In narrative texts *ja* 'and' functions as a connective between the parts of the narrative. In these functions, *ja* is no longer a conjunction but a multipurpose particle. [11](ISK 2004: 778.) In children's speech during a play session *ja* is often used as a starter of an utterance, especially when a child comments on actions in the play, as in Example 4:

---

11 Despite different functions all occurrences of *ja* 'and' are categorized as conjunctions in this study.

Example 4

*CHI: *ne voi ollat tässä ## tässä tuolilla.*
%eng: they can be here ## here on the chair.
%act: puts a cup on a chair
*EXP: *mm+m .*
*CHI: ***ja*** *tää <voi ollat tä`ssä> [=! whispering] .*
%eng: **and** this one can be here.
%act: puts a plate on a chair
*EXP: *mm+m -? .*
*CHI: ***ja*** *tää voi ollav vaikka vaikka vaikka <pöy`ällä> [=! whispering] ##.*
%eng: **and** this one can be, for example, on the table ##.
%act: takes a jug and puts it on the table

(4 Henna; 287–297)

The second category includes components with optional elaboration, that is, there are no obligatory requirements for either morphological or syntactic elaboration, but each component allows both. The components belonging to this group are the noun phrase (NP), adjective phrase (AP), adverb phrase (ADVP), and verb construction (V). However, they vary substantially in their potential for elaboration. Noun phrases and verb constructions have the greatest potential for elaboration. For NPs all forms of nominal inflection are used and, in addition, the head of a NP can be modified with, for example, an adjective or a demonstrative pronoun. Because of case agreement between a head and a modifier in a noun phrase the final result may easily attain a three- or four-layered structure including both syntactic and morphological elaboration. For verb structures, a wide range of inflection in person, tense and mood is available. Moreover, syntactic elaboration of verb structures is very common in the negative form, verb chains, phrasal verbs and verb alliances. The simultaneous occurrence of syntactic and morphological elaboration results in three- to four-layered structures in the same way as in noun phrases. Thus, both noun phrases and verb constructions may vary from an unelaborated word to a multilayered, wide structure with both morphological and syntactic expansion, as illustrated in Examples 5a–d:

Example 5

a) Unelaborated NP and V

1.	NP PRO:DEM	V

	Nää	*lentää.*
	Nämä	lentää.
	These	fly
	'These are flying.'	

(31 Aleksi; 58/186)

b) Three-layered V and NP

1.		V		NP
2.	AUX:NEG	V	A	N
3.		neg	ill	iness

Ei *nuku* *omaaN* *käNNyssä.*
ei nuku&NEG oma-an* sängy-ssä.
Not sleep&NEG own-ILL bed-INESS
'(He/she) is not sleeping in his/her own bed.' (8 Aino; 7/266)

c) Four-layered NP

1.			NP
2.		AP	N
3.	ADV	PRO:ADJ	pl&partit
4.		pl&partit	

Ihan *saman`eisia* *aitoja.*
Ihan samanlais-ia aito-ja.
Exactly similar-PL&PARTIT fence-PL&PARTIT.
'Exactly similar fences.' (13 Joel; 10/499)

d) Four-layered V together with three-layered and two-layered NPs

1.	NP N			V		NP
2.	pl		V	V	PRO:ADJ	N
3.		AUX:NEG	V	1inf	partit	partit
4.			neg			

Vauvat *ei* *voi* *syyät* *tämmöstä* *ruokaa .*
Vauva-t ei voi&NEG syö-dä tämmöis-tä ruoka-a.
Vaby-PL not can&NEG eat such-PARTIT food-PARTIT.
'Babies can't eat this kind of food.' (3 Laura; 1/980)

Compared with noun phrases and verb constructions, adjective and adverb phrases are far more restricted in their potential for elaboration, although morphological and syntactic devices are allowed in extension. Adjectives do have similar inflectional forms to the other nominal words (nouns, pronouns, and numerals), and their modification with syntactic devices is possible as well. However, because our focus is now only on independent syntactic components the possibilities for an independent adjective phrase to occur in an utterance do restrict elaboration[12]. The most probable position for an independent AP to occur is together with a copula in a characterizing clause (e.g. *Tämä on kaunis* 'This is beautiful'). In this position an adjective phrase may be in the nominative or partitive case only, which substantially reduces the inflectional potential. For modification within adjective phrases, it is common to use intensifiers, which are typically frozen adverbs such as *tosi* 'really' and *ihan* 'completely; quite'. (An example can be found in Example 5c, although it is not an independent AP-component.) Therefore, although theoretically behaving similarly to noun phrases, independent adjective phrases have far fewer possibilities to grow into

12 It is common to have an adjective or an adjective phrase modifying a noun as in example 5c. However, in this position, an adjective is not an independent component, since it is a part of a noun phrase.

three- to four-layered structures, due to the restrictions on placing these phrases in an utterance.

In the case of adverb phrases there are also morphological restrictions on their elaboration, but they are of a different origin. Adverbs are typically either uninflected, or have an incomplete inflectional paradigm. For example, the locative proadverbs are inflected only in the local cases. Moreover, there are some syntactic restrictions in expanding an adverb phrase. Deictic adverbs, such as *tänään* 'today', *täällä* 'in here' and interrogative adverbs, such as *missä* 'where' and *miksi* 'why', cannot occur with modifiers, for semantic reasons. (ISK 2004: 638.) Thus, it seems that a typical adverb phrase is an unelaborated single adverb. However, during a play session, typical utterances including adverb phrases refer to locality, which gives more opportunities for inflected proadverbs to occur. Another possible way to elaborate adverb phrases morphologically is to use clitic particles. Elaboration of adverb phrases is therefore not impossible, though limited.

The third category of elaboration potential involves adposition phrases (PP) and infinitive phrases (INFP) as components requiring elaboration. An adposition phrase is always a combination of an adposition and a noun phrase. Additionally, adpositions require their complements to be in a specific inflectional form, although an appropriate form is not yet always available in children's language. In the case of adposition phrases there is therefore automatically at least syntactic elaboration. Morphological elaboration may also occur in accurately produced utterances such as that in Example 6.

Example 6

1.	CONJ	NP PRO	ADVP ADV	V			PP
2.				past		NP	AD
3.				2s	PRO:DEM	N	
4.					gen	gen	

Kus	*sää*	*seitte*	*tulit*	*sen*	*rokotiitiN*	*kanssa.*
Kun	sinä	sitten	tul-i-t	se-n	krokotiili-n	kanssa.
When	you	then	come-PAST-2S	it-GEN	crocodile-GEN	with.

'When you arrived with the crocodile.' (10 Santeri; 2/538)

In infinitive phrases the elaboration requirements are reversed. The head of an infinitive phrase is an inflected infinitive, which in this data only appears as a 1st infinitive (the so-called A-infinitive) or a 3rd infinitive (the so-called MA-infinitive). Among all phrase types the infinitive phrase is the one that most resembles a clause, and thus it may be elaborated both morphologically and syntactically in a rather complex way (in addition to the obligatory infinitive head). This gives a lot of potential for creating structural complexity within a component (Example 7).

Example 7

1.	NP N	INFP		
2.		3INF		NP
3.		ill	PRO:DEM	N
4.			partit	

Tetä *ottamaan* *tätä* *uokaa.*
Setä otta-ma-an tä-tä ruokaa.
Uncle take-3INF-ILL this-PARTIT food.
'Uncle, (come and) get this food.' (17 Tatu; 4/502)

In terms of elaboration potential the most powerful components are infinitive phrases (INFP) and adposition phrases (PP). However, both of these are very infrequent in the data, as Figure 32 shows. Therefore their contribution to the complexity of utterances (and to the MLU results) is not important. In the IPSyn analysis, however, both these component types are specified in the sentence structure section (items SS15 and SS16), and the distinct parts of an infinitive phrase are specified in the verb construction section as well (V8, V12, V13, V14, V16). Therefore they play a more important role in IPSyn than they do in MLU.

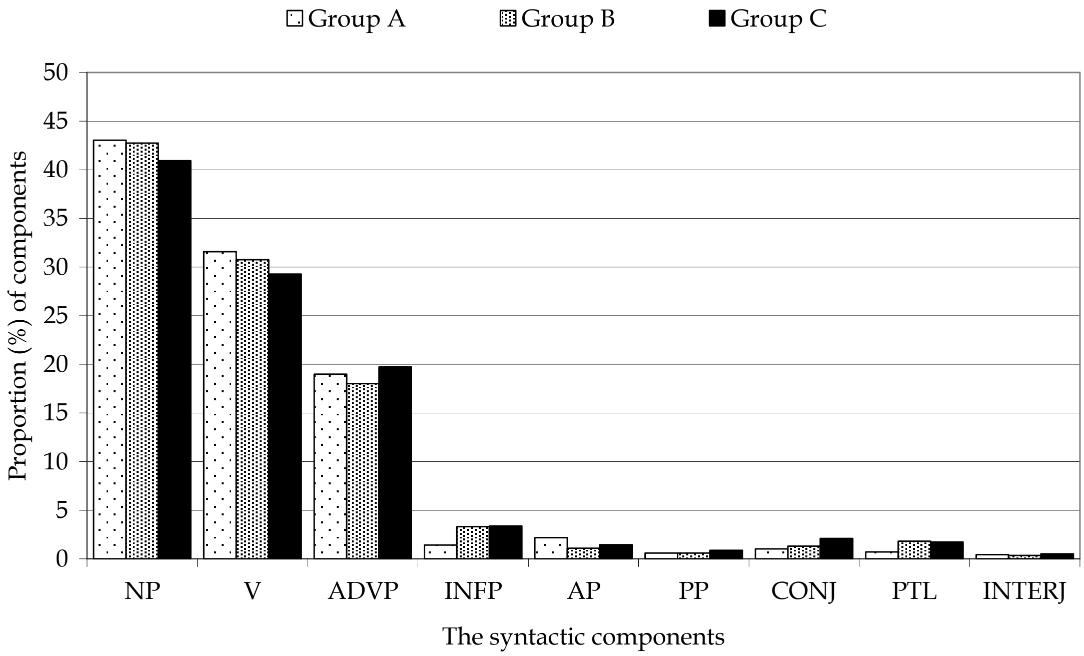


FIGURE 32 The proportions (%) of independent syntactic components in each group. The results are based on the total number of component tokens (1833, 3153 and 2487 for groups A, B and C respectively).

Figure 32 shows that verb constructions (V), noun phrases (NP) and adverb phrases (ADVP) are the most frequent component types. Verb constructions and noun phrases provide a great deal of scope for elaboration and thus, on the basis of both high frequency and the elaboration potential, they are probably the most powerful component types. Therefore the structural complexity of an utterance is likely to depend on the complexity of these particular components.

Adverb phrases do not have a comparable significance since there are so many factors restricting their elaboration.

Generally speaking, rather than differences, Figure 32 shows marked similarity between the groups. It strengthens the finding that there is a similar basic pattern of utterance construction in each group even though MLU and IPSyn as measures of complexity reveal differences.

6.3 Conclusions

In this chapter the focus was on the independent syntactic components in the children's utterances. Investigation of the number of components in their utterances as well as the types and combinations of components highlighted similarities rather than revealing major differences between the three groups of children. With the respect to the number of components, two- to four-component utterances were most frequent, although this shifted from two- to four-component utterances in group A to three- and four-component utterances in group C. This indicates a smooth transition towards longer utterances, taking syntactic components as the unit of calculation. It seems that increase in the number of components is only marginally responsible for growth in structural complexity. Hypotheses 1, 2, 3 and 5 are partly supported but the hypotheses overestimated the magnitude of the increase in the number of components and, consequently, how much the number of components contributes to structural complexity. In fact, structural complexity grows moderately by the addition of independent syntactic components to utterances.

When the preferred combinations of components in the utterances were investigated group by group, it was found that similar utterance types were produced in every group. Moreover, it was noted that the shorter the utterances (in number of components), the more predominant the preferred combinations became. Only in utterances with four and five components were the results distributed more evenly, and combinations with a clear predominance were not found. All three groups preferred utterances consisting of a verb construction, a noun phrase, and an adverb phrase. A similar pattern in the construction of utterances characterises all the groups, although they represent different levels of structural complexity as measured by MLU and IPSyn. Thus, the differences in complexity arise from sources other than a difference of components in the utterances.

Finally, the proportions of each syntactic component in utterances were investigated. Verb constructions, noun phrases, infinitive phrases and adposition phrases were argued to have the widest potential for growing into morphosyntactically complex structures. In terms of the proportions of distinct component types in their utterances, the groups behaved almost identically. The most frequently used components were noun phrases, verb constructions, and adverbial phrases. They account for approximately 90% of all component tokens for each group, and NPs and Vs alone accounted for 70%. Their high frequency on one hand and wide potential on the other hand raise the noun phrase and the verb construction to the position of the components with the greatest scope for contributing to the structural complexity of utterances.

All the results discussed in this chapter point to one conclusion. The importance of components in contributing to structural complexity lies in the fact that they provide a foundation for growth in structural complexity. Distinct components have different potential for elaboration, and it is the utilization of this potential that mostly affects structural complexity. The morphosyntactic elaboration of components, which has been excluded from investigation so far, seems to be the most likely candidate for explaining differences in structural complexity, particularly because the most frequently used components (NPs and Vs) are among those with the greatest potential for elaboration.

7 THE LAYERED STRUCTURE OF TWO-, THREE- AND FOUR-COMPONENT UTTERANCES

In the preceding chapter it was concluded that the morphosyntactic elaboration of syntactic components seems to be the most probable candidate for explaining growth in structural complexity. A direct consequence of elaboration is a layered structure of utterances which is concretely illustrated in the Utterance Analysis. Focusing on the layered structure of components is the next step in the search for factors that may contribute to growth in structural complexity.

Focusing on layers is also focusing on possible trade-offs between the number of components and their elaboration in utterances and asking whether utterances with more components have less elaboration than utterances with fewer components. This is investigated first by detecting the most elaborated component in each two-, three-, and four-component utterance (Section 7.2) and then by examining what kind of compositions the layered components form in entire utterances (Section 7.3). Another approach to possible trade-offs starts from the characteristics of distinct component types (7.4) and aims to find out how much children exploit the elaboration potential of each component type and whether it changes when the number of components in an utterance changes.

7.1 The foundation of the analyses of layered structure and elaboration

The layered structure of utterances is a result of morphological and syntactical elaboration, but talking about layers does not yet specify the exact means which lead to, for example, a two- or three-layered structure. Thus, the number of layers provides a preliminary picture of each individual component or utterance, without concerning any details of elaboration. Example 8 shows an utterance consisting of only unelaborated components, in other words, a single-layered utterance. To become two-layered at least one component requires either morphological or syntactic elaboration. In Example 9 both the verb construction and the noun phrase are inflected, and in Example 10 the noun phrase is elaborated syntactically by extending it with an adjective.

Example 8

1.	ADVP ADV	NP N	V

Tittä *auto`* *pkääntyy.*
Sitten auto kääntyy.
Then car turns.
'Then the car turns.' (6 Ronja; 50/673)

Example 9

1.	NP PRO	V	NP N
2.		1s	ill

Minä *meen* *säNkyy.*
Minä mene-n sänky-yn.
I go-1S bed-ILL.
'I'm going to bed.' (30 Elina; 67/248)

Example 10

1.	V	NP PRO		NP
2.			ADJ	N

On *tuo* *iso* *jalkapallo.*
On tuo iso jalkapallo.
Is that big football
'That is a big football.' (18 Saku; 27/1604)

Between the single- and the two-layered components lies an important borderline, that between unelaborated and elaborated components. There is another important borderline between two-layered components and those with three or more layers, because reaching the latter level usually requires utilization of more than a single elaboration device. One inflection or syntactic extension is enough to produce a two-layered component, but in a three-layered component, usually both a syntactic extension and an inflection are required, as in Example 11. However, this generalization is not without exceptions, since two inflectional endings in a single word form (Example 12) or two syntactic extensions (Example 13) can result in a three-layered structure as well, but these means of elaboration are not as common as the morphosyntactic ones. Thus, the presence of a layered structure already implies the use of different elaboration devices, without yet describing it in detail.

Example 11

1.		NP PRO	V
2.	AUX:NEG		V
3.	1s		neg

em *mie* */muista.*
e-n minä muista&NEG.
Not-1S I remember&NEG.
'I don't remember. ' (18 Saku; 28/82)

Example 12

1.	V
2.	2s
3.	clit

Tarttetpa.
Tarvitse-t-pa.
Need-2S-CLIT.
'Yes, you need.' (17 Tatu; 75/877)

Example 13

1.	NP PRO:DEM	V			NP
2.				AP	N
3.			PTL	ADJ	

S+ *on* *niim* *pieni* *vauva.*
Se on niin pieni vauva.
It is so little baby.
'It is such a little baby.' (3 Laura; 68/1527)

There are two major bases for the analyses reported in this chapter. One is the hypotheses derived from the discrepancy between MLU values and IPSyn scores and the other is the results from the investigation of the independent syntactic components in utterances.

Layers or elaboration are touched upon in all the hypotheses. Hypotheses 1, 2 and 5 suggest that the children in group A are familiarizing themselves little by little with elaboration devices. Becoming acquainted with a new morphological or syntactic device shows up as trying it out in an utterance. IPSyn reflects such experimentation whereas MLU does not because the new devices are used in separate utterances rather than being combined into longer ones. Thus, the differences within group A are assumed to be due to different resources (shown by the variation in IPSyn scores from 42 to 68) and not to the different use of resources (shown by the similarity in MLU values, ranging from 3.688 to 4.388). Hypotheses 1 and 2 are supported if the results concerning layered structure and elaboration in group A are as follows:

a) All children in group A produce utterances with a similar number of elaboration layers, but the number of layers remains small.
b) Elaboration mostly involves only one syntactic component per utterance.

Hypotheses 1 and 2 are already indirectly supported by the results for independent syntactic components presented in the preceding chapter. The children in group A mostly produced utterances consisting of two or three syntactic components (34.2% and 42.4% respectively of all utterances). Within their limited MLU range (3.688–4.388), this leaves little room for elaboration of two- or three-component utterances and almost none for four-component utterances. Only in their single-component utterances (12.5%) are there real possibilities for elaboration. Therefore group A's output should have relatively few layers overall.

Group C is similar to group A in that the relation between the MLU values and IPSyn scores is similar; however the level of the results in these groups differs significantly. Children in group C represent those with a high score on both scales. This has led to the predictions expressed in Hypothesis 3 that for group C utterance length does not restrict morphological and syntactic elaboration and that structural complexity has grown vertically by elaboration (in addition to the horizontal growth in complexity). If the morphosyntactic elaboration is not restricted this means that MLU does not restrict layered structure so much. In group C, the average utterance length varies between 6.050 and 6.350 morphemes. In the analysis of syntactic components reported in the previous chapter, it was found that the majority of utterances produced by the children in group C consist of three or four components (50.3% and 29.4% respectively). To reach an average length of six morphemes a three-component utterance needs to have each component elaborated into two layers on average. For two-component utterances (9.3% of all utterances in group C), both of the components could reach the third layer by elaboration. This usually implies use of both syntactic and morphological devices within a single component. Thus, growth in MLU provides space for a variety of elaboration and for the development of further layers.

In the previous chapter the properties of each syntactic component were discussed in terms of morphosyntactic complexity. It was found that the syntactic components with the most elaboration potential are the noun phrase, verb construction, infinitive phrase, and adposition phrase, whereas there are severe limitations on the elaboration of adverb and adjective phrases. It was also found that the most common syntactic components in all three groups are noun phrases and verb constructions, together with adverb phrases. These same components were also found in the most utilized component combinations. In practice this means that if a three-component utterance consists of a noun phrase, a verb construction and an adverb phrase this provides more space for the noun phrase and the verb construction to expand into further layers, because the adverb phrase has very restricted elaboration potential. In the case of a prototypical three-component utterance which results in six or even more morphemes, the small space requirement of the adverb phrase remarkably facilitates growth into three- and four-layered utterances on the part of the other two components.

Finally, concerning group B, the hypotheses propose that growth in complexity in the structure of utterances is mainly achieved by combining more and more resources. This prediction has already received support from the component analysis, which showed that the number of components does not change markedly between the groups. Thus, the components *per se* cannot be responsible for the substantial lengthening of the utterances in group B. Therefore the most probable candidate is expansion by elaboration. It is logical that elaboration has two directions for expansion, first horizontally (more and more syntactic components within an utterance), and second, vertically within single components by combining already acquired elaboration resources. The former results in extent of elaboration whereas the latter leads to depth of elaboration. Whether expansion in both directions is simultaneous or not will be seen in the forthcoming analyses.

The hypotheses based on the divergent results in terms of MLU and IPSyn and on the results of the component analysis constitute the foundation of the analysis of the layered structure of whole utterances and individual syntactic components. The focus is on the following three questions:

1) What is the maximum number of layers in an utterance?
2) How many components in an utterance are elaborated and how many layers do the elaborated components reach?
3) What component types are elaborated and to what depth?

Answers to the first question come from an analysis where only the deepest layer in each utterance is considered (Section 7.2). For the second question, all the components and the layers they reach are taken into account to determine the spread of horizontal expansion (Section 7.3). Finally, answering the third question requires focusing on one component type at a time (section 7.4). Every token of a noun phrases (NP), verb structure (V), and adverb phrase (ADVP) will be classified according to the number of layers attained. These component types are chosen because they are the most common types in all three groups. The analysis focuses on two- to four-component utterances only because utterances of this extent account for the vast majority (A, 85.6%; B, 86.9 %; C, 89.0%) of all the utterances produced in each of the groups. Thus, the investigation of layers concentrates on the core of the data, both in terms of component selection and the extent of the utterances. All the results are based on the Utterance Analysis.

7.2 The depth of elaboration in two-, three- and four-component utterances

In the Utterance Analysis each utterance was investigated component by component. Each part of a component was identified and its relations to the other parts within the component were described. As a result of this procedure the layered structure of components and the layers within utterances were exposed. Investigating layered structure provides information both about the maximum depth of layers children reach in their utterances and about how many components are elaborated and how deeply. First I will look at the maximum depth of layers, that is, elaboration in two-, three- and four-component utterances. This analysis concentrates only on those components with most layers. In Examples 14 and 15 the verb constructions are the most deeply elaborated components and both of them reach the third layer, although they are rather different as constructions. In Example 14 the elaboration involves only inflection but in Example 15 there is a verb chain including inflection in both members of the chain. Thus, it could be argued that the verb structure in Example 15 is more complex than that in Example 14, but for the analysis of the maximum number of layers this does not play a role yet. However, it gives a good example of how different means can be used to reach the same number of layers.

Example 14

1.	NP PRO	V	ADVP ADV	NP PRO:DEM
2.		past		all
3.		1s		

Minä	*antein*	*nyt*	*tsille.*
Minä	anno-i-n	nyt	si-lle.
I	give-PAST-1S	now	it-ALL.

'I gave to it now.' (38 Pekka; 9/1719)

Example 15

1.	ADVP ADV		V	NP PRO	NP PRO:DEM
2.		V	V		adess
3.		2s	1inf		

Nyt	*voit*	*lent`äät*	*tinä*	*tällä.*
Nyt	voi-t	lentä-ä	sinä	tä-llä.
Now	can-2S	fly-1INF	you	this-ADESS.

'Now you can fly with this. ' (34 Leo; 2/213)

What we know for each group so far is the average length of utterance calculated in morphemes, as well as the number of components and the most common component combinations in utterances. When these pieces of information are combined the boundaries for component elaboration begin to shape. For two-component utterances it was reported in Chapter 6 that they account for 34.2 % (246/720) of all utterances in group A, 18.3% (190/1040) in group B and 9.3% (67/720) in group C. In all three groups the most common component combinations in two-component utterances were V + NP (A 53.7%; B 55.8%; C 59.7%), V + AdvP (A 13.0%; B 12.1%; C 13.4 %) and NP + AdvP (A 13.0%; B 11.1%; C 7.5%). Thus, the three favourite combination types account for approximately 80 % of the two-component utterances in each group. It was stated that verb structures and noun phrases are the most important components in terms of elaboration potential. How much space this potential can account for depends both on utterance length and on which other components are present in a given utterance. In the case of NP and V featuring in the same utterance they both contend for space and neither of them can fully realize their potential, especially when the utterances are short, as with group A and partly with group B. However, in utterances where an adverb phrase is the other component there is a strong possibility that layers will appear because adverb phrases are restricted in elaboration and thus allow space for the elaboration of the other component, that is, a noun phrase or a verb construction in two-component utterances.

In Table 15 and in Figure 33 the distribution of the maximum number of layers in two-component utterances in all the three groups is presented as percentage values. The figure and the table show first of all that the vast majority of two-component utterances are elaborated in all groups. In fact, in group C there are no unelaborated utterances at all and in groups A and B they make up only 14.2 % and 15.3 % respectively. Secondly, group C is clearly ahead of the other groups in the maximally reached layers. As many as 67.2% of all two-component utterances produced by children in group C reach the third layer, which means that there is both syntactic and morphological elaboration of the components. Moreover, 19.4 % of utterances reach the fourth layer, that is, represent even more sophisticated elaboration. At the same time, the children in groups A and B produce mostly utterances reaching maximally the second layer (A 54.1%, and B 40.5%), although there is a substantial proportion of utterances reaching the third layer as well (A 27.6% and B 35.8 %).

TABLE 15 The proportions (%) and numbers of two-component utterances categorized on the basis of the maximum number of layers. The largest proportions are highlighted with grey. The proportions in percentages appear in graphic form in Figure 33.

		One-layered	**Two-layered**	**Three-layerd**	**Four-layered**	**Five-layered**	**Total**
Group A							
	Utterances	35	133	68	10	0	246
	%	14.2%	54.1%	27.6%	4.1%	0.0%	100.0%
Group B							
	Utterances	29	77	68	15	1	190
	%	15.3%	40.5%	35.8%	7.9%	0.5%	100.0%
Group C							
	Utterances	0	7	45	13	2	67
	%	0.0%	10.4%	67.2%	19.4%	3.0%	100.0%

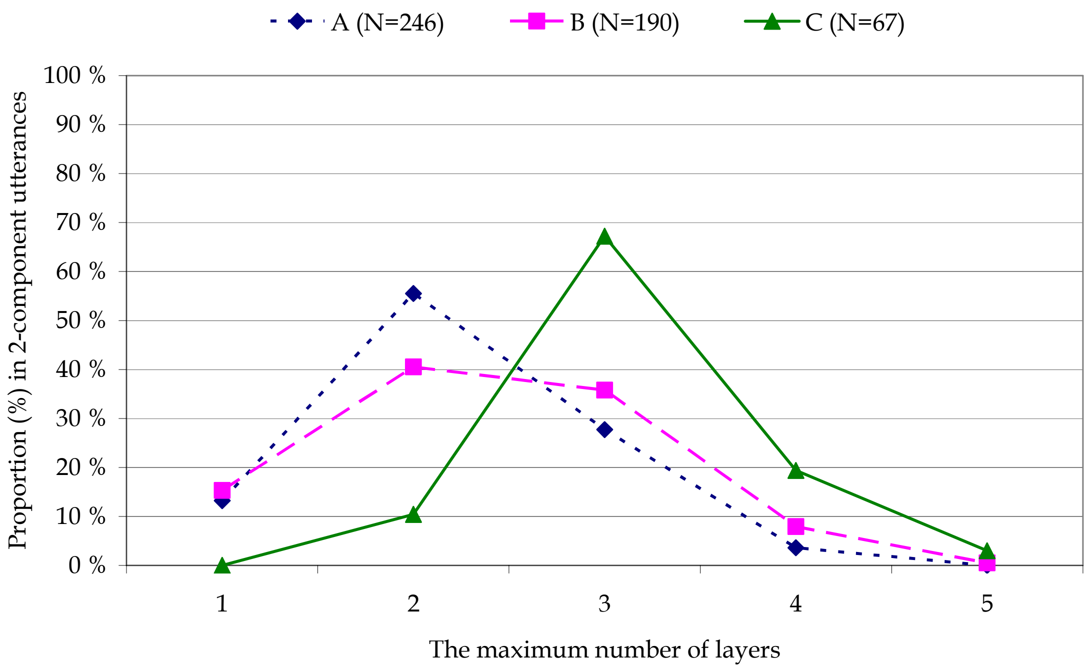


FIGURE 33 The distribution of the maximum number of layers in two-component utterances. The single-layered utterances are those without any elaboration where as two- to five-layered utterances include morphological, syntactic or both types of elaboration. The percentage values are shown in Table 15.

According to the hypotheses it was expected that group B would settle itself between groups A and C. It was suggested that in group B the growth of MLU associated with rather fixed IPSyn scores implied the combining of structures in utterances. Logically this should also result in deeper layers. Evidence of such a development in the growth of complexity can be seen in the two-component utterances, since the maximally three-layered utterances constitute a substantial proportion of the utterances in group B. However, there is still a slight tendency toward combining structures rather than a result clearly distinguishing this group from the performance of group A. This might imply that the combining of structures in group B is more of a horizontal nature, that is, involving several components within an utterance instead of combining structural means in one component. The spreading of elaboration to the whole utterance is not optimally shown in an investigation of maximum number of layers. The suggestion that elaboration first spreads horizontally (several components of an utterance are elaborated) and only after that vertically (going deeper in individual components) will be tested in Section 7.3 which investigates the breadth of layered structure in utterances.

In the case of three-component utterances, the groups look more similar than with two-component utterances. Substantial differences are evident only in utterances with a maximally three-layered structure (Table 16 and Figure 34). In group C the proportion of utterances of this type is 48.1% whereas in groups A and B the proportions are 18.74% and 34.4% respectively. When compared to two-component utterances (Figure 33), group C has changed most: the proportion of maximally two-layered utterances has grown from 10.4% to

44.5% whereas the proportions of three- and four-layered utterances have decreased from 67.2% and 19.4% to 48.1% and 5.5%. In other words there is a shift from three-layered utterances to three- and two-layered ones. In group B there is also a slight shift in the pattern of the depth of utterances. In the case of two-component utterances two- and three-layered ones are almost equally represented but with three-component utterances the preference has shifted more clearly to two-layered utterances.

TABLE 16 The proportions (%) and numbers of three-component utterances categorized on the basis of the maximum number of layers. The largest proportions are highlighted with grey. The same proportions appear in graphic form in Figure 34.

		One-layered	Two-layered	Three-layered	Four-layered	Five-layered	Total
Group A							
	Utterances	73	170	57	5	0	305
	%	23.9%	55.7%	18.7%	1.6%	0.0%	99.9%
Group B							
	Utterances	37	277	176	20	2	512
	%	7.2%	54.1%	34.4%	3.9%	0.4%	100.0%
Group C							
	Utterances	6	161	174	20	1	362
	%	1.7%	44.5%	48.1%	5.5%	0.3%	100.1

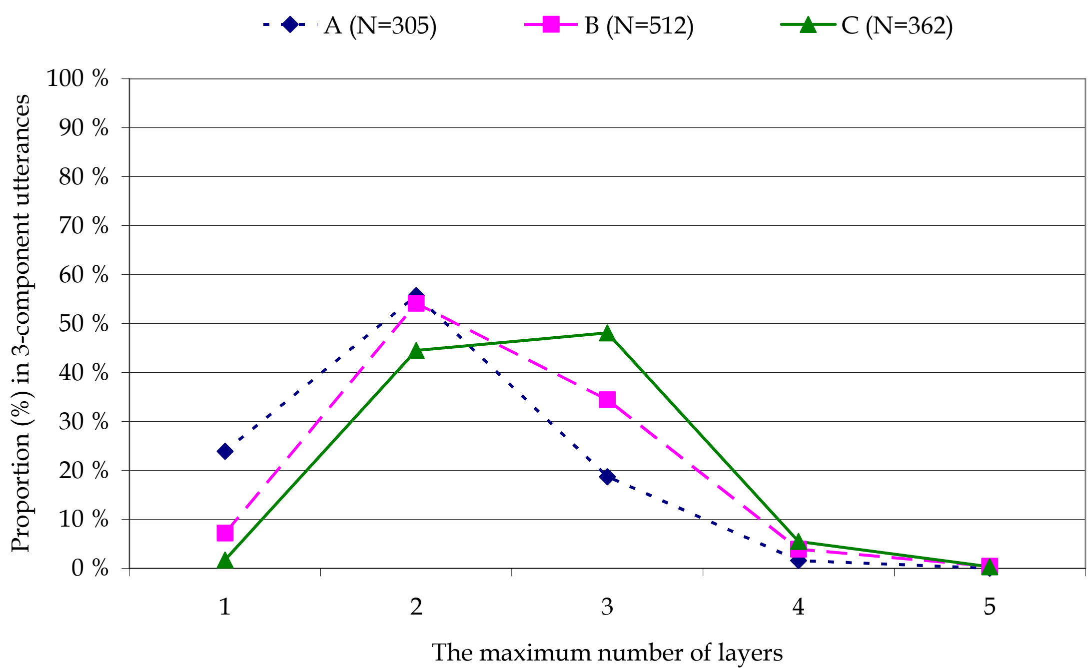


FIGURE 34 The distribution of the maximum number of layers in three-component utterances. The same proportions (%) are shown in Table 16.

In four-component utterances the tendency towards similarity across the groups continues, and the results in all groups are almost identical. The results of this investigation are presented in Table 17 and Figure 35. Approximately 60% of utterances in all groups reach a maximum of two layers of elaboration, which requires either morphological or syntactic elaboration but not both in the most deeply elaborated component in a four-component utterance.

TABLE 17 The proportions (%) and numbers of four-component utterances categorized on the basis of the maximum number of layers. The largest proportions are highlighted with grey. The same proportions appear in graphic form in Figure 35.

		One-layered	Two-layered	Three-layered	Four-layered	Five-layered	Total
Group A							
	Utterances	12	41	10	2	0	65
	%	18.5 %	63.1 %	15.4 %	3.1 %	0.0 %	100.1 %
Group B							
	Utterances	32	118	47	4	1	202
	%	15.8 %	58.4 %	23.3 %	2.0 %	0.5 %	100.0 %
Group C							
	Utterances	20	124	60	8	0	212
	%	9.4 %	58.5 %	28.3 %	3.8 %	0.0 %	100.0 %

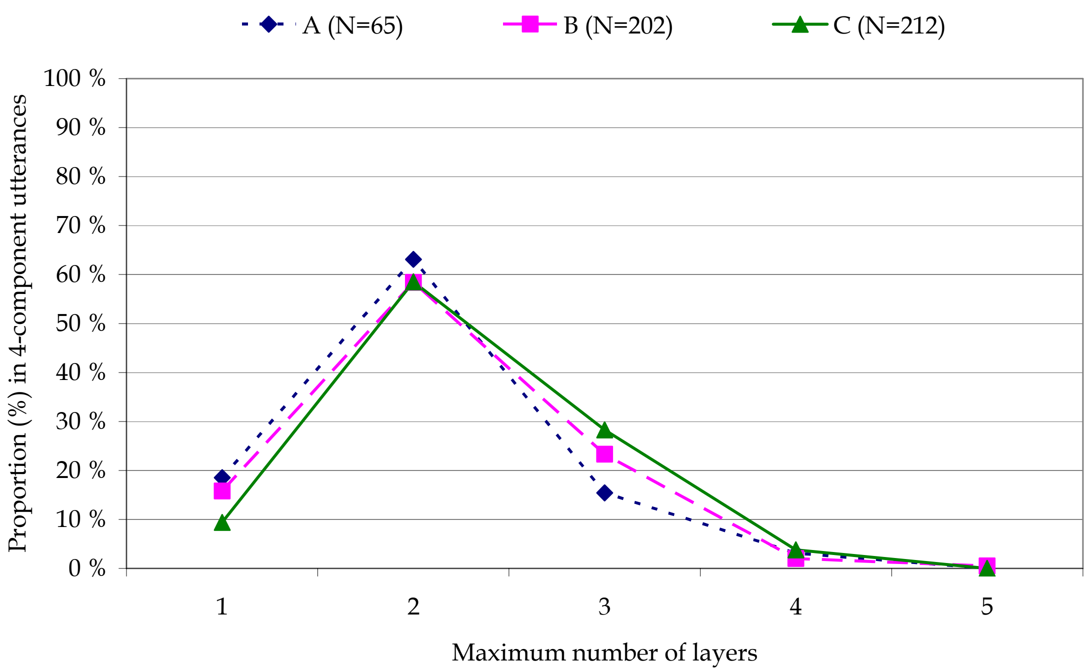


FIGURE 35 The distribution of the maximum number of layers in four-component utterances. The same proportions (%) are presented in Table 17.

In the single-layered (unelaborated utterances) and three-layered utterances (where the most deeply elaborated component includes both morphological and syntactic elaboration) the groups perform within approximately 10% of each other, with groups B and C biased more towards three-layered elaborations and group A more towards lack of elaboration. These are practically the only differences found in the depth of elaboration in four-component utterances.

The investigation of the maximum elaboration depth in two- to four-component utterances produced by children in groups A, B, and C revealed that the more components there are in an utterance the more alike the performances of the groups become. This suggests that differences in structural complexity between the groups are most likely to be found in two-component utterances, at least where maximum depth of elaboration is concerned. This seems to suggest a model of structural complexity that diverges from that suggested in the hypotheses. It was assumed that differences between the groups are evident in all kinds of productions and the possibility of the groups merging with each other was not raised. In other words, the groups were expected to show similar differences between each other irrespective of the number of components in utterances. The most striking changes in performance are seen in group C. The proportion of maximally three-layered utterances decreases from 67.2% in the two-component utterances to 48.1% in three-component utterances and finally to 28.3% in four-component utterances, whereas the proportions of maximally two-layered utterances increase from 10.4% via 44.5% to 58.5 % respectively. In groups A and B these changes are much more moderate, varying in maximally two-layered utterances from 54.1 % to 63.1% in group A and from 40.5% to 58.4% in group B. In the maximally three-layered utterances the ranges of changes are 27.6–15.4% (group A) and 35.8–23.3% (group B) respectively.

The trade-off effect (Bock 1982; Crystal 1987) was one of the reasons why utterances rather than so-called complete clauses are analysed here. It is argued that increase in complexity in one area may cause decrease in complexity in another area, and this is exactly what seems to be happening in the productions of the children in group C. When the number of syntactic components increases the complexity of elaboration decreases, at least as far as the maximum elaboration depth is concerned. Why is the trade-off effect evident only in group C and why do children in groups A and B perform constantly at a rather similar level? There are at least two possible explanations, which can later be tested. First, only the children in group C have advanced in their skills to the point where they are accustomed to more sophisticated elaboration, but even they are able to make good use of elaboration devices only in utterances where there are no more than two components to be elaborated. The other possible answer to this question is that the three groups represent three different developmental phases of elaboration. However, this analysis of the maximum depth of layers does not yet make things sufficiently clear. Group A seems to represent children who are able to perform some kinds of elaboration, either morphological or syntactic, but combining these devices is still very rare. The next step, represented by the children in group B, involves combining different elaborations mostly by spreading across the different components, and only occasionally within a single component. This keeps the maximum number of layers in group B closer to those in group A than in group C. Therefore it is the number of elaborated components that should clearly separate groups A and B rather than the maximum depth of the layers. Finally, the children in group C are working with phrases, where they can combine elaboration devices easily in one component, but this mostly happens only when there are not too many components in an utterance. To find out whether either of these explanations gets further support I will take a closer look at the extent to which elaboration has spread within utterances.

7.3 The breadth of elaboration in two-, three- and four-component utterances

Investigating the extent of elaboration involves studying the number of elaborated components per utterance and the number of layers in these components. Let us look again at the utterances by Pekka (ID 38) and Leo (ID 34) in Examples 14 and 15. In the previous analysis the only important factor was the most deeply elaborated component in an utterance. In the case of Examples 14 and 15 the utterances were categorized as maximally three-layered, that is, the most deeply elaborated component(s) of the utterance has (have) three layers. Now that the focus moves from the maximum number of layers to the full extent of elaboration, each component and all its layers are important. In Example 14 there are four syntactic components (two NPs, V and ADVP), two of them are unelaborated (one of the NPs and the ADVP) whereas the other noun phrase reaches the second layer and the verb structure the third layer (with its inflection). A condensed description of Example 14 would thus be 3+2+1+1 expressing the layers of each of the four components in the utterance, in order of elaboration depth. Similarly the utterance in Example 15 would have a similar description of its components and layers.

Example 14

1.	NP PRO	V	ADVP ADV	NP PRO:DEM
2.		past		all
3.		1s		
	Minä Minä I	*antein* anno-i-n give-PAST-1S	*nyt* nyt now	*tsille.* si-lle. it-ALL.

‘I gave to it now.’ (38 Pekka; 9/1719)

Example 15

1.	ADVP ADV		V	NP PRO	NP PRO:DEM
2.		V	V		adess
3.		2s	1inf		

Nyt	*voit*	*lent`äät*	*tinä*	*tällä.*
Nyt	voi-t	lentä-ä	sinä	tä-llä.
Now	can-2S	fly-1INF	you	this-ADESS.

'Now you can fly with this.' (34 Leo; 2/213)

The starting point for these analyses lies again in Hypotheses 1–5 introduced in Section 4.2.2. In group A, the results in the previous analysis already showed that combining elaboration devices within a component is not yet very common, since the majority of their two-, three- and four-component utterances reached maximally a two-layer level of elaboration. All two-layered components are the result of either a single inflection or a syntactic extension but not both. Due to the shortness of their utterances (MLU), it was predicted in Hypotheses 1 and 2 that elaboration would not yet spread to many components in group A.

The results for group B in the previous analysis only partly met expectations because the maximum level of layers was still rather low and the results were more like those for group A than group C. However, combining resources at the level of the whole utterance rather than at component level may turn out to be the factor differentiating the structural complexity of the productions of the children in groups A and B. If this is the case, group B differs from group A in producing a larger number of elaborated components in their utterances, as Hypothesis 4 predicts.

In group C, an increase in layers was clearly observable by comparison with the other two groups, but only for two-component utterances. According to Hypothesis 3, the spread of elaboration to several components in an utterance is to be expected because the average length of utterances will not restrict elaboration so much. Testing these hypotheses will continue now in two ways. The first task is to find out how many elaborated components there are in two-, three- and four-component utterances for each of the three groups. The second task is to investigate what kinds of elaborated components are combined in the utterances produced by each group. At this point in the investigation components are still defined according to the number of layers they include. Do deeply layered components have other elaborated components beside them in an utterance or are unelaborated component partners preferred? The analysis starts with investigation of the number of elaborated components first in two- and then in three- and four-component utterances. Groups A, B and C are compared in terms of how much elaboration they prefer in their utterances. For two-component utterances the number of elaborated components varies substantially between the groups. This is shown in Table 18 and Figure 36. In group A the majority (67.5%) of two-component utterances have only one elaborated component. The other two elaboration categories (0-

column: unelaborated utterances; 2-column: both components elaborated) share the rest of the utterances almost equally (14.2% and 18.3%).

In group B the children mostly produce two-component utterances with only one component elaborated (47.9%). However, utterances with both components elaborated account for a substantial proportion (36.8%) of utterances as well. In the category of unelaborated utterances the children in groups A and B perform very similarly.

Group C differs from the two other groups in three ways. First the number of two-component utterances (67) is much lower than that in groups A (246) and B (190). Second, almost two thirds (65.7%) of two-component utterances have both components elaborated, while in groups A and B the preferred elaboration category was utterances with one elaborated component. Third and finally, the children in group C do not have any unelaborated two-component utterances in the data sample.

TABLE 18 Numbers of elaborated components in two-component utterances. The results are presented in graphic form in Figure 36.

		Elaborated components			
		0	**1**	**2**	**Total**
Group A					
	Utterances	35	166	45	246
	%	14.2 %	67.5 %	18.3 %	100.0 %
Group B					
	Utterances	29	91	70	190
	%	15.3 %	47.9 %	36.8 %	100.0 %
Group C					
	Utterances	0	23	44	67
	%	0.0 %	34.3 %	65.7 %	100.0 %

TABLE 19 Numbers of elaborated components in three-component utterances. The results are presented in graphic form in Figure 37.

		Elaborated components				
		0	**1**	**2**	**3**	**Total**
Group A						
	Utterances	73	153	71	8	305
	%	23.9 %	50.2 %	23.3 %	2.6 %	100.0 %
Group B						
	Utterances	37	247	200	28	512
	%	7.2 %	48.2 %	39.1 %	5.5 %	100.0 %
Group C						
	Utterances	6	148	172	36	362
	%	1.7 %	40.9 %	47.5 %	9.9 %	100.0 %

TABLE 20 Numbers of elaborated components in four-component utterances. The results are presented in graph form in Figure 38.

		Elaborated components					
		0	1	2	3	4	Total
Group A							
	Utterances	12	22	17	13	1	65
	%	18.5 %	33.8 %	26.2 %	20.0 %	1.5 %	100.0 %
Group B							
	Utterances	32	67	73	28	2	202
	%	15.8 %	33.2 %	36.1 %	13.9 %	1.0 %	100.0 %
Group C							
	Utterances	20	76	86	27	3	212
	%	9.4 %	35.8 %	40.6 %	12.7 %	1.4 %	99.9 %

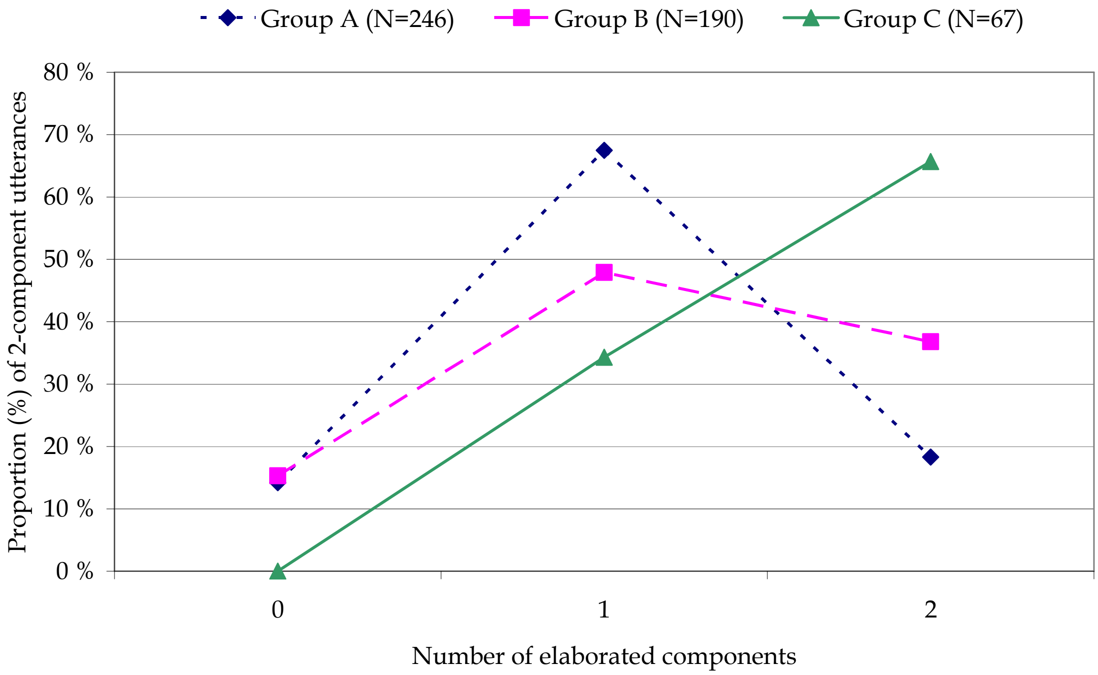


FIGURE 36 The proportions (%) of two-component utterances with zero, one, or two elaborated components. The same proportions are shown in Table 18.

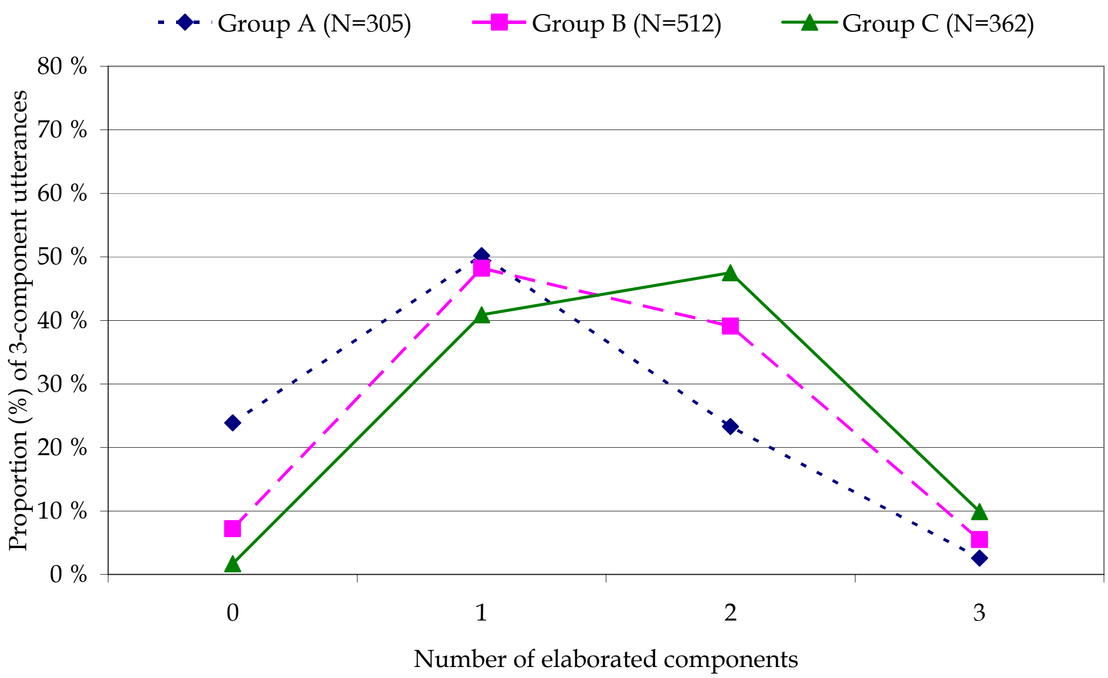


FIGURE 37 The proportions (%) of three-component utterances with zero, one, two, or three elaborated components. The same proportions are shown in Table 19.

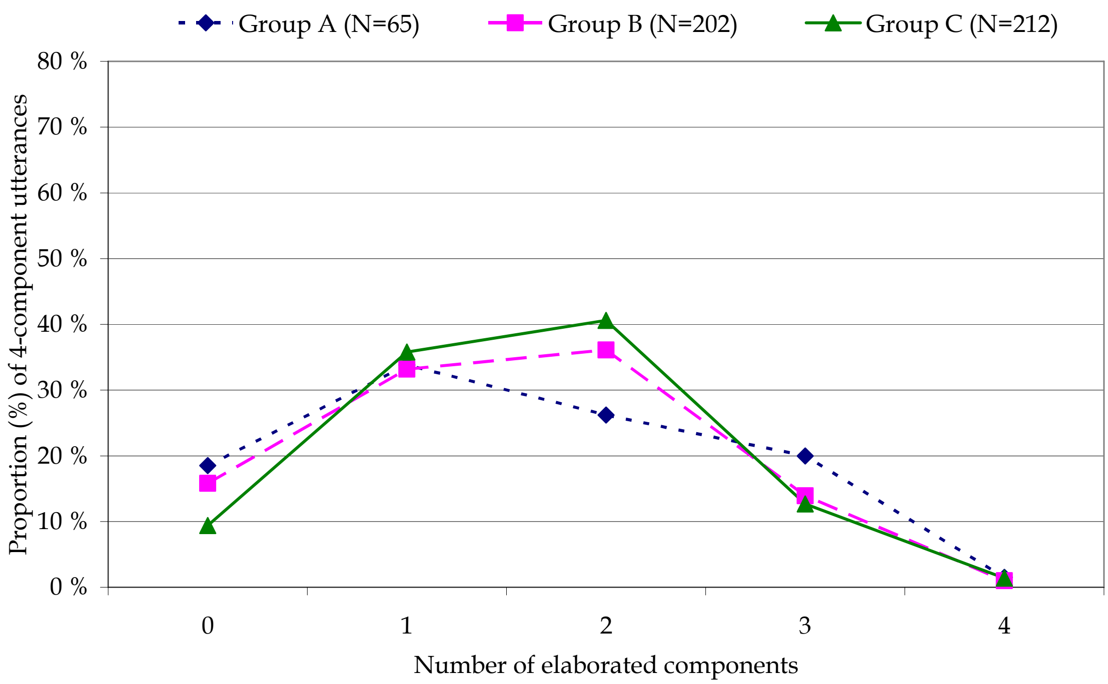


FIGURE 38 The proportions (%) of four-component utterances with zero, one, two, three, or four elaborated components. The same proportions are shown in Table 20.

All these observations point in one direction. It seems that the children in group C produce structurally more advanced two-component utterances than the children in groups A and B. The results from all three groups combined show a continuum from group A to group B and finally to group C in terms of the number of elaborated components.

In three-component utterances the differences between the groups decrease remarkably (see Table 19 and Figure 37). The situation in group A is rather similar to that in two-component utterances: elaboration of only one component per utterance is preferred (50.2% of three-component utterances), and unelaborated utterances (23.9%) together with utterances including two elaborated components (23.3%) are in equal proportion. In two-component utterances group B was a bridge between A and C, but now it has very much approached the performance of group C. The children in groups B and C show almost equal preference for one (B 48.2%; C 40.9%) and two (B 40.9%; C 47.5%) elaborated components.

In four-component utterances it is as if there were only two groups rather than three, A being one and B and C combined the other. The children in groups B and C continue to prefer utterances including one or two elaborated components, while the other elaboration categories are more or less marginal. However, group A shows a more evenly distributed performance across categories, although also in this group one third of utterances belong to the category of including only one elaborated component.

Analyses of the number of elaborated components support Hypotheses 1 and 2. In group A the children favour the elaboration of only one component per utterance and the preference is shown in all three utterance categories. So far, the results also support Hypothesis 4 concerning group B. The children in group B seem to perform very similarly to those in group C as far as the number of elaborated components is concerned. In other words, they have extended morphosyntactic elaboration to a broader ground in utterances. The next analysis dealing with the depth of elaboration will show whether the extension of elaboration to more components per utterance is the only property that connects groups B and C.

Now that we know how utterances are built out of elaborated and unelaborated components in each group and in each of the three utterance categories, it is time to look at the component combinations in utterances more closely. If we have a three-component utterance with two of its components elaborated, how many layers does each of these components have? What if all three components are elaborated: do the components have fewer layers than in the case of only two elaborated components? How do the combinations of layered components change when the total number of components in an utterance decreases to two, or increases to four? In other words, are there deeper elaborations in two-component than in three- or four-component utterances and does the depth of elaboration change when the number of elaborated components per utterance changes? To get a picture of this the layers of each component in two-, three- and four-component utterances were counted and a profile of the layer combinations of each utterance were condensed into a numerical expression.

TABLE 21 Layer combinations in two-component utterances (U) for all three groups. Pairs of numbers refer to the two components in the utterances, and each number expresses how many elaboration layers a component has. The profiles of combinations for each group are shown in Figure 39.

			Group A			Group B			Group C	
			U	%		U	%		U	%
2 elaborated components										
	5+3		0			0			1	1.5 %
	4+2		2	0.8 %		6	3.2 %		7	10.4 %
	3+3		2	0.8 %		4	2.1 %		4	6.0 %
	3+2		13	5.3 %		38	20.0 %		26	38.8 %
	2+2		28	11.4 %		22	11.6 %		6	9.0 %
1 elaborated component										
	5+1		0			1	0.5 %		1	1.5 %
	4+1		8	3.3 %		9	4.7 %		6	9.0 %
	3+1		53	21.5 %		26	13.7 %		15	22.4 %
	2+1		105	42.7 %		55	28.9 %		1	1.5 %
Unelaborated utterances										
	1+1		35	14.2 %		29	15.3 %		0	
Total			**246**	**100.0 %**		**190**	**100.0 %**		**67**	**100.1 %**

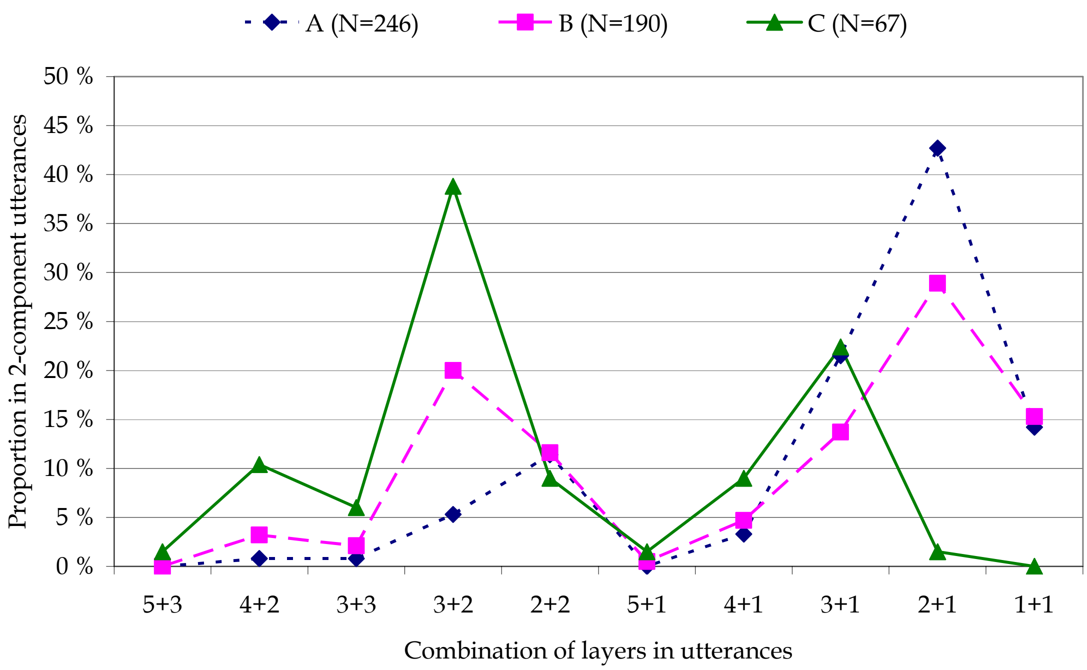


FIGURE 39 The proportion of distinct layer combinations in two-component utterances for all three groups. The figure is based on the results shown in Table 21.

For example, 4+2+1 is a description of a three-component utterance. One of the components is unelaborated (1 refers to a single-layered structure), and the other two are elaborated. One of the elaborated components has four layers and the other two. The components are expressed numerically in order of depth and not in order of appearance in the actual utterance.

In Table 21 and Figure 39, the two-component utterances produced by children in all three groups are categorized according to the layers in their component combinations. The overall picture shows the same as was found already in the previous analysis concerning the number of elaborated components: group A and C differ remarkably from each other and group B forms a bridge between them. However, the present analysis reveals also that the difference is not only in the number of elaborated components but also in how deeply they are elaborated. When only one component is elaborated groups B and C prefer a 2+1-combination (A 42.75; B 28.9%), in other words the elaborated component has two layers as in Example 16. This is the simplest possible form of elaboration and it is a result of either a single inflection or a syntactic extension in one component.

Example 16

1.	V	NP PRO:DEM
2.	pass	

Pistetäät *tä2ä2.*
Piste-tään tämä.
Put-PASS this.
'Let's put this one.' (6 Ronja; 35/57)

If we look at group C, the preferred form of utterances with one elaborated component is 3+1 (22.4%) which requires double inflection or doubled syntactic extension, or a combination of syntactic extension and inflection as in Example 17:

Example 17

1.		ADVP ADV:DEM	V
2.	V		V
3.	cond		1inf

Vois *tonnel* *laittaa.*
Vo-isi tuonne laitta-a.
Can-COND there put-1INF.
'It could be placed over there.' (19 Tuomo; 56/497)

Thus, children in group C tend to produce more complex structures. This tendency continues in utterances with both components elaborated. Group C does not prefer the simplest alternative (2+2) but the slightly more complex combination of 3+2 layers shown in Example 18. Interestingly, in this case group B seems to follow group C in combination quality, and one fifth of its two-component utterances also fall into this category.

Example 18

1.	NP PRO:DEM		V
2.	partit	V	V
3.			1inf

Tätä *voi* *korjata.*
Tä-tä voi korja-ta.
This-PARTIT can repair-1INF.
'This can be repaired.' (19 Tuomo; 55/469)

The children in group A produce few utterances with both components elaborated, and most of them represent the simplest possible combination alternative of 2+2.

In the case of three-component utterances the distribution of utterances across layer combinations looks much more similar in all three groups than with two-component utterances (see Figure 40 and Table 22). However, group C differs from the other two groups in that it has two equally common layer combinations in utterances with one or two elaborated components. In both categories they are the two simplest possible combination alternatives. In cases of only one elaborated component they are 2+1+1 (18.0%) and 3+1+1 (20.2%), and in the category of two elaborated components 2+2+1 (22.1%) and 3+2+1 (19.9%). In the same circumstances groups A and B prefer only the simplest alternatives 2+1+1 (A 38.7%; B 28.3%) and 2+2+1 (A 15.1%; B 22.9%) although the more complex combinations are produced too. In all three groups utterances with three elaborated components as well the deepest layer combinations with two elaborated components are very marginal.

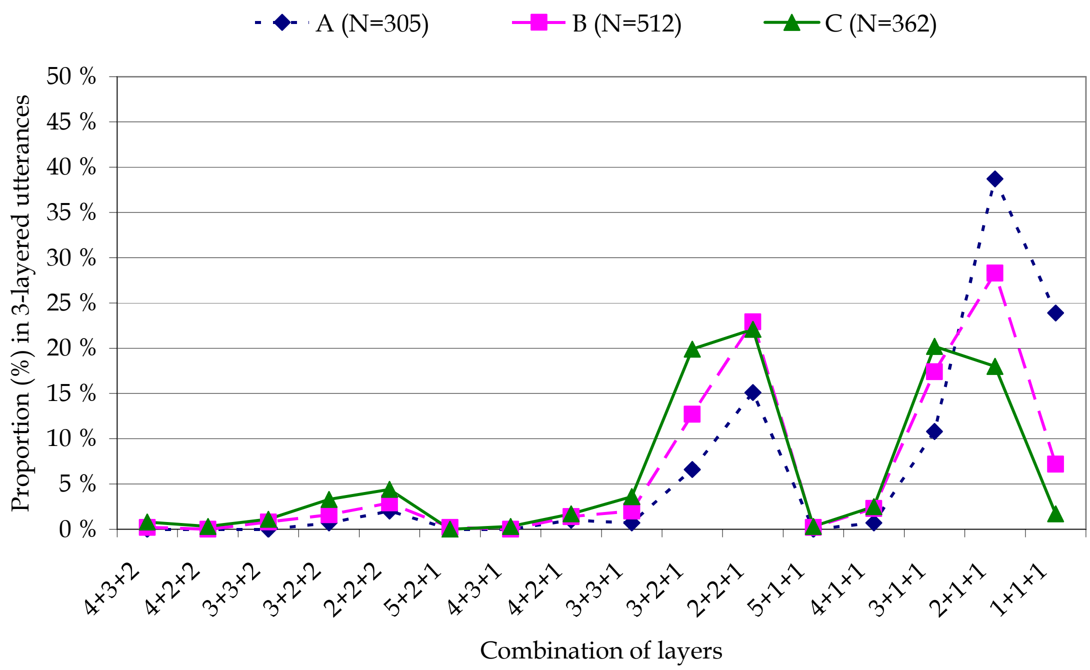


FIGURE 40 The proportion of the different layer combinations in the three-component utterances. The proportions and numbers of utterances are shown in Table 22.

TABLE 22 Layer combinations in three-component utterances (U) for all three groups. Sequences of numbers (e.g. 3+3+2) refer to the three components in the utterances, and each number expresses how many elaboration layers a component has. The combinations are shown in Figure 40.

		Group A		Group B		Group C	
		U	**%**	**U**	**%**	**U**	**%**
3 elaborated components							
	4+3+2			1	0.2 %	3	0.8 %
	4+2+2					1	0.3 %
	3+3+2			4	0.8 %	4	1.1 %
	3+2+2	2	0.7 %	8	1.6 %	12	3.3 %
	2+2+2	6	2.0 %	15	2.9 %	16	4.4 %
2 elaborated components							
	5+2+1			1	0.2 %		
	4+3+1					1	0.3 %
	4+2+1	3	1.0 %	7	1.4 %	6	1.7 %
	3+3+1	2	0.7 %	10	2.0 %	13	3.6 %
	3+2+1	20	6.6 %	65	12.7 %	72	19.9 %
	2+2+1	46	15.1 %	117	22.9 %	80	22.1 %
1 elaborated component							
	5+1+1			1	0.2 %	1	0.3 %
	4+1+1	2	0.7 %	12	2.3 %	9	2.5 %
	3+1+1	33	10.8 %	89	17.4 %	73	20.2 %
	2+1+1	118	38.7 %	145	28.3 %	65	18.0 %
Unelaborated utterances							
	1+1+1	73	23.9 %	37	7.2 %	6	1.7 %
Total		**305**	**100.2 %**	**512**	**100.1 %**	**362**	**100.2 %**

For group B moving from two-component to three-component utterances seems to mean crossing some kind of a borderline. In two-component utterances it performed very similarly to group C in layer combinations. Children in both groups combined equally deeply elaborated components. However in three-component utterances the situation has changed. Although groups B and C are similar in the way their three3-component utterances are distributed in categories expressing the number of elaborated components in them, elaboration within components is different in these groups. Group B prefers a very simple elaboration but in group C deeper elaboration is already equally involved. In other words, children in group C often produce utterances where components are more deeply elaborated than in utterances by the children in group B. Group B is on its way to the situation which is associated with group C in Hypothesis 3: “Multidimensional complexity has grown both horizontally and vertically by adding new syntactic components and elaborations.”

In four-component utterances the three groups perform almost identically. They all prefer the simplest combination types and the distribution of their utterances across the combination categories follows similar trends, as Table 23 and Figure 41 show. The simplest possible alternative is preferred with only two-layered elaboration.

TABLE 23 Layer combinations in four-component utterances (U) for all three groups. The combinations appear in graph form in Figure 41.

		Group A		**Group B**		**Group C**	
		U	**%**	**U**	**%**	**U**	**%**
4 elaborated components							
	3+2+2+2			2	1.0 %	2	0.9 %
	2+2+2+2	1	1.5 %			1	0.5 %
3 elaborated components							
	4+4+2+1	1	1.5 %	1	0.5 %		
	4+3+2+1					1	0.5 %
	4+2+2+1					1	0.5 %
	3+3+2+1	2	3.1 %	6	3.0 %	7	3.3 %
	3+2+2+1	2	3.1 %	6	3.0 %	5	2.4 %
	2+2+2+1	8	12.3 %	15	7.4 %	13	6.1 %
2 elaborated components							
	4+3+1+1			1	0.5 %	1	0.5 %
	4+2+1+1	1	1.5 %	2	1.0 %	3	1.4 %
	3+3+1+1	1	1.5 %	2	1.0 %	2	0.9 %
	3+2+1+1	2	3.1 %	21	10.4 %	28	13.2 %
	2+2+1+1	13	20.0 %	47	23.3 %	52	24.5 %
1 elaborated component							
	5+1+1+1			1	0.5 %		
	4+1+1+1					2	0.9 %
	3+1+1+1	3	4.6 %	10	5.0 %	16	7.5 %
	2+1+1+1	19	29.2 %	56	27.7 %	58	27.4 %
Unelaborated utterances							
	1+1+1+1	12	18.5 %	32	15.8 %	20	9.4 %
Total		**65**	**99.9 %**	**202**	**100.1 %**	**212**	**99.9 %**

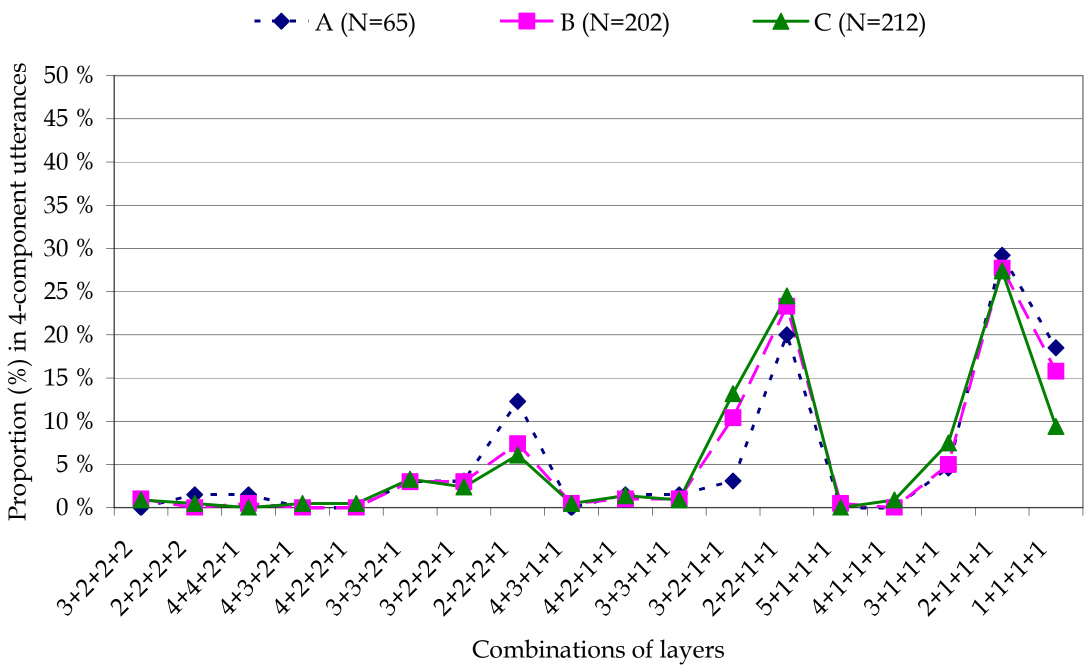


FIGURE 41 The proportions of the different layer combinations in the four-component utterances.

The analysis of the number of elaborated components and the combination of layers in utterances is a step forward in searching for factors contributing to structural complexity. The present results give further support to these previously obtained in the analysis of maximum number of layers. The strong tendency towards uniformity between the groups in step with an increase in the number of components is also evident when all components and layers are considered. These results give further evidence of a trade-off effect as well. The more syntactic components there are in an utterance the less elaborated the components are.

The starting point for this analysis was Hypotheses 1–5 which were originally based on the contradictory results from MLU and IPSyn. To a greater of lesser extent all of the hypotheses were supported by the results of the analyses of elaboration in utterances. Hypotheses 1 and 2 suggest that short utterance length combined with varying structural resources results in utterances with an immature multilayered structure. The results of the present analysis show that children in group A i) elaborate fewer components per utterance than children in the other groups, and ii) prefer the simplest elaboration alternatives 2+1, 2+1+1, and 2+1+1+1 alongside unelaborated utterances. Thus the performance of group A is in line with Hypotheses 1 and 2.

Hypothesis 3 suggests that high MLU combined with large structural resources results in utterances with more and more components including deep elaboration. In group C, the target group of the Hypothesis, this seems to be true of two- and partly of three-component utterances. The utterances include more elaborated components than with the two other groups, and elaboration preferences are not restricted to the simplest alternatives. However, the four-component utterances show a different trend. Although there are more of them in group C than in other groups, qualitatively they are not more complex than those produced by children in the other groups. Thus, deep elaboration in several components covers only utterances with three or fewer components.

Hypothesis 4 focuses on children in group B. "Combining more and more resources in utterances" seems to be mostly true but it has different manifestations in distinct utterance categories. In number of elaborated components per utterance, group B is similar to group C. In other words, elaboration has spread to more components than in group A. In two-component utterances group B very much resembles group C also in depth of elaboration, but in three-component utterances the combinations of elaborated components do not reach the same level of complexity as in productions by group C. "Combining of resources" is now only horizontal. Finally, in four-component utterances, all the children perform very similarly, and only Hypotheses 1 and 2 concerning group A are supported.

Hypothesis 5 gives an overall suggestion of how morphosyntactic complexity develops. The results presented in this section support the idea of a stepwise development introduced in the hypothesis, but moreover, they strongly emphasize that elaboration spreads first into more components and only after that into deeper layers.

Data selection is a factor which may cause biased results. It is likely that at least the children who have the highest MLU values lose some of their two-component utterances because these have too few morphemes to qualify amongst the 80 longest utterances. To discover the possible consequences of the method of data selection the children's data was checked in terms of how short utterances are included.

The shortest possible two-component utterance must have two morphemes. Thus, to be sure that all two-component utterances produced during the recording are included in the selected data, the shortest utterance to qualify must be monomorphemic. Only in the cases of Aleksi (ID 31; member of groups A and B) and Seppo (35; member of groups A and B) are the shortest utterances monomorphemic and thus all their two-component utterances have passed the selection criteria. This already shows that all groups are affected by the data selection procedure, and not only those with the highest MLU values.

If the shortest utterance in a child's data is bimorphemic, we know that at least all the two-component utterances including elaboration and all the three-component utterances have qualified for the final data. However, there is a possibility that some of the unelaborated two-component utterances are left outside the data. A survey of the data reveals that Tatu (17), Saku (18) and Sampo (36) in group A and Mika (15) in group B as well as Leo (34) belonging to groups A and B have bimorphemic utterances in their data.

If the shortest utterance in a child's data is trimorphemic, this means that all unelaborated two-component utterances are excluded from the selected final data. There may also be loss of two-component utterances with one elaborated component and unelaborated three-component utterances. Children who belong to this category are Ronja (6), Anna (22) and Elisa (29) in group A, Joel

(13) and Sanni (24) in group B and finally Anniina (23) and Janna (27) in group C. At this point in the survey all the children in group A have been listed according to their shortest qualifying utterances. This means that even with group A some of the two-component and probably even three-component utterances are not included and this implies that the results are at least to some extent biased towards elaborated utterances. For group C the bias is much stronger because in the case of two-component utterances all unelaborated utterances are left out of the analysis because none of them qualified among the 80 longest utterances.

If the shortest utterances consist of four morphemes, this implies that only elaborated two- and three-component utterances qualify and even some of the unelaborated four-component utterances are probably excluded. Aino (8), Santeri (10) and Pekka (38) in group B and Tuija (1), Laura (3), Tuomo (19) and Elina (30) in group C as well as Jaakko (11) representing both groups belong to this category. It seems that the data selection emphasizes elaboration and that the children with mostly long utterances benefit most from this.

Finally, there are two children Riikka (2) and Taru (7), representing groups B and C whose shortest utterance in their selected data sample is five morphemes long. This means that all of their unelaborated two-, three- and four-component utterances are left out of the data as well as all two-component utterances with one elaborated component. Other elaborated utterances may also have been left out the selection. These two children are responsible for increasing the proportion of deeply elaborated utterances in groups B and C.

The data selection affects the results and especially emphasizes elaboration, mostly in groups B and C but also in group A. Thus the effect is present throughout the data and not only in group C, for example, although in that group the effect is the strongest.

7.4 The layered structure in NP, V and ADVP

7.4.1 Noun phrases

Thus far, particularly the two- and three-component utterances have revealed an interesting difference between the three groups. Differences in layered structure are most evident in two2-component utterances and they gradually disappear as the number of components increases. The same pattern is repeated irrespective of whether the maximum number of layers or the layered structure of utterances in general is considered. However, the effect of the data selection procedure must be kept in mind when the results are considered. The next question is whether any particular component type is more responsible for the differences in elaboration, which concerns how deeply children elaborate distinct component types.

In the component analysis it was found that NP, V, and ADVP account for approximately 90% of all the component tokens in each group's data, and the distribution of these component types was very similar in each of the groups. Moreover, NP and V were defined as the most powerful components in terms of elaboration potential. Thus, it is logical to think that these component types are the most likely positions for structural differences to appear. Next, I will take a closer look at the two- and three-component utterances focusing on the layered structure of NP-, V-, and ADVP-components to find out if the differences between the groups are centred on a particular component type. The four-component utterances are left out of the analysis because there were no group differences in their case.

In the two-component utterances, the groups do differ substantially from each other when the layered structure of NP-components is focused on. This is shown in Table 24 and Figure 42. Almost 65% of the noun phrases are unelaborated in the utterances produced by the children in group A, whereas in group C the proportion of unelaborated NP-components is only 13%. In group B, the proportions of unelaborated and elaborated (two-, three- and four-layered) noun phrases are quite close to each other (47,6% and 52,4% respectively). The contrast between elaborated and unelaborated tokens alone would be enough to conclude that the elaboration of the NP-components does interact with differences in structural complexity in groups A, B, and C. However, a closer look at the elaborated noun phrases gives even further evidence of the case.

TABLE 24 The layered structure of NP-components in two- and three-component utterances. The proportions of NP-tokens of different layers are shown in graphic form in Figures 42 and 43.

		Group A		Group B		Group C	
		NP tokens	%	**NP tokens**	%	**NP tokens**	%
Two-component U							
	1 layer	139	64.4 %	81	47.7 %	7	13.0 %
	2 layers	61	28.2 %	55	32.4 %	25	46.3 %
	3 layers	15	6.9 %	30	17.6 %	19	35.2 %
	4 layers	1	0.5 %	4	2.4 %	3	5.6 %
Total		216	100.0 %	170	100.1 %	54	100.1 %
Three-component U							
	1 layer	286	70.4 %	394	57.3 %	206	45.8 %
	2 layers	99	24.4 %	228	33.1 %	168	37.3 %
	3 layers	20	4.9 %	62	9.0 %	69	15.3 %
	4 layers	1	0.2 %	4	0.6 %	7	1.6 %
Total		406	99.9 %	688	100.0 %	450	100.0 %

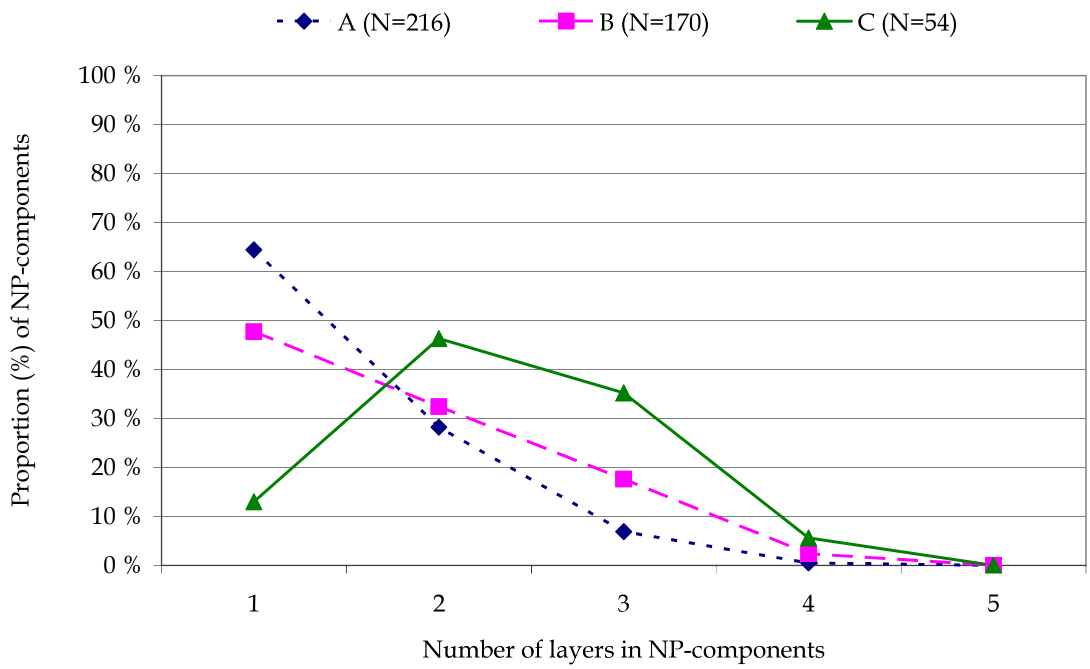


FIGURE 42 The layered structure of the noun phrases in two-component utterances.

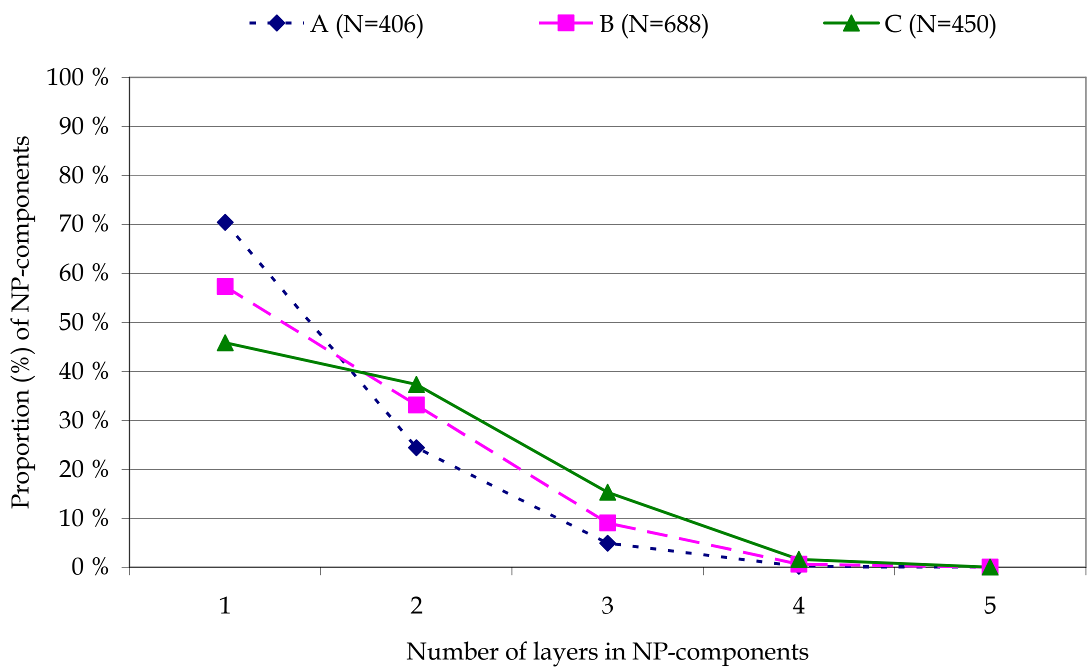


FIGURE 43 The layered structure of the noun phrases in three-component utterances.

Almost half of the NPs produced by the children in group C in two-component utterances are two-layered constructions, and more than one third are three-layered. In these two categories, the children in groups A and B are far behind those in group C. In three-component utterances, however, all three groups elaborate NPs in nearly the same ways, as Table 24 and Figure 43 show. This is mostly due to decreases in elaboration in group C, whereas elaborations in groups A and B are nearly the same as found in two-component utterances. The relation between the elaborated and unelaborated tokens of NP components in the group C is now notably different. The result is that the profiles of the three groups are much more similar. However, despite this decrease in elaboration, the children in group C still perform better in the categories of two- and three-layered NP components. The superior performance of group C compared with the two other groups is evident but, in addition to that, group B now also stands out from the other two groups. Its place between them is detectible in both Figures 42 and 43.

These results suggest that children make use of the elaboration potential within noun phrases to various degrees, especially in two-component utterances but also in three-component ones. It is not a question of not having some of the elaboration types in a group's performance. More likely, the differences in layered structure of NPs are due to differences in which of the types of elaboration is most used. Although the children in group A do produce elaborated NPs to two, three and even four layers they nevertheless clearly prefer unelaborated noun phrases. The children in group B are moving more towards elaborated NPs, but still approximately half of the NPs are unelaborated. The strongest use of elaboration is in group C. This is most evident in the two-component utterances, whereas the addition of another component (from two to three) weakens the amount of elaboration.

7.4.2 Verb constructions

Where verb constructions are concerned the results are somewhat different from those reported for noun phrases. First, as shown in Table 25 and Figure 37, group C is one layer ahead of groups A and B. Almost 50% of all verb constructions produced by the children in group C are three-layered components whereas the children in groups A and B produce mostly two-layered ones. However, the differences disappear almost totally when the three-component utterances (Table 25 and Figure 45) are considered. Another noteworthy factor is a greatly increased proportion of unelaborated verb constructions in three-component utterances. For noun phrases only group C showed a big increase in unelaborated tokens, but in verb constructions, all the groups behave similarly.

The investigation of the layered structure of noun phrases gave reasons to believe that there are three phases of elaboration and that each group represents a separate phase. However, the results for verb constructions suggest a different situation. The groups differ only marginally for three-component utterances. If the proportion of unelaborated components is the criterion for separating developmental phases there seem to be three of them, but for verbs no particular group clearly represents any phase. It is more likely that the groups are situated somewhere in between, behaving differently in distinct utterance categories.

TABLE 25 The layered structure of V-components in two- and three-component utterances. The proportions of V-tokens of different layers appear in graphic form in Figures 44 and 45.

		Group A		Group B		Group C	
		V tokens	%	V tokens	%	V tokens	%
Two-component U							
	1 layer	47	26.6 %	37	26.2 %	8	14.5 %
	2 layers	76	42.9 %	58	41.1 %	10	18.2 %
	3 layers	43	24.3 %	36	25.5 %	26	47.3 %
	4 layers	11	6.2 %	8	5.7 %	8	14.5 %
	5 layers	0	0.0 %	2	1.4 %	3	5.5 %
Total		177	100.0 %	141	99.9 %	55	100.0 %
Three-component U							
	1 layer	165	58.1 %	262	51.9 %	177	49.3 %
	2 layers	83	29.2 %	152	30.1 %	98	27.3 %
	3 layers	34	12.0 %	76	15.0 %	73	20.3 %
	4 layers	2	0.7 %	13	2.6 %	10	2.8 %
	5 layers	0	0.0 %	2	0.4 %	1	0.3 %
Total		284	100.0 %	505	100.0 %	359	100.0 %

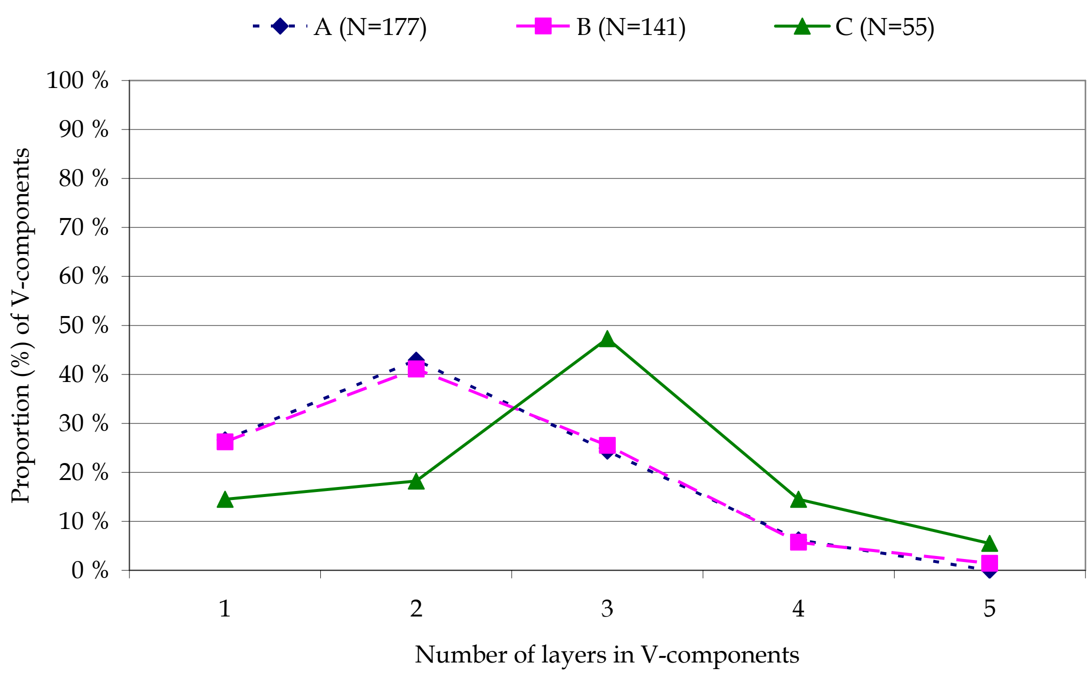


FIGURE 44 The layered structure of the verb constructions in two-component utterances.

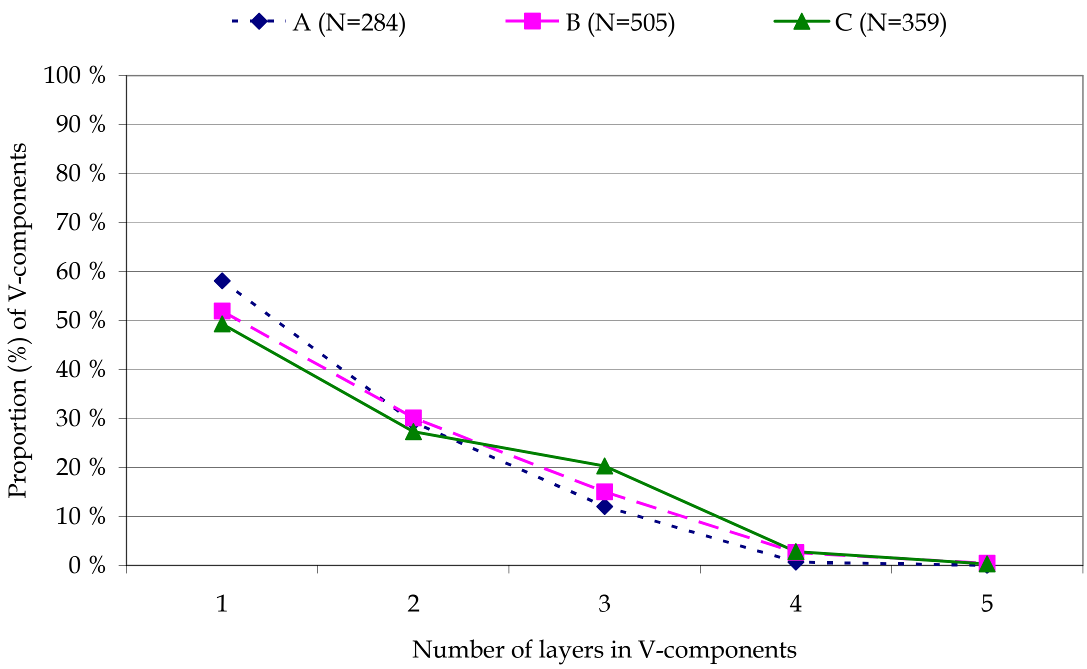


FIGURE 45 The layered structure of the verb constructions in three-component utterances.

In the case of three-component utterances, all three groups seem to be in a phase where they prefer unelaborated verb constructions, whereas with two-component utterances, they have moved into the next two phases, where the majority of verb constructions are elaborated. This time, the depth of elaboration is the factor that separates group C from the others. Thus, under favourable conditions (that is when there are not too many components to elaborate) separate phases of complexity seem to be evident in the data.

7.4.3 Adverb phrases

Adverb phrases are different from noun phrases and verb constructions in their potential for elaboration. There are far fewer possibilities to inflect or extend them syntactically than there are for NPs and Vs. Thus, the results of layered structure of adverb phrases shown in Table 26 and Figures 46 and 47 are as expected. There is a clear borderline between constructions consisting of one or two layers and constructions with three or more layers, the former being the dominating type of adverb phrases with over 90% in all the three groups. However, there are two details in the results that differ from those for noun phrases and verb constructions. First, groups B and C maintain their profile, whereas the performance of group A changes when the number of components increases. In the case of noun phrases and verb constructions the situation was the opposite, with group C being more similar to A's profile. Another difference is that groups B and C are very similar. With noun phrases, group B was more similar to group A, and with verb constructions this similarity was even greater, especially for two-component utterances.

TABLE 26 The layered structure of ADVP-components in two- and three-component utterances. The proportions of ADVP-tokens of different layers appear in graphic form in Figures 46 and 47.

		Group A		Group B		Group C	
		ADVP	%	ADVP	%	ADVP	%
2-component U							
	1 layer	37	52.9 %	26	59.1 %	7	50.0 %
	2 layers	29	41.4 %	18	40.9 %	6	42.9 %
	3 layers	4	5.7 %	0	0.0 %	1	7.1 %
	4 layers	0	0.0 %	0	0.0 %	0	0.0 %
Total		70	100.0 %	44	100.0 %	14	100.0 %
3-component U							
	1 layer	122	68.2 %	111	50.0 %	87	46.0 %
	2 layers	54	30.2 %	96	43.2 %	88	46.6 %
	3 layers	2	1.1 %	12	5.4 %	14	7.4 %
	4 layers	1	0.6 %	3	1.4 %	0	0.0 %
Total		179	100.1 %	222	100.0 %	189	100.0 %

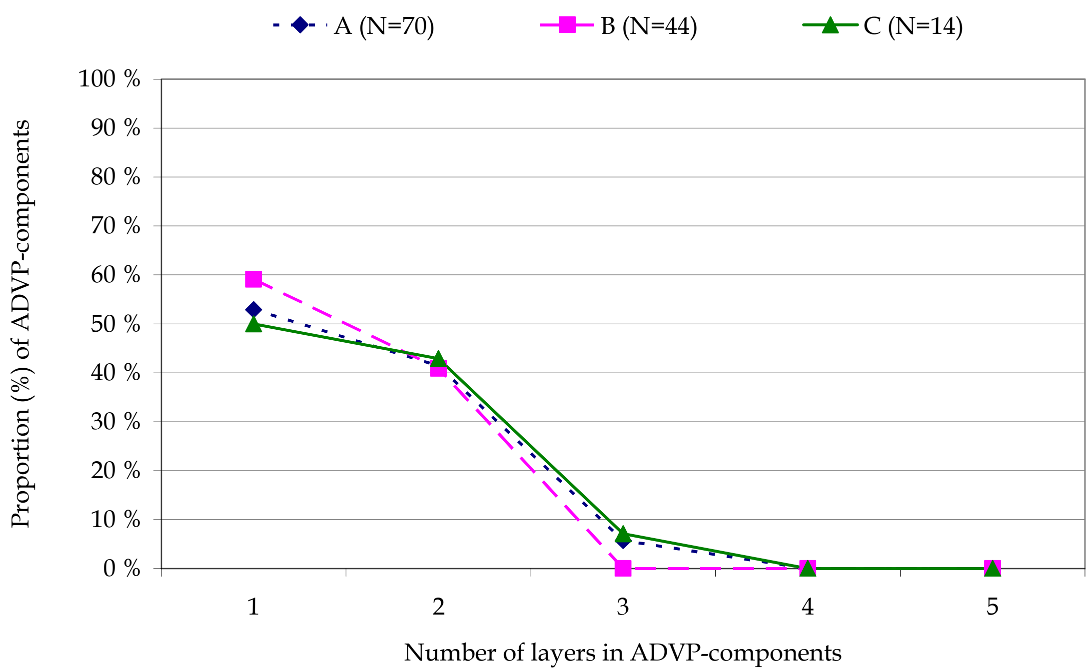


FIGURE 46 The layered structure of the adverb phrases in two-component utterances.

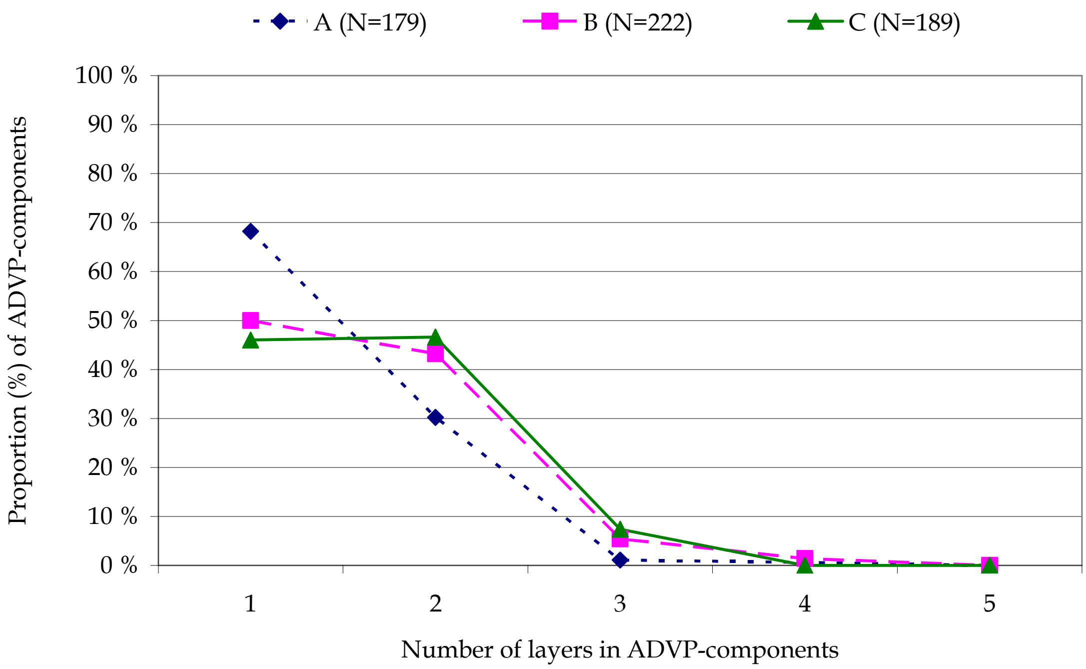


FIGURE 47 The layered structure of the adverb phrases in three-component utterances.

7.4.4 The layered structure of components in the light of a trade-off effect and the componential structure of utterances

The results concerning the layered structures of NP-, V-, and ADVP-components show that similar structures do occur in all groups. It is the proportion of each layered type that varies across the groups, and each group prefers certain types of layering. Thus, the group differences are relative rather than dichotomous between the existence and non-existence of a given type of layering. The variation between the groups is most evident for two-component structures, and this result is consistent with the results of the analysis of the maximum number of layers and the layered structure of utterances in general. The children in group C produce more multilayered components than the children in groups A and B. This is the case especially for noun phrase components but also for verb constructions. In practice this indicates that the groups use the elaboration potential of noun phrases and verb constructions in different ways. In the case of adverb phrases the groups behave very similarly, which was expected, since adverb phrases are expected to have a rather narrow potential for elaboration and thus there is not much scope for variation across the groups.

The utterances with only two components seem to be the ones where especially the noun phrases and verb constructions reveal the largest differences between the groups. For both components the children in group C seem to be able to produce a larger proportion of multilayered tokens than the children in groups A and B. However, this is also the place where the data selection procedure enters the picture. Restricting the data to the 80 longest utterances from each child affected especially group C (although other groups were affected too), and caused a loss of less elaborated two-component and even three-component utterances. The bias in the data shows also in the very small number of two-component utterances in general. Thus the results for two-component utterances are better than they should be, especially for group C, and probably give too advanced an impression of the language used by the children.

Where three-component utterances are concerned the results for the groups are more similar, especially for verb constructions. This group similarity is mostly due to differences in the utterances of group C which is now remarkably similar to the other two groups. At the three-component level groups A and B maintain approximately the same level of layering in the noun phrases as in the two-component utterances; they also change less in their verb constructions than group C. It looks as if the children in group C have succeeded in establishing more complex layered structures, but only when the number of components in utterances does not exceed two. However, it must be kept in mind that the data selection is likely to exaggerate the difference across two- and three-component utterances especially for group C. Groups A and B have mastered an elementary stage of elaboration but a smaller number of components in an utterance does not yet help much to expand the elaboration, as is the case for group C. Where adverb phrases are concerned a similar phenomenon can be observed, but when the performance in two- and three-component utterances is compared, the group showing the biggest differences is group A. On this basis it could be concluded that only group A is still in a phase where an increase in number of components in an utterance affects the layered structure of adverb phrases, whereas the other two groups have reached a phase where the number of components no longer affects it.

TABLE 27 A three-phased approach to the growth of structural complexity in the layered structure of noun phrases (NP), verb structures (V) and adverb phrases (ADVP).

Phase	Identification	Component	Group on this phase
1	an elementary stage of elaboration is established;	NP	A and B
	limitation to a small number of components does	V	A and B
	not help in expanding elaboration	ADVP	
2	an elementary stage of elaboration is established;	NP	C
	a small number of components does help	V	C
	somewhat in expanding elaboration	ADVP	A
3	a stage of elaboration where the number of	NP	
	components does not have much effect	V	
	on the layered structure of components	ADVP	B and C

The growth in complexity in the elaboration of syntactic components seems to proceed through three separate phases. The variables that contribute to moving from one phase to the next are the total number of components in an utterance, the maintenance of layered structure in the components and the elaboration potential of a component. The phases are listed in Table 27.

The views that children are going through specific phases in their acquisition process and that the addition of complexity in one part of an utterance prejudices growth in some other area is consistent with the trade-off effect (Bock 1982; Crystal 1987). Progressing from one phase to another is a signal that the trade-off effect is becoming weaker. The results seem to suggest that the elaboration potential of a component has an impact on the weakening. The trade-off effect is weaker in components such as adverb phrase which have a narrow potential for elaboration, whereas the component types with a wider range of elaboration potential are a locus for a trade-off effect for a longer period of time.

Although the trade-off effect seems to offer a good explanation for the behaviour of the groups, functional factors must also be considered. It was reported earlier that the most common combinations of syntactic components in the three-component utterances were V + NP + AdvP and V + NP + NP[13], and they accounted for approximately 70 % of the three-component utterances in all the three groups. The former is suitable for expressing location (Examples 19 and 20) and the latter for introducing and naming things (Examples 21 and 22). Both of these functions are essential during a play session, especially when a child gets new toys to play with. Each recorded play session in the data was divided into three sections, and each shift to a new section was effected by bringing in a new set of toys. This, of course, increased the need for naming and introducing things.

Example 19

1.	ADVP ADV:DEM	V	NP N
2.	adess		pl&partit

Siellä *on* *lelu`ja.*
Sie-llä on lelu-ja.
There-ADESS is toy-PL&PARTIT.
'There are toys in there. ' (31 Aleksi; 10/38)

Example 20

1.	ADVP ADV:INT	V		NP
2.	iness		PRO:DEM	N
3.			gen	

Missä *on* *sel* *lautanen?*
Mi-ssä on sen lautanen?
Wh-INESS is it-GEN plate?
'Where is its plate?' (19 Tuomo; 35/144)

13 The symbols V + NP + AdvP and V + NP + NP are intended to include all possible orders in which the three components can occur. Thus, for example V + NP + NP can actually manifest as NP + NP + V, NP + V + NP, and V + NP +NP.

Example 21

1.	NP PRO:DEM	V		NP
2.			A	N

Tä *o* *uut* *poni.*
Tämä on uusi poni.
This is new pony.
'This is a new pony.' (8 Aino; 54/923)

Example 22

1.	NP PRO:DEM	V	NP N

Se *on* *i/s`i:.*
Se on isi.
It is Daddy.
'It is Daddy.' (13 Joel; 68/435)

Yet another important function for the three-component utterances is to describe action and tell what is happening or will happen next. Children begin their narrative utterances with adverbs such as *sitten* 'then, next, after that' as in Examples 23–25 from Jaakko (11).

Example 23

1.	ADVP ADV	V	NP N
2.		pass	adess

Hitte *ajataah* *hakko3lilla.*
Sitten aje-taan traktori-lla.
Then drive-PASS tractor-ADESS.
'Next, we'll drive a tractor.' (11 Jaakko; 54/181)

Example 24

1.	ADVP ADV	V	NP N
2.		pass	

Hitte *ot`etaah* *hiekkaa.*
Sitten ote-taan hiekkaa.
Then take-PASS sand.
'Then, we'll take some sand.' (11 Jaakko; 55/189)

Example 25

1.	ADVP ADV	V	NP N
2.		pass	adess

Hitte *hyyää* *haalukalla.*
Sitten syö-dään haaruka-lla.
Then eat-PASS fork-ADESS.
'Then, we'll eat with a fork.' (11 Jaakko; 56/210)

What is typical of all these utterances is that the verb is in a simple form. In the existential utterances and in the naming function the verb is a copula, and even in describing the play, the "here and now" situation favours the present tense and only a little other elaboration. This is one possible reason for the relative lack of elaboration of verb constructions in the three-component utterances. While all three groups preferred elaborated verb constructions in two-component utterances, in three-component utterances the situation was very different, and the proportion of unelaborated verbs was substantially larger.

Expressing location as well as introducing toys and naming things may also affect the noun phrases. A noun phrase is used to name a toy in locative utterances. The names of the toys are in the nominative case, and instead of inflected NPs, the proadverbs such as *siellä* 'in there', *tuolla* 'over there' and *täällä* 'in here' are used for expressing location. This is possible and convenient in a situation where both participants see the whole scene, and there is no need to describe location more precisely, since an adverb together with, say, a pointing gesture is enough to convey the meaning. This, of course, decreases the need for NPs to be elaborated, and this is reflected in the results as large number of unelaborated noun phrases in locative utterances.

The trade-off effect is thus not the only explanation for changes in the layered structure when the number of components increases. The verbal expressions of a child only acquiring his or her first language have also functional frames, determined by the situation a child is involved in. Thus, the final version of an utterance is a product of interaction between skills and context.

7.5 Summary

The layered structure of a syntactic component and of an utterance is a consequence of the use of morphological and syntactic devices in elaborating syntactic components. Thus, it creates a frame for morphosyntactic complexity. In this chapter the layered structure of utterances was studied from three angles. First, the maximum depth of layers was surveyed to find out how far children can go in elaborating components. Secondly, the extent of elaboration was investigated to see how many and how deeply elaborated components are combined in utterances. And finally, noun phrases, verb constructions, and adverb phrases were focused on, and the layered structure of these most heavily used syntactic components was studied to see how their elaboration potential is exploited. The survey of layered structure concentrated on the two-, three- and four-component utterances since they account for almost 90% of all utterances produced in each group.

The basic aim of this study is to discover what kinds of differences there are between groups A, B and C. The differences found so far are not structural features that one group has acquired while the others have not. The children in all groups produce similar numbers of components in their utterances as well as similarly layered structures, but preferences do vary. In other words, there seems to be no absolute difference in morphosyntactic complexity but some structures are more heavily represented in one group than in the others.

In Hypotheses 1 and 2 it was predicted that in group A the layered structure would remain at a rather shallow level. The results support these hypotheses. In more than half the two-, three- and four-component utterances produced by the children in group A, at most two layers are reached. In actual elaboration this means either only one inflection or one syntactic extension in a syntactic component. Furthermore, it was hypothesised that elaboration is not spread widely within the utterances. This prediction too turned out to be correct in general terms, since most of these children's two- and three-component utterances had only one elaborated component.

In the case of group C, it was hypothesised (Hypothesis 3) that no factors would limit the layered structure of components and utterances. Thus, more complex structures were expected to occur in utterances produced by children in group C than by other children. For two-component utterances this hypothesis turned out to be correct. Elaboration is both deeper, with much use of three- and four-layered structures, and wider, with emphasis on two elaborated components. However, in three-component utterances the maximum depth has shifted to two- and three-layered structures equally. In four-component utterances there is a clear dominance of two-layered structures. Exactly the same phenomenon occurs in the case of elaboration spreading within utterances, but on a smaller scale. Thus the children in group C do not maintain the structural complexity of their utterances when the number of components increases.

Group B is situated between A and C in MLU values and IPSyn scores. This group was seen as a kind of a bridge between the other two groups. In Hypothesis 4 it was predicted that these children would combining more structures and this proved to be the case in the longer utterances detected by MLU. However, two different kinds of combination are possible, and they may occur either simultaneously or separately. First, adding new structures may be horizontal, that is, more components are elaborated. This would not necessarily show up in the maximum layer survey but it would in the investigation of elaboration spread. Secondly, adding more elaboration may be vertical, that is, several elaboration devices can be combined within a single component. This would generate higher figures in the study of maximum number of layers. The hypothesis predicting either adding more elaboration or combining more structures in utterances turned out to be correct and the investigation of layered structure revealed that it was more a question of horizontal combining, although in two-component utterances, some indication of vertical elaboration is also present.

Group B shows children moving towards a subsequent stage. It has features similar to those of both groups A and C but it also takes its own place between the two other groups in, for example, the maximum number of layer results for two-component utterances. In my opinion, group B gives evidence of development that does not happen on all fronts simultaneously; more likely different features develop at their own pace. This picture of piecemeal growth in structural complexity gets further support from the findings for layered structures in the noun phrases and verb constructions. They are both seen as components with a wide potential for elaboration. However, they behave very differently in the surveys of layered structure. With all three groups, verb constructions are elaborated in the majority of the two-component utterances, whereas only with group C are the noun phrases elaborated in the same way. The children in the other two groups produce mostly unelaborated noun phrases. It seems that children begin to use the potential of verb constructions earlier than that of noun phrases.

The developmental pace and structural complexity of utterances intertwine in an interesting way. The children in group A produce utterances that can be characterized as rather similar to each other in terms of layered structure. An increase in the number of syntactic components does not significantly affect the layered structure of the components or the number of elaborated components per utterance. The performance of group A is therefore stable. However, especially in group C, but in group B as well, a different kind of behaviour is evident. In two-component utterances, group C shows far more complex structures than the other groups, but when the number of components increases the utterances begin to resemble more and more those of the children in group A. The children in group C seem to have advanced skills only when there are very few components in an utterance, and they rapidly return to the phase in which group A constantly remains. A convenient and an easy way to explain such a change in utterance complexity is the trade-off effect which helps to avoid situations where the total linguistic load of an utterance exceeds a child's capacity for processing and producing it. Thus, when the number of components increases their elaboration cannot be so great in order to maintain an acceptable level of structural complexity. The fact that the trade-off effect is evident mostly in productions by group C suggests that these children are moving on to a level beyond the phase in which group A still firmly remains. They can now expand elaboration, especially in two-component utterances, whereas for the children in group A a small number of components does not yet lessen the complexity load enough to help in expanding elaboration.

The trade-off effect seems to explain quite well the differences in group performance regarding the number of components. However, there may be yet another factor that is intertwined with developmental pace and structural complexity, namely the typical functions of utterances during a play session. What seems to be balancing of linguistic load might also be a reflection of the functions for which the utterances are commonly used. A three-component utterance is perfect for naming and introducing things. When a child gets toys he/she is not already familiar with, the first wish is to name them and introduce them to the other participants in the situation. In my data, bringing a new set of toys into the session was repeated three times, thus raising the need for expressing such functions. For such utterances, it is typical to have a simple, copula-like verb and noun phrases consisting of names for things or demonstrative pronouns. If *This is X* -type utterances are common among the three-component utterances, it is not surprising that the survey of layered structure indicates features that are typical of the trade-off effect: the number of components increases by one but elaboration decreases substantially.

There is yet another type of utterance that may similarly affect layered structure and has connections to a rather low level of noun phrase elaboration in general. Expressions referring to location (*X is in Y*) and existentiality (*There is an X in Y*) are also very typical in a play session. What is distinctive about these sessions is that because all the participants are able to see the whole scene and thus most of the locations it is possible to refer to them with a proadverb, sometimes supplemented by pointing. The location which is referred to is disambiguated by the context and the actions in the situation, and there is no need to express the location by an actual noun inflected in an appropriate locative case. This reduces the need for elaborated noun phrases and increases the use of adverb phrases respectively.

In this chapter the role of data selection as a factor capable of biasing the results has been discussed on several occasions. In this study the data was limited to the 80 longest utterances from each child. This restriction has affected the data for all three groups but with the greatest impact in the case of group C. In practice this means that the results may be biased, especially towards emphasizing the elaboration of two-component utterances at the expense of unelaborated ones. The data selection is probably one reason why group C performs so differently in two- and three-component utterances. The difference between these two utterance categories may be exaggerated because the data is more biased in the case of two- than three-component utterances.

The results of this investigation of the layered structure of utterances and components together with the interpretations made on the basis of the results lead to two major conclusions. The first is that despite all the differences that different scales of measurement may uncover in children's linguistic behaviour, the underlying structures are still rather similar from group to group. The differences are more likely reflections of the children's differing skill in taking advantage of the capacities they all have. The other major conclusion is that although we are only concerned about structural features and especially about growth in the complexity of structural features, the functional and situational factors are still there. Language is never sterilized in this sense, and thus, the contribution of these factors has to be considered in the interpretation of the results.

8 THE INNER STRUCTURE OF SYNTACTIC COMPONENTS

In this chapter a detailed analysis of utterances consisting of two, three and four components continues. So far it has been found that all three groups (A, B and C) prefer similar component combinations, and these are mostly mixtures of NP, V and AdvP (Chapter 6). Thus the number of components in utterances or the type of components used did not differentiate the groups.

The next subject for investigation was the elaboration of components. In Chapter 7 two approaches were taken to this issue, investigating first how many elaboration layers there are the most elaborated component in each utterance (depth of elaboration; Section 7.2), and secondly how many components are elaborated within each utterance and how many layers each of them has (breadth of elaboration; Section 7.3).

In this chapter the investigation into the elaboration of the components continues, focusing in detail on the means whereby two-, three-, four- and even five-layered component structures are created. Do groups A, B, and C differ in the kind of elaboration they use to create, for example, a component with three layers and are these three-layered components elaborated differently are in a two-component utterance than when part of three- or four-component utterances?

8.1 The purpose of the analysis

So far the differences between groups A, B and C have become most evident in two-component utterances where groups B and C have more layers in the components than group A. When the number of components in utterances increases from two to three the differences become substantially smaller. In four-component utterances, the groups perform almost identically with respect to the maximum number of layers, in layers of different component types, and in the layered structure of whole utterances. However, the layered structure of components and utterances is a result of elaboration and the same number of layers may be attained through different kinds of elaboration. In terms of structural complexity the various routes to a given number of layers are not equivalent. Thus, an analysis focusing on the number of layers in components and the number of layered (elaborated) components within utterances gives only a rough estimate of structural complexity rather than a detailed description.

The purpose of this chapter is to go beyond the rough estimates gained in the previous layer analysis. This will be done by analysing the inner structures of syntactic components and examining which elaboration devices are used for each component. The results will shed light on two new aspects of the components through exposing the variation within each layer and investigating the elaboration devices in the components used by each group.

Another issue that has been taken up several times in discussion of the results is the question of trade-off effects, which seem to be an appropriate way to explain the increasing similarity in the layered structures in the children's productions, especially in the case of group C. A kind of structural balance in utterances is maintained by decreasing complexity in layers when the number of components is increased. However, certain pragmatic factors may also contribute to the final results. It is possible that a similar tendency toward equalizing structural complexity is also present in the inner structure of the components. This has not yet been shown in the analysis of layered structure since within a single layer there are several options for elaboration, some simpler than others. It will be interesting to see whether a four-layered component in a four-component utterance tends to be simpler than a four-layered component in a two-component utterance. If this is the case, the trade-off effect will get further support from the inner structure of the components.

Until now the emphasis of this study has been on investigating item by item the manifestation of morphosyntactic complexity in children's productions. In this chapter this emphasis will be maintained. The questions that need to be answered are: what kind of morphosyntactic structures are present on levels of two, three, four and five layers, and which of them are produced by the groups.

So far the relationship between the quantitative analyses of MLU and IPSyn and the structural complexity exposed by Utterance Analysis has not been discussed. Now it is finally time to consider this issue. The discussion will require concentrating on the individuals within each group and comparing the results of this analysis with MLU values and IPSyn scores child by child.

The elaboration of components forms the core of structural complexity in this study. This is the factor that creates layers in the components and this is why it is important to investigate it in detail and to look for connections with MLU and IPSyn. This is also the factor that may vary most between children in the same group. Thus it is a more effective way of comparing results with other measuring scales than layered structure alone would be.

Before taking up these questions in detail it is necessary to pay attention to structural complexity. To make any estimates of growth in complexity and of mutual relations in terms of complexity it is necessary to consider what makes some structures more complex than others. On what grounds can structures be placed on a scale of complexity? What kind of a scale of complexity do children create with their utterances and the components in them?

8.2 Elaboration types

8.2.1 Some remarks on the notation of elaboration types

All the results in this chapter start from Utterance Analysis, where each component is placed in a chart of elaboration. Elaboration devices and the relations between them are illustrated by means of rows and columns in the chart. Such a chart is a useful means of revealing the structures in each utterance and component, but it is also rather difficult to refer to in prose. Therefore, the descriptions of components will now be converted into written formulas with the help of square brackets, parentheses, letters and numbers. The principles of this notation are explained in Table 28.

TABLE 28 The symbols used in converting the conventions of Utterance Analysis into a written formula.

Symbol		**Meaning**
[ ]	Square brackets	component boundaries
( )	Parentheses	elaboration boundaries within a component
S2	S + number after the S-symbol	use of a syntactic device in elaboration; the number indicates how many words the elaboration contains
S3		result of elaboration = 3 words
M		use of morphological device in elaboration
M2	number after the M-symbol	the number indicates how many suffixes are attached to a single word
2S2	number before a S-symbol	a doubled use of a syntactic device on a given layer for elaborating different parts of a component
2M	number before an M-symbol	a doubled use of a morphological device for elaboration on a given layer (of two different words of a component)
+	plus	connects a) devices within a layer, b) layers within a component, and c) components within an utterance

In the abstract notation of components each component is surrounded with square brackets and each layer with parentheses[14]. Elaborations are described with symbols which refer to morphological (M) and syntactic (S) devices and to the extent to which they are used to expand a component on a given layer. Examples 26–28 illustrate the conversion from an Utterance Analysis into a written formula:

Example 26

	[M]	+	[S2 + (2M)]
1.	V		PP
2.	pass	N	AD
3.		gen	ill

Mennääl *laktori`n* *sissää.*
Men-nään traktori-n sisä-än.
Go-PASS tractor-GEN inside-ILL.
'Let's go inside the tractor.' (10 Santeri; 35/1235)

In Example 26 there is an utterance consisting of two components, a verb structure (V) and an adposition phrase (PP). Their inner structure is described in square brackets and the brackets are connected with a plus sign to show that these two components form an utterance. In the V-component there is only one elaboration. The verb is inflected in the passive, and the two-layered V-component of this utterance in described as [M]. The adposition phrase component has three layers. The second layer shows a syntactic elaboration: the PP-component consists of two parts (a noun *traktori* 'tractor' and a postposition *sisä-* 'inside'), and this is expressed by S2. However, both parts of PP are further elaborated by inflection (one in the genitive and the other in the illative case), and this is shown on the third layer. Moving to the third layer necessitates the use of parentheses in the notation: [S2 + (2M)] means that the syntactic parts of the PP-component are inflected and there is only one suffix attached to both. The whole 2-component utterance can be expressed as [M] + [S2 + (2M)].

Example 27

			[S2 + ((S2 + (M))+M)]
1.			NP
2.		AP	N
3.	ADV	PRO:ADJ	pl&partit
4.		pl&partit	

Ihan *saman`eisia* *aitoja.*
Ihan samanlais-ia aito-ja
Exactly similar-PL&PARTIT fence-PL&PARTIT.
'Exactly the same kind of fences.' (13 Joel; 10/499)

In Example 27 there is a single-component utterance, but the only component (NP) is elaborated in four layers. The first elaboration on the second layer

14 In fact, parentheses are used to delimit layers only beginning when the third layer is reached. An unelaborated, one-layered component is not bracketed, and in the case of two layers square brackets delimit the layer at the same time as they separate components from each other. Therefore, the third layer is the first one to need an extra marker for separating layers.

shows syntactic elaboration into two parts (a noun and an adjective phrase; S2). The adjective phrase is a further combination of two parts (an adverb and a pronominal adjective) as the third layer shows, and finally, the pronominal adjective is inflected in the partitive plural. The elaboration of the AP-element included in the component on the third and fourth layers is expressed as (S2 + (M)), the outer parenthesis confirming the notation to the AP-element and the inner parenthesis referring to elaboration on the fourth layer. The notation [S2 + (S2 + (M))] accounts for the two elements of the NP-component and the elaboration of the AP-element, but notation of the elaboration of the N-element of NP-component is still lacking. This is done by adding one more M and parentheses yet again to delimit elaboration on the whole third layer. The final description of this single-component utterance is [S2 + ((S2 + (M)) + M)].

Example 28

	1	+	[S2 + (M2 + M)]
1.	INTERJ		NP
2.		PRO:INT	N
3.		partit	pl&partit
4.		clit	
	Hui	*mitähäl*	*leluja?*
	hui	mi-tä-hän	lelu-ja?
	Wow	what-PARTIT-CLIT	toy-PL&PARTIT?

'Wow, what kind of toys?' (2 Riikka; 24/39)

In Example 28 there is a two-component utterance consisting of an interjection and an NP, but only the NP-component is elaborated. The NP-component consists of two elements (interrogative pronoun and noun; S2) which are both inflected, the former with two inflections (M2) and the latter with one (M). The structure of the NP-component is expressed as [S2 + (M2 + M)].

This more convenient way of referring to component structures is not the only benefit of this notation system. It also enables the categorization of the structures and helps to reveal similarities in morphosyntactic features at a more abstract level. Therefore it is an important tool for discussing structural complexity.

In Table 29 there is an overview for all three groups of how many components are elaborated in two-, three-, and four-component utterances and how deeply these components are elaborated. The table also shows the proportions each elaboration type (two-, three-, four- and five-layered elaboration) accounts for in each group and how many elaborated tokens in there are in the data as a whole. Finally, the bottom line shows total number of components (elaborated and unelaborated) in each group and in the whole data. In Sections 8.2.2–8.2.5 each elaboration type (two-, three-, four- and five-layered) will be discussed in detail.

TABLE 29 Overview of frequencies and proportions (%) of the different elaboration types of the elaborated components in groups A, B, and C. The two bottom lines show the total number of elaborated components and the total number of all components in two-, three- and four-component utterances.

	Group A		Group B		Group C		Total	
	Tokens	%	Tokens	%	Tokens	%	Tokens	%
2-layered comp.								
	507	75.2 %	897	70.8 %	688	65.5 %	2092	69.9 %
3-layered comp.								
	145	21.5 %	317	25.0 %	313	29.8 %	775	25.9 %
4-layered comp.								
	22	3.3 %	48	3.8 %	46	4.4 %	116	3.9 %
5-layered comp.								
	0	0.0 %	5	0.4 %	4	0.4 %	9	0.3 %
Elaborated comp.								
	674		1267		1051		2992	
All components								
	1667		2724		2070		6461	

8.2.2 Two-layered elaboration

Any kind of elaboration will develop a component into at least a two-layered structure with the final number of layers depending on the quality of elaboration. To become two-layered a component requires only a single elaboration. Thus, a single suffixation or a combination of, for example, a demonstrative pronoun and a noun in a noun phrase are ways to create a second layer.

In the data there are three different kinds of elaboration that result in two-layered components. These are: the use of a suffix (M), the combination of an uninflected modifier and a head (S2), and the combination of three uninflected words in a phrase (S3). These two-layered elaboration types are illustrated in Examples 29–31.

Example 29
Two-component utterance, both components representing two-layered [M]-elaboration

1.	NP PRO:DEM	V
2.	clit	past

Täm:ki *tuli .*
Tämä-kin tul-i
This-CLIT come-PAST.
‘This one came, too.’ (3 Laura; 80/104)

Example 30
Three-component utterance, NP-component representing two-layered [S2]-elaboration

1.	V	NP PRO		NP
2.			ADJ	N

On *tuo* *iso* *jalkapallo.*
On tuo iso jalkapallo.
Is that big football
'That is a big football.' (18 Saku; 27/1604)

Example 31
Four-component utterance, NP-component representing two-layered [S3]-elaboration

1.	ADVP ADV:INT			ADVP ADV	NP	V
2.		PRO:DEM	PRO:INDEF		N	

Mit`ä *te* *ykkiln* *hittet* *tuo3li* *on?*
Missä se yksi sitten tuoli on?
Where it one then chair is?
'Where's that chair again?' (11 Jaakko; 34/936)

Table 30 shows that two-layered components account for approximately 70% of all the elaborated components produced by the children in all three groups. There are a total of 1774 occasions of [M]-elaboration in the data. The vast majority of elaborations belong to this category. The other two types are used far less, [S2] being the means of elaboration on 249 occasions and [S3] in only 8 components. The high frequency of [M] is not surprising, since it includes all kinds of inflection in all word classes. The radically lower incidence of [S2] and [S3] is not surprising, either, since they both require all elements to be uninflected which means that in the case of noun phrases they are suitable for only a few functions in Finnish sentences, mostly in subject position or in object position together with a passive verb.

TABLE 30 The frequencies of two-layered elaboration types for each group. The proportions (%) of each elaboration type are based on the total number of elaborated component tokens in the group's data consisting of all two- to four-component utterances.

	Group A		**Group B**		**Group C**		**Total**	
Type	Tokens	%	Tokens	%	Tokens	%	Tokens	%
[M]	465	69.0 %	776	61.3 %	593	56.4 %	1834	61.3 %
[S2]	41	6.1 %	117	9.2 %	92	8.8 %	250	8.4 %
[S3]	1	0.2 %	4	0.3 %	3	0.3 %	8	0.3 %
2-layered component tokens (total)								
	507	75.2 %	897	70.8 %	688	65.5 %	2092	69.9 %
Elaborated component tokens (total)								
	674		1267		1051		2992	

In verb constructions there are even fewer possibilities for the use of [S2]-components. These are mostly phrasal verbs (an idiomatic combination of a verb and a particle such as *mennä pois* 'go away' or *pitää kiinni* 'to hold on to'; ISK 2004: 448) or erroneous structures where a required inflection is missing for some reason.

8.2.3 Three-layered elaboration

Transition from a two-layered component to a three-layered one requires a further elaboration of one of more elements in a two-layered component. Therefore the possibilities for expanding a component grow remarkably. Three-layered elaboration types, their frequencies and proportions (%) based on the total number of all elaborated component tokens are shown in Table 30.

There is only one three-layered elaboration type which is based on the use of suffixes alone. This is the [M2] elaboration type, where two suffixes are attached to a single stem. For instance, in Example 30 there is an inessive ending *-ssA* and a particle clitic *-s* attached to the interrogative adverb stem *mi-*. Other possible suffix combinations in noun phrases are two particle clitics, and a plural marker together with either a case ending (with the exception of the partitive plural, c.f. Laalo 2002) or a particle clitic. In verb constructions, suffixes for person and past tense as well as for person and mood can be combined.

TABLE 31 The frequencies of different three-layered elaboration types in each group. The proportions (%) of each elaboration type are based on the total number of elaborated component tokens in the group's data consisting of all two- to four-component utterances.

	Group A		**Group B**		**Group C**		**Total**	
Type	**Tokens**	%	**Tokens**	%	**Tokens**	%	**Tokens**	%
[M2]	32	4.8 %	56	4.4 %	55	5.2 %	143	4.8 %
[S2+(M)]	85	12.6 %	176	13.9 %	155	14.8 %	416	13.9 %
[S2+(2M)]	23	3.4 %	77	6.1 %	94	8.9 %	194	6.5 %
[S2+(S2)]	1	0.2 %	0		5	1.5 %	6	0.2 %
[S2+(S2+M)]	0		2	0.2 %	1	0.1 %	3	0.1 %
[S3+(M)]	1	0.2 %	1	0.1 %	1	0.1 %	3	0.1 %
[S3+(2M)]	1	0.2 %	3	0.2 %	1	0.1 %	5	0.2 %
[S3+(3M)]	2	0.3 %	2	0.2 %	0		4	0.1 %
[S3+(S2+2M)]	0		0		1	0.1 %	1	0.0 %
3-layered component tokens (total)								
	145	21.5 %	317	25.0 %	313	29.8 %	774	25.9 %
Elaborated component tokens (total)								
	674		1267		1051		2992	

Particle clitics may occur together with a suffix for person, tense or mood, as well as with another particle clitic to form an [M2]-elaboration in a verb component. Other component types where [M2]-elaboration is possible are ADVP, INFP, and AP where particle clitics provide a good source of possibilities either alone or together with case suffixes. Although [M2]-elaboration has a lot of potential for expressing meaning and possibilities in combining suffixes, it is rather rare in the data. The total number of component tokens representing [M2]-elaboration is only 143 (4.8 %) out of the total of 2991 elaborated components.

When three-layered elaboration is based on [S2]-elaboration on the second layer, there are four distinct variants. In two of them, [S2+(M)] and [S2+(S2)], only one of the second layer elements is elaborated further. In the remaining two types, [S2+(2M)] and [S2+(S2+M)], both elements are elaborated. Furthermore, in the elaboration type mentioned last, both syntactic and morphological devices are involved, whereas in the other variants only one is in use. Table 31 gives more information about why two-layered [S2]-component tokens are rather rare. [S2+(M)] and [S2+(2M)] are far more frequently used than the other two [S2]-based types ([S2 + (S2)] and [S2 + (S2 + M)]). These elaboration types which include only morphological extensions have substantially more tokens (total of 609), because they are useful, for example, for expressing locality (Example 32), and ownership or belonging together (Example 33). In verb constructions [S2+(M)] and [S2+(2M)] are involved in negation (Example 34) and other constructions with an auxiliary as well as in verb chains (Example 35). Another use for these component types is the perfect tense (Example 36). In fact the perfect tense is a morphological form, but since it is constructed with the help of an auxiliary it contains two separate elements and is classified as S2 and not as M. Other structures where [S2+(M)] and [S2+(2M)] appear are infinitive phrases (Example 37) and adposition phrases (Example 38).

Example 32
[S2+(2M)] expressing locality

1.	NP PRO:INT	V		**NP**
2.			**PRO:DEM**	**N**
3.			**elat**	**elat**

Kuka *luo* *tästä* *tuttipullosta?*
Kuka juo tä-stä tuttipullo-sta?
Who drink this-ELAT feeding bottle-ELAT.
'Who drinks **from this feeding bottle?'** (18 Saku; 6/123)

Example 33
[S2+(M)] expressing belonging together

1.		**NP**	V	INFP	
2.	**N**	**N**		V	3INF
3.	**gen**			1inf	ill

Äitiv *vauva* *hanuu* *mennän* *nukkumaa.*
Äidi-n vauva haluaa men-nä nukku-ma-an.
Mother-GEN baby want go-1INF sleep-3INF-ILL.
'**Mother's baby** wants to go to bed' (3 Laura; 4/615)

Example 34
[S2+(2M)] expressing negation

1.		NP PRO	**V**
2.	**AUX:NEG**		**V**
3.	**1s**		**neg**

Em *mie* *muista .*
En minä muista.
No-1S I remember-NEG.
'**I don't remember**.' (18 Saku; 28/82)

Example 35
[S2+(M)] in a verb chain construction

1.	NP PRO:DEM		ADVP ADV:DEM	**V**	NP N
2.		**AUX**		**3INF**	pl&partit
3.				**iness**	

Se *on* *siitä* *koljavassa* *leNkaita*
Se on siitä korjaa-ma-ssa renka-ita.
It is there repair-3INF-INESS tyre-PL&PARTIT
'It is there, **repairing** tyres.' (19 Tuomo; 11/1194)

Example 36
[S2+(2M)] in the perfect tense

1.	NP PRO			NP	**V**
2.		**AUX**	N(P)	N	**V**
3.		**1s**	gen	iness	**part**

Mää *oon* *Nenoh* *huoneessak* *käyny.*
Minä ole-n Nero-n huonee-ssa käy-nyt.
I have-1s Nero-GEN room-INESS visit-PART.
'**I've been** to Nero's room.' (3 Laura; 6/1345))

Example 37
[S2+(M)] in an infinitive phrase

1.	NP PRO:DEM	V	**INFP**	
2.	clit		**3INF**	**N**
3.			**ill**	

Täma`kit *tuee* *juomaN* *kahvia.*
Tämä-kin tulee juo-ma-an kahvia.
This-CLIT come drink-3INF-ILL coffee.
'This one's coming **to drink some coffee**, too. ' (7 Taru; 28/1010)

Example 38
[S2+(2M)] in an adposition phrase

1.	NP PRO:DEM	V		**PP**
2.			**N**	**AD**
3.			**gen**	**ill**

He *menee* *hauvav* *vielee.*
Se menee vauva-n viere-en.
It go baby-GEN beside-ILL.
'It goes **next to the baby**.' (11 Jaakko; 18/467)

Unlike the morphological extension of [S2] introduced above, tokens of three-layered elaboration types with only syntactic ([S2+(S2)]) or both syntactic and morphological extensions ([S2+(S2+M)]) are very rare (only nine instances). All structures that are extensions of the [S3]-type are very infrequent as well (altogether thirteen instances). Table 31 shows that in most cases there is only one token of each type in each group. These figures indicate that most of the component types are present in the list because only one child in a group produces them once. Thus, individual children do have an impact on the elaboration type inventory, while the results concerning all the members of a group are limited to only those types that are clearly the most frequent.

8.2.4 Four-layered elaboration

In four-layered elaboration the possibilities for variation are even greater than with three-layered elaboration but at the same time the frequencies are very low, as Table 32 shows. The proportions (%) of the four-layered elaboration types in every group are only marginal. On the group level all that can be said is that in the utterances of children in groups B and C there are more occurrences of four-layered components than in the utterances of group A, but both the proportional differences between the groups as well as the actual number of tokens of each type are very small.

TABLE 32 The frequencies of four-layered elaboration types for each group. The proportions (%) of each component type are based on the total number of elaborated component tokens (T) in the group's data consisting of the two- to four-component utterances.

	Group A		**Group B**		**Group C**		**Total**	
Type	**T**	%	**T**	%	**T**	%	**T**	%
[M3]	0		2	0.2 %	2	0.2 %	4	0.1 %
[S2+(M2)]	8	1.2 %	17	1.3 %	6	0.6 %	31	1.0 %
[S2+(M2+M)]	1	0.1 %	6	0.5 %	9	0.9 %	16	0.5 %
[S2+(2M2)]	1	0.1 %	0		0		1	0.0 %
[S2+(S2+(M))]	4	0.6 %	8	0.6 %	11	1.1 %	23	0.8 %
[S2+(S2+(2M))]	4	0.6 %	3	0.2 %	3	0.3 %	10	0.3 %
[S2+((S2+(M))+M)]	2	0.3 %	7	0.6 %	8	0.8 %	17	0.6 %
[S2+((S2+(2M))+M)]	0		2	0.2 %	5	0.5 %	7	0.2 %
[S2+(S3+(3M))]	1	0.1 %	1	0,1 %	0		2	0.1 %
[S3+(M2+M)]	1	0.1 %	0		0		1	0.0 %
[S3+(M2+2M)]	0		1	0,1 %	0		1	0.0 %
[S3+((S2+(M))+M]	0		1	0,1 %	1	0.1 %	2	0.1 %
[S3+((S2+(2M))+M]	0		0		1	0.1 %	1	0.0 %
4-layered component tokens (total)								
	22	3.3 %	48	3,8 %	46	4.4 %	116	3.9 %
Elaborated component tokens (total)								
	674		1267		1051		2992	

Four-layered elaboration occurs especially in verb structures, noun phrases, and infinitive phrases but there are also some cases in adposition phrases and even adverb phrases. In verb phrases four-layered components mostly express negative or interrogative functions, but these components differ from previously noted components with the same function in that they include modality, necessity or some other aspect requiring the presence an. Examples are shown in 39–41.

Example 39
V = [S2+(S2+(2M))] expressing epistemic modality and negation

1.	NP PRO:DEM			**V**
2.		**AUX:NEG**		**V**
3.			**V**	**V**
4.			**neg**	**1inf**

Se *ei* *taiam* *mattua.*
Se ei taida mahtu-a.
It no seem fit-1INF.
'I **don't think it'll fit**.' (29 Elisa; 13/639)

Example 40
V = [S2+(S2+(2M))] expressing negation of necessity

1.			**V**	NP N
2.	**AUX:NEG**		**V**	partit
3.		**V:NES**	**V**	
4.		**cond**	**1Inf**	

Eikä` *pitäisi* *antaa* *ykkaaketta.*
Eikä pitä-isi anta-a yskänlääket-tä.
No must-COND give-1INF cough medicine-PARTIT.
'(You) s**hould not give** (him/her) any cough medicine.' (7 Taru 25/842)

Example 41
V = [S2+(M2+M)] epistemic modality and interrogative function

1.		NP PRO:DEM	**V**
2.	**V**	partit	**V**
3.	**clit**		**1inf**
4.			**clit**

Voiko *näitä* *ko:ljatakki?*
Voi-ko näi-tä korja-ta-kin?
Can-CLIT these-PARTIT repair-1INF-CLIT?
'**Can** these **be repaired, too**?' (19 Tuomo 22/615)

Modality, necessity, and negation can be applied to verbs in the passive (Examples 42a and b) as well as to verbs consisting of two or more elements, such as phrasal verbs (Examples 43a and b) or verb chains (Example 44).

Examples 42

a) V = [S2+(M2)] negation in passive (colloquial form corresponding to 1PL)

1.		**V**	NP PRO:DEM
2.	**AUX:NEG**	**V**	
3.		**neg**	
4.		**pass**	

Ei	*aukas~ta*	*se.*
Ei	aukais-ta	se.
Not	open-NEG&PASS	it.

'**Let's not open** it. ' (13 Joel; 62/982)

b) V = [S2+(M2)] negation in passive

1.		NP PRO:DEM	NP N	**V**
2.	**AUX:NEG**		adess	**V**
3.				**neg**
4.				**pass**

Ei	*se*	*yöllä*	*tarvita.*
Ei	se	yö-llä	tarvi-ta.
Not	it	night-ADESS	need-NEG&PASS

'It **is not needed** during the night.' (35 Seppo; 14/447)

Example 43

a) V = [S2+(S2+(M))] necessity together with a phrasal verb

1.	NP PRO:DEM		**V**	
2.	gen	**V:NES**	**V**	
3.			**V**	**PTL**
4.			**1Inf**	

Tom	*pitää*	*mennäp*	*pois.*
Tuo-n	pitää	men-nä	pois.
That-GEN	must	go-1INF	away.

'That one **has to go away**.' (1 Tuija; 40/503)

b) V = [S2+(S2+(M))] expressing necessity together with a phrasal verb

1.	NP PRO		**V**	NP PRO:DEM	
2.	gen	**V:NES**	**V**		
3.			**V**		**PTL**
4.			**1Inf**		

Mum	*pitää*	*ot`aa*	*tää*	*pois.*
Minu-n	pitää	otta-a	tämä	pois.
I-GEN	must	take-1INF	this	away.

'I **have to take** this **away**.' (15 Mika; 16/271)

Example 44

V = [S2+(M2+M)] expressing epistemic modality together with a verb chain

1.		NP PRO	**V**	NP N	NP PRO:DEM
2.	**AUX**		**V**	partit	adess
3.	**1s**		**pass&part**		
4.	**clit**				

Sa`kko	*mää*	*leikattua*	*nurmikkoo*	*tällä ?*
Saa-n-ko	minä	leikat-tua	nurmikko-a	tä-llä?
Can-1S-CLIT	I	cut-PASS&PART	lawn-PARTIT	this-ADESS

'Can I **cut** the lawn with this? (19 Tuomo; 1/513)

Four-layered noun phrases typically consist of two to three elements, one of them being a determining or a modifying element (Examples 45a and b). Another typical feature is plurality together with a case marking (Example 46). These noun phrases are generally extended with morphological devices, and the resulting elaboration often includes two suffixes on a single element.

Example 45
a) NP = [S2+(M2)], a noun phrase including a determining pronoun

1.	NP PRO:DEM	V		**NP**
2.		past	**PRO:DEM**	**N**
3.				**pl**
4.				**clit**

Nää:	*laitto*	*nää*	*pöyvätkii .*
Nämä	laitto-i	nämä	pöydä-t-kin.
These	put-PAST	these	table-PL-CLIT.

'These (people) put **these tables, too**.' (31 Aleksi; 7/366)

b) NP = [S2+(M2+M)], a noun phrase including a modifying adjective

1.	ADVP ADV	NP PRO	V		**NP**
2.		all		**ADJ**	**N**
3.				**pl**	**pl**
4.					**clit**

Sittem	*mulle*	*tulee*	*likaset*	*käitti.*
Sitten	minu-lle	tulee	likaise-t	käde-t-kin.
Then	I-ALL	become	dirty-PL	hand-PL-CLIT.

'Then I get **dirty hands**, **too**.' (Laura; 3/1663)

Example 46
NP = [S2+(M2+M)], a noun phrase including a determining pronoun and suffixes expressing plurality and case

1.	NP N		NP		**NP**
2.	all	ADV	N	**PRO:DEM**	**N**
3.			partit	**all**	**pl**
4.					**all**

Pellolle	*vähäl*	*luokaa*	*noillep*	*pelloille.*
Pello-lle,	vähän	ruoka-a	noi-lle	pello-i-lle.
Field-ALL	some	food-PARTIT	those-ALL	field-PL-ALL.

'Onto a field, some food **onto those fields**.' (19 Tuomo; 3/787)

Four-layered infinitive phrases are less common than verb constructions and noun phrases. However, several children in groups B and C produce infinitive phrases which resemble clausal constructions. It has been stated that the infinitive phrase is the most clause-like structure among the phrases (ISK 2004: 488, 836), and in four-layered components this starts to show, for example, in that they include other phrases (Examples 47–48).

Example 47
INFP = [S2+((S2+(2M))+M)], an infinitive phrase including another infinitive phrase

1.	NP PRO:DEM	V	**INFP**		
2.			**3inf**		**INFP**
3.			**ill**	**PRO**	**3inf**
4.				**partit**	**ill**

Nee	*lähhee*	*käytää*	*sinua*	*kattomaan.*
Ne	lähtee	käy-mä-än	sinu-a	katso-ma-an.
They	go	visit-3INF-ILL	you-PARTIT	see-3INF-ILL.

They go **for a visit to see you**. (2 Riikka; 5/1064)

Example 48
INFP = [S2+((S2+(2M))+M)] an infinitive phrase including a noun phrase

1.	NP PRO:DEM		V	**INFP**		
2.		AUX:NEG	V	**V**		**NP**
3.			neg	**1inf**	**PRO:DEM**	**N**
4.					**partit**	**partit**

S+	*ei*	*ossaa*	*ajaat*	*tätä*	*lahkalia.*
Se	ei	osaa	aja-a	tä-tä	traktori-a.
It	not	can&NEG	drive-1INF	this-PARTIT	tractor-PARTIT.

'It doesn't know how to **drive this tractor**.' (19 Tuomo; 4/753)

It has been claimed that children acquire words and functionally coherent phrases as well as create their own constructions with open slots in them. Across these item-based constructions, which also have an important role in adult linguistic competence, there are many kinds of abstractions. (Tomasello 2003: 325–327.) In her study Kauppinen (1998), for example, shows, how a Finnish speaking child acquires the conditional mood through the various functions of the form. Similarly, it can be argued that variation in the functions of negation and asking questions, as well as the need to express modality, necessity and permission expose a child to the morphosyntactic structural variants that can be used to express these functions. Absolute structural complexity thus grows through functional variation and extension.

8.2.5 Five-layered elaboration

The elaborated components produced by the children in this study are mostly two- and three-layered ones. In four-layered components, the numbers of most elaboration types are marginal, each occurring in utterances by only one or two children in a group. Therefore it is not surprising that the existence of five-layered components is even rarer. In fact, there are only four elaboration types representing this category, and none of them is found in utterances by children in group A. In addition, there are only four children who have five-layered components in their utterances, namely Santeri (ID10 / group B), Laura (3/C), Janna (27/B and C), and Riikka (2/B and C). Thus, not even in groups B and C can we talk about a skill of the whole group. The five-layered elaboration types and their frequencies are introduced in Table 33.

All five-layered components represent verb constructions. Typical features of these constructions are the use of the negative auxiliary and the passive form (Example 49), but interrogative utterances are possible as well (Example 50). A clitic particle is often the final morphological element that causes a component to become five-layered.

TABLE 33 The frequencies of the five-layered elaboration types for each group. The proportions (%) of each type are based on the total number of elaborated component tokens (T) in the group's data, which consists of all two- to four-component utterances.

	Group A		**Group B**		**Group C**		**Total**	
Type	**T**	%	**T**	%	**T**	%	**T**	%
[S2+(M3)]	0		0		1	0.10 %	1	0.03 %
[S2+(M3+M)]	0		2	0.16 %	0		2	0.07 %
[S2+(S2+(M2))]	0		1	0.08 %	1		2	0.07 %
[S2+(S2+(M2+M))]	0		2	0.16 %	2	0.19 %	4	0.13 %
5-layered component tokens (total)								
	0		5	0.39 %	4	0.38 %	9	0.30 %
Elaborated component tokens (total)								
	674		1267		1051		2994	

Example 49
V = [S2+(M3)] expressing negation in passive (a colloquial for corresponding to 1PL)

1.			AP	**V**
2.	**AUX:NEG**	ADV	ADJ	**V**
3.			partit	**neg**
4.				**pass**
5.				**clit**

Ei *noin* *isoo* *otetakkaa.*
Ei noin iso-a ote-ta-kaan.
Not so big-PARTIT take-NEG&PASS-CLIT.
'**We won't take** such a big one.' (3 Laura; 33/1315)

Example 50
V = [S2+(M3+M)]; a polite question

1.		NP PRO:DEM	**V**	NP PRO:DEM
2.	**V**		**V**	
3.	**cond**		**1 Inf**	
4.	**2s**			
5.	**clit**			

Poisiks+ sää ajaat tällä ?
Vo-isi-t-ko sinä aja-a tällä?
Can-COND-2S-CLIT you drive-1INF this.
'Would you please drive this one.' (10 Santeri; 4/655)

8.2.6 The evaluation of the complexity of distinct elaboration types

Examining the elaboration types introduced in the previous sections has already made it clear that there are differences in the structural complexity of the components. There is no doubt that, for example, a five-layered [S2+(M3+M)]-component (Example 50) is a more complex structure than a three-layered component [S2+(M)] (Examples 33, 35 and 37). In the former type, both elements of S2 are elaborated and, besides, there are three suffixes in one element, whereas in the three-layered component only one of the S2-elements is elaborated and this is done with a single suffix. However, the situation becomes much more complicated when the distinctions are more fine-grained or if we have to evaluate the structural complexity of components which contain different elaboration devices. We need to decide which is more complex, a morphological device or a syntactic one. Is it more complex to elaborate morphologically the already existing elements in a component than to add new syntactic elements? Do the different devices compensate each other, and if they do, to what extent do they do so? In posing these questions we are facing the problem of comparability (Miestamo 2006), that is, how to value the morphological and syntactic elements involved in a morphosyntactic construction.

To avoid speculations on whether morphological devices are more complex than syntactic ones or vice versa, another kind of description must be applied instead of a strict ordering of components in terms of structural complexity. A suitable device for this purpose is a net-like construction where the relations of structures to their nearest variants are more important than placing all the elaboration types on a line according to complexity. In the net-like description model the components are organized according to their derivational relations into structural families. Each family is thus a subnet of structures with derivational paths between them, and distinct families will not have to be compared. In this context derivation means adding something to the original structure to create a new structure. This makes a derived structure a more complex one than the original one. In my opinion, a net-like model of complexity relations between structurally different components corresponds better with the actual properties of language than a linear order ranging from a simple structure to a complex one. It shows that an absolute comparison is not relevant or even possible across structural families and yet it allows absolute comparison within each of the families. At the same time it shows multidimensional relations between the structures within a family.

At this point one should be reminded that a net which branches into structural families, paths between structures and derivation must be taken metaphorically. It does not constitute an image of how a child constructs her language. A child is not moving along paths from structure to structure. A net is not a psycholinguistic description of acquisition process. In exactly the same way as structures are not derived from each other, they are not each others' derivatives. Talking about derivation in this context only emphasizes the large similarities and small differences between the components that are near each other in a structural sense. Thus the net, paths, and derivation are all used only to describe structures, not to describe the processes in acquiring them or potential processes in creating different structures.

A starting point for describing the net of structural families is in the three distinctive types of two-layered components [M], [S2], and [S3]. They represent structures that are created with minimum help from elaboration (see Section 8.2.2). All the elaboration types found on the third, fourth and finally on the fifth layer are further elaborations of these three basic types. Each structural family originating from them will now be discussed separately.

The structural family originating from two-layered [M]-elaboration is the most simply constructed and the smallest of the three subnets. This is because further elaboration is possible only with another morphological device which is an additional suffix of some kind. The [M]-family has only one branch and it proceeds straightforwardly from [M] via [M2] to [M3] (Figure 48). Components representing elaboration types of [M]-family are always single word-forms, and they differ from each other only in number of suffixes (1, 2 or 3). The 2-layered [M]-elaboration, which is the starting point of the whole [M]-family of elaboration, is by far the most frequently used elaboration type in the whole data (see Table 30), while the other members of the family are rather rare (see Table 31 for frequency and proportion of [M2], and Table 32 for [M3]).

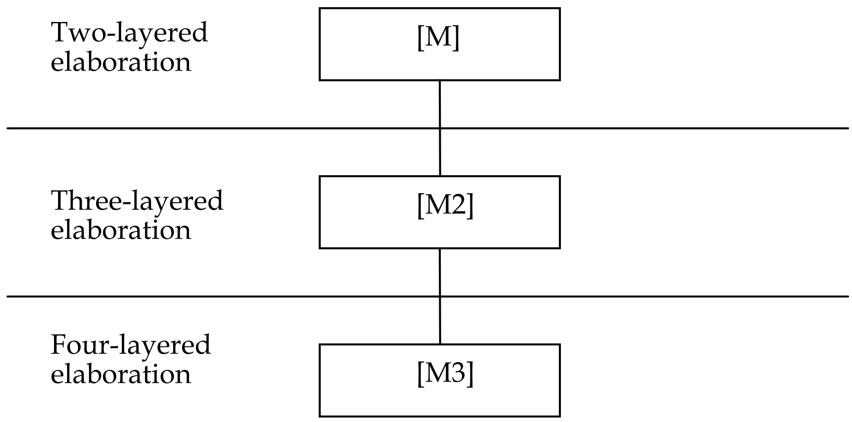


FIGURE 48 The members of the [M]-family. Each elaboration is illustrated with a connecting line between the structures. The lower a structure is in the net the more complex it is taken to be in terms of morphosyntactic structure.

The structural family starting from two-layered [S2]-elaboration has multiple possibilities in further elaborations when compared to variation possibilities in the [M]-family. In a component representing a two-layered [S2] elaboration type there are two elements (words) and both of them can be extended in several ways (see Figure 49). This obliges the [S2]-family to branch off into two sub-branches from the very beginning, one with only morphological elaboration (the first member of this branch is [S2+(M)]) and the other with morphological and syntactic elaboration (starting from [S2+(S2)]). These branches divide into further sub-branches according to the number of elaborated elements and the devices used in elaboration. The [S2]-family grows into a multidimensional network where there are connections between component types in several directions. However, there are no connecting lines between the sub-branches beginning from [S2+(M)] and [S2+(S2)], although they do have common elements. For example, [S2+(M2)] is a part of [S2+(S2+(M2))] as well, but these two elaboration types are not connected to each other. The reason for this is that the obvious linkage is different from the connections plotted in Figure 49. The lines refer to connections between structures *x* and *y*, where *x* is the one that is elaborated further and *y* is the result of the elaboration. In the case of [S2+(M2)] and [S2+(S2+(M2))] and other similarly connected pairs of structures, the relationship is not derivative. Rather there is a structure *x* that might be interpreted as a means to achieve the structure *z*. Thus, it is a question of a means and result and not of a starting point and result as in the relations illustrated in Figure 49.

There are three separate branches originating from [S2], distinguished according to how a basic, two2-layered [S2]-component consisting of two words is elaborated further. The first branch on the left (Figure 49) is for those extensions where [S2] is elaborated only morphologically to attain a third, fourth and even fifth layer. This branch is split into two sub-branches, one for morphological elaboration of a single word only and the other for morphological elaboration of two words.

The branch on the right consists of component structures which start elaboration of [S2] with another [S2]. Like the branch on the left, the branch on the right is split in two according to the number of elements elaborated in the basic [S2]-structure, but there are several connections between these sub-branches as well.

The third line branching off from the two-layered [S2] is the one leading straight to the fourth layer and including only the [S2+(S3+(3M)] elaboration type. The reason for separating it as an independent branch is that in this structure elaboration of the original [S2] starts with [S3] but neither with [M], as on the left, nor with [S2] as on the right. Figure 49 shows only those elaboration types that were found in the data, but not every possible structure from the simplest to the most complex. This is why the potential intermediate forms between the [S2] and [S2+(S3+(3M))] are not included here.

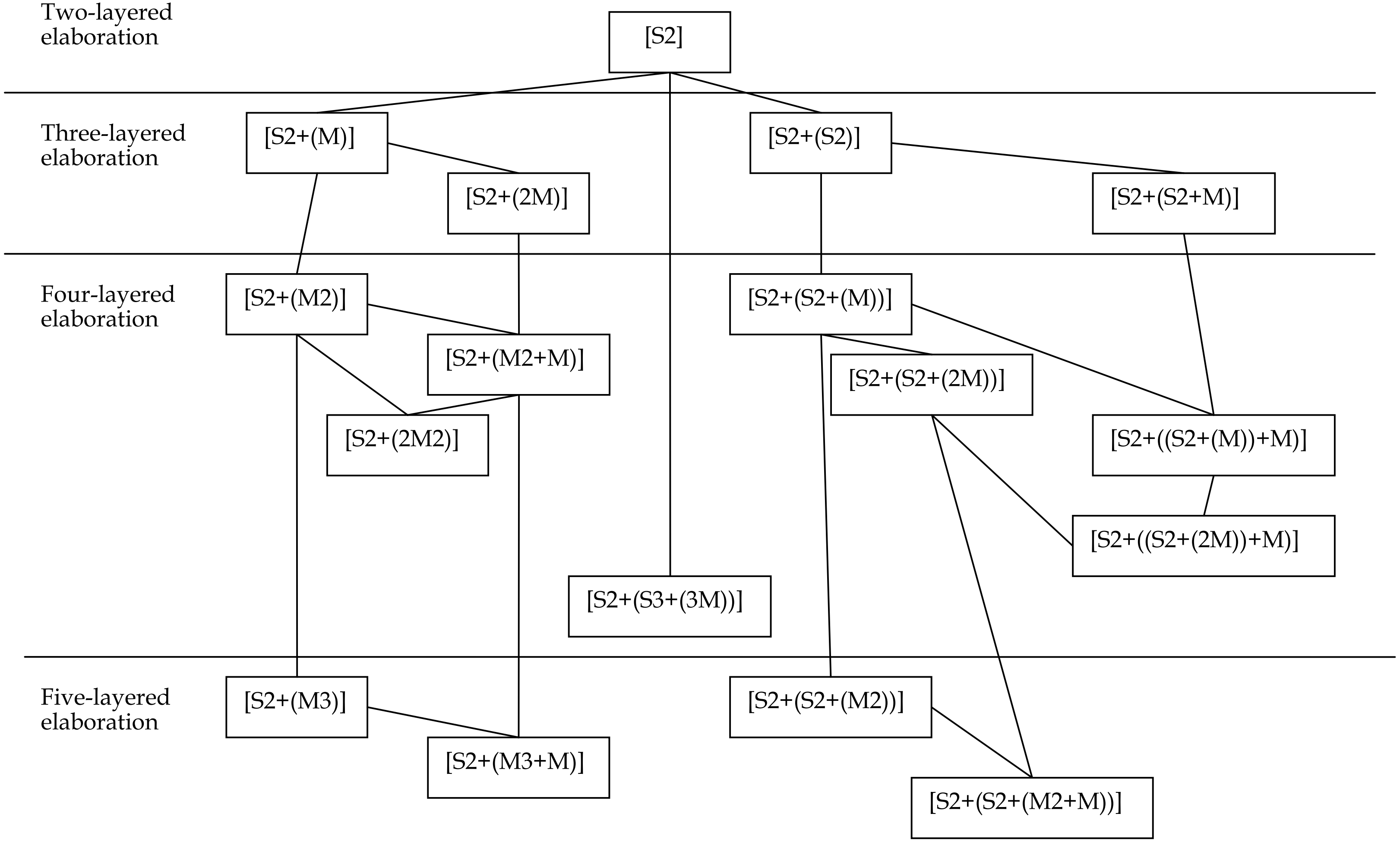


FIGURE 49 The elaboration types originating from the two-layered [S2]-type arranged on layers and connected with lines referring to a derivational relationship. Comparison in terms of structural complexity can be made within each branch, and differences in level show differences in complexity. Elaboration types placed lower in the picture are evaluated as more complex structures than those higher up

The structural branches starting from the two-layered [S2] show how much complexity may vary within a layer. Placing one elaboration type lower than another within a layer suggests that it is more complex than one placed higher, although the distance between the structures is only suggestive. Within a branch on a given layer an influential factor is the number of elaborations occurring on that layer. Thus for example [S2+(2M2)] is more complex than [S2+(M2+M)], because it has two elaborations which both reach the fourth level whereas in the latter structure there is only one elaboration on the fourth level while the other one reaches only the third level (Figure 50). Such a comparison is suitable within a branch but a comparison between the branches would rather quickly encounter problems of comparability.

Figure 49 showing the [S2]-branches also sheds light on another aspect of the role of layers. Until now the number of layers has been used as an estimate of structural complexity. The elaborations attaining different layers have been categorized into different complexity levels as well and addition to layers has implied a more complex structure.

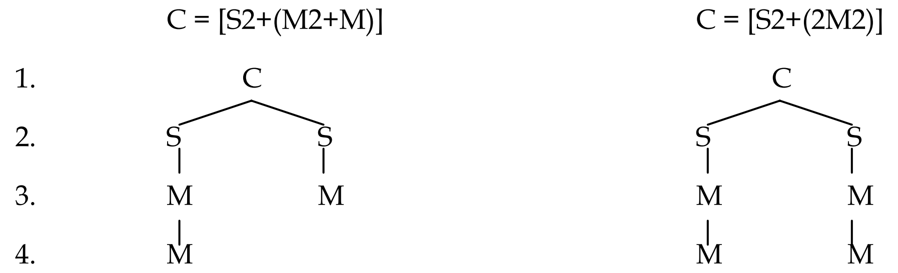


FIGURE 50 Comparison of [S2+(M2+M)] and [S2+(2M2)]. Both component structures have four layers but the former reaches the fourth layer with only one of its elements whereas the other reaches the layer with both its elements and thus the latter is considered to be a more complex structure. C = component; S = result of syntactic elaboration; M = result of morphological elaboration; 1.–4. = layers.

However, examining elaboration types such as [S2+(S3+(3M))] on the fourth layer and [S2+(M3)] on the fifth layer gives an opposite impression. The former type has three elaborated elements which all reach the fourth level and this is done by extending the basic two-layered [S2]-structure syntactically (S3) and morphologically (3M). In the latter case there is only one element which is elaborated and there is only one elaboration device: morphological extension (M3). This makes the role of layers rather ambiguous and the estimation of complexity based on layered structure becomes even rougher than was expected.

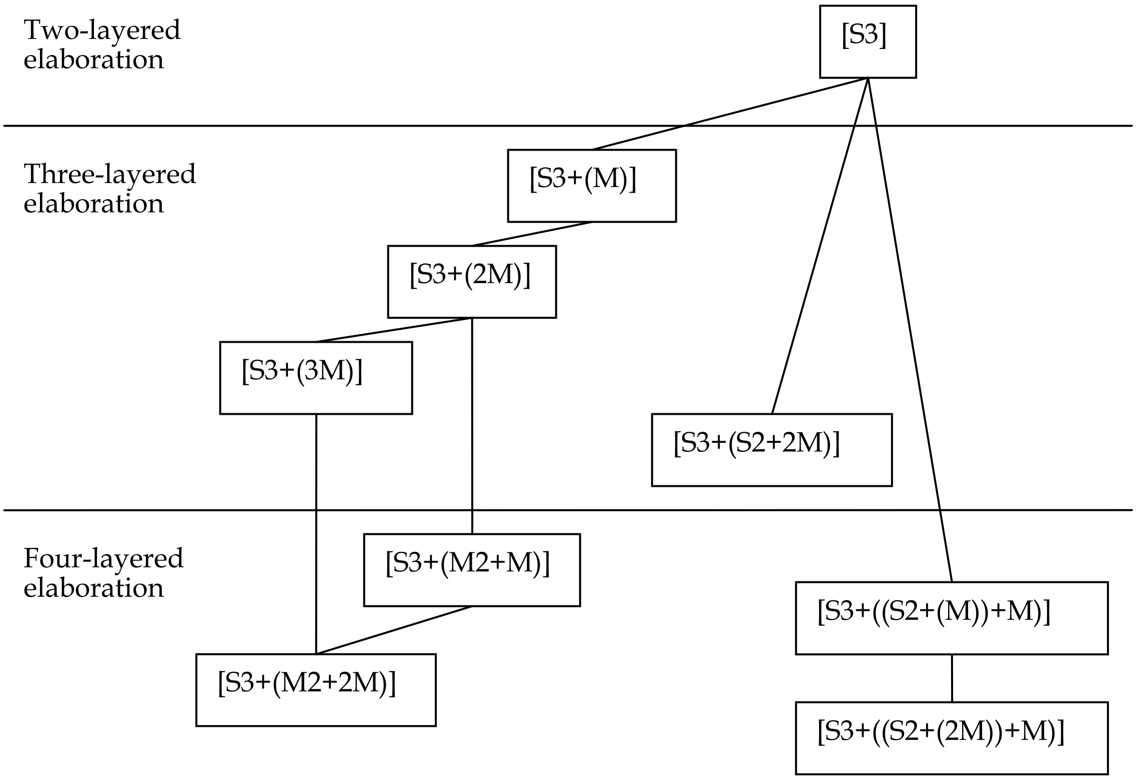


FIGURE 51 The elaboration types originating from a two-layered [S3]-component arranged on layers and connected with lines referring to a derivational relationship.

Besides the [M]-family and the [S2]-family, there is a third family of elaboration types to be laid out and evaluated in terms of structural complexity. The [S3]-family is illustrated in Figure 51. The basic two-layered [S3]-elaboration expands in two ways, either with morphological devices alone (the left branch) or with both syntactic and morphological devices (the branches on the right). Actually the two right hand branches do have a shared origin on the third layer. They are both extensions of an unattested elaboration type such as [S3+(S2+M)], but since this type did not occur in the children's utterances both branches start directly from the [S3]-elaboration.

Although there are many fewer elaboration types originating from [S3] than from [S2] the complexity variation within this family is obvious. In the [S2]-family, most of the component types were elaborated both syntactically and morphologically, but in the [S3]-family, morphological devices alone seem to be the main means of elaboration. It must be kept in mind that between the branches and even between the layers the distances in the figure are not meant to be comparable. The illustration of complexity differences with distances and placement of one structure either higher or lower than another is only a suggestive description of complexity relations.

8.3 Complexity of component tokens in groups A, B and C

In earlier chapters we have learned among other things that the children in groups A, B and C use mostly utterances consisting of two, three and four components (Chapter 6) and that there seem to be some trade-offs between the number of components in utterances and the layered structure of the components: in two -component utterances the components tend to reach deeper layers than in utterances consisting of three or four components (Chapter 7). In this chapter so far the focus has been on the elaboration types that are used to create component tokens in utterances. The investigation ws first of all into what kind of two-, three-, four- and five-layered elaboration types there are in children's utterances and how much each of them is used (Sections 8.2–8.5). Therreafter, it was suggested how these elaboration types can be classified into [M]-, [S2]-, and [S3]-families and how each family can be arranged in a net to describe the complexity relations between the members. Now it is time to put these pieces of knowledge together and make comparisons between groups A, B and C in terms of how complex the elaboration types of [M]-, [S2]-, and [S3]-families they use to construct their component tokens in two-, three- and four-component utterances. The distribution of tokens of each elaboration type in two-, three- and four-component utterances is compiled in table form in Appendix 4. A comparison of groups A, B and C will be made separately for each distinct elaboration family, starting with the [M]-family (see Figure 48).

In the [M]-family there are only three distinct elaboration types ([M], [M2] and [M3]) and each of them reaches a different layer. The [M]-type[15] was found to be the most frequent elaboration type in every group, with proportions ranging from 56% to 69% (Table 30). The other members of the [M]-family occur much less frequently, although [M2]-elaboration has spread to two-, three- and four-component utterances in all groups. Group differences are present only in case of [M3]-elaboration which occurs only in utterances by children in groups B and C. Thus, the inventory of elaboration types in the [M]-family seems to show that groups B and C have progressed a step further than group A in structural complexity, as illustrated in Figure 52. However, it must be remembered that there are only two tokens representing [M3]-type in groups B and C. Therefore, this is not yet convincing evidence of more complex component structure in groups B and C.

15 It is important to keep [M]-family and [M]-type separated. The [M]-family refers to the group of three elaboration types which are purely morphological ([M], [M2] and [M3]). The [M]-type is the basic elaboration in the [M]-family, and it describes components that consist of a single word form including only one suffix. The case is the same with the other elaboration families too, because each family is named according to its basic elaboration type.

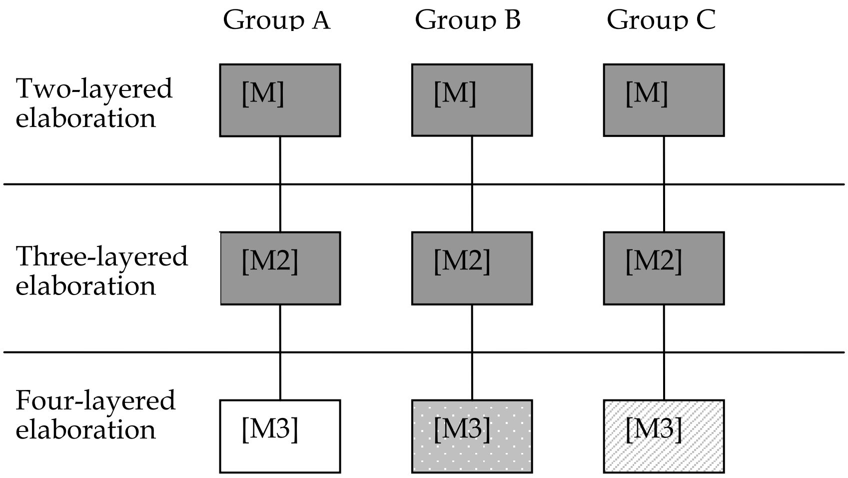


FIGURE 52 The occurrences of components representing the [M]-family elaboration types in utterance categories of two, three or four components. The extent of use is signalled with different shades of gray. White = no occurrences of the component type; diagonal lines in pale gray = tokens found in only one utterance category (either in two-, three-, or four-component utterances); white dots on medium gray = tokens found in two utterance categories (in two- and three-component, three- and four-component or two- and four-component utterances); dark grey = tokens found in all utterance categories.

A similar trend can be seen in the [S2]-branches, and this time in a more obvious fashion. In Figures 53–55 all components in the [S2]-family are presented group by group. Again, wideness of use of each component type is taken into account, and marked with different shades of gray in the figures.

What is common in all three groups is the wide use of the elaboration types [S2], [S2+(M)] and [S2+(2M)]. All these two- and three-layered elaboration types occur in two-, three3- and four-component utterances in every group. Together with tokens of the [M]-family, they form the core resources, and account for as much as 91.1% (614 tokens) of all elaborated components in group A, 90.4% (1146 tokens) in group B, and 88.9% (934 tokens) in group C. Despite the differences between the groups evidenced by MLU and IPSyn, the structural similarities in components are remarkable since the differences are concentrated in those elaboration types with only marginal occurrence rates.

A comparison between the groups shows that the inventory of elaboration types and the breadth of use of each type obviously increase between groups A and C, but groups B and C are very similar. If we look at what elaboration types are used, then all groups are similar except in respect of the fifth layer alone.

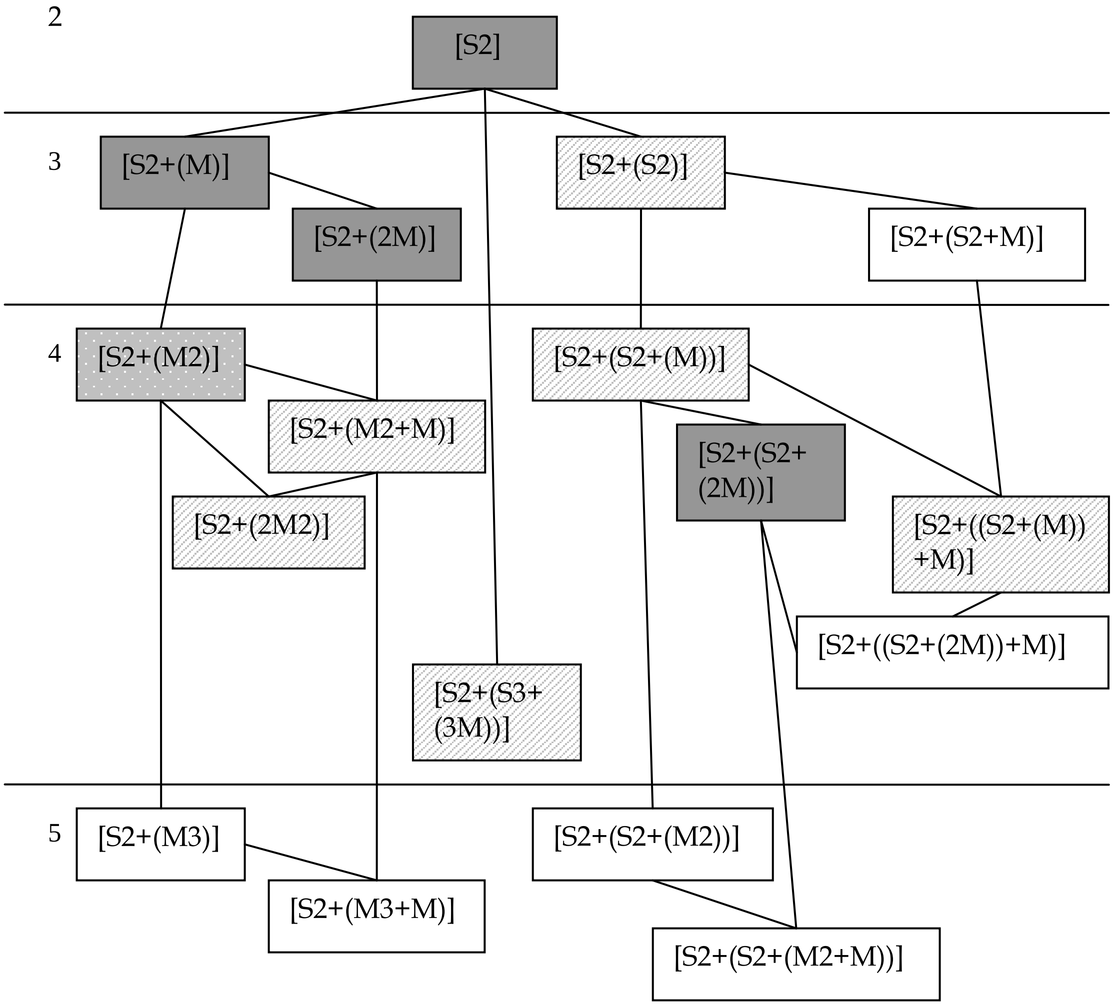


FIGURE 53 Occurrences of the [S2]-family elaboration types in the utterances of group A. White = no occurrences of the elaboration type; diagonal lines in pale gray = tokens in one utterance category (either in two-, three-, or four-component utterances); white dots on medium gray = tokens in two utterance categories (in two- and three- component, three- and four-component or two- and four-component utterances); dark grey = tokens in all three utterance categories.

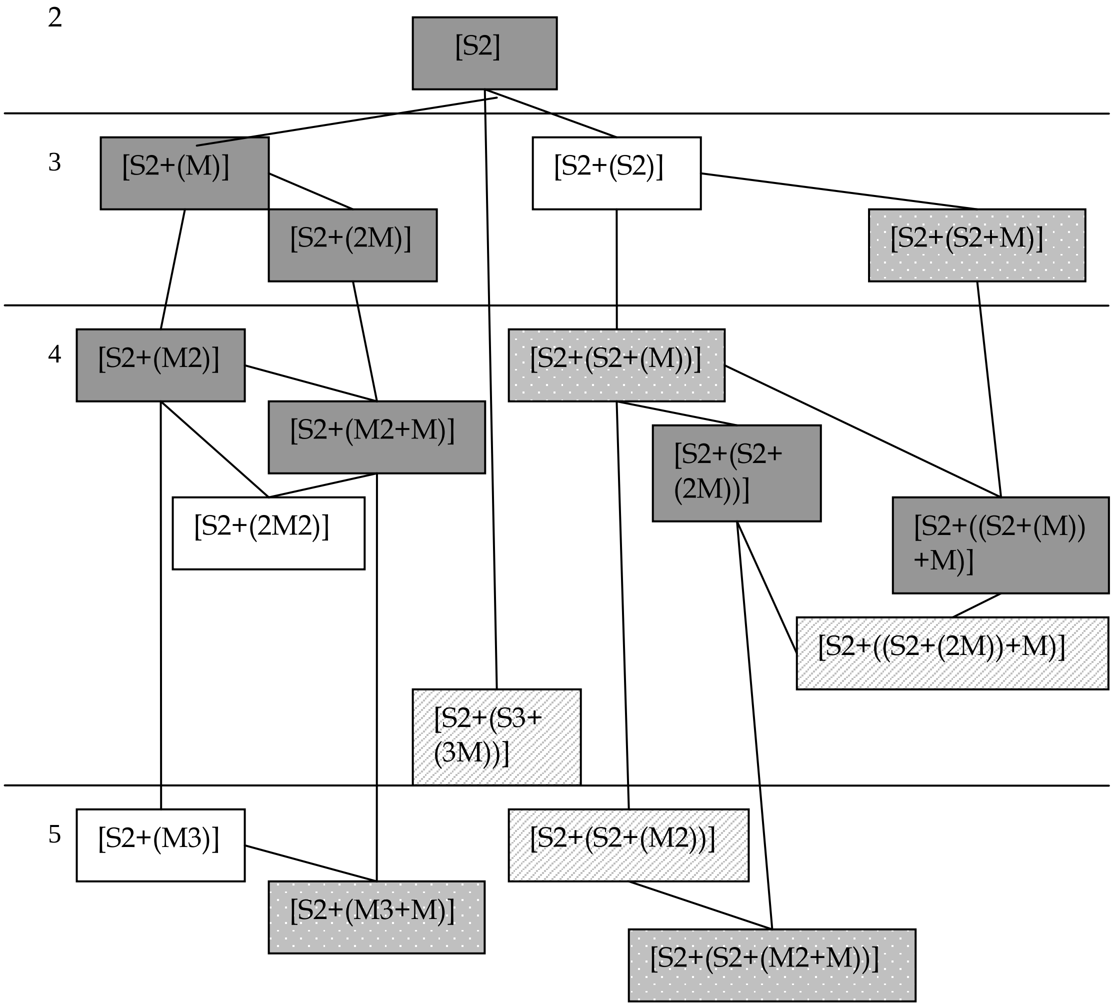


FIGURE 54 Occurrences of the [S2]-family elaboration types in the utterances of group B. White = no occurrences of the elaboration type; diagonal lines in pale gray = tokens in one utterance category (either in two-, three-, or four-component utterances); white dots on medium gray = tokens in two utterance categories (in two- and three- component, three- and four-component or two- and four-component utterances); dark grey = tokens in all three utterance categories.

But as soon as the breadth of the use of the elaboration types is considered groups B and C show a more firmly established inventory of types on the fourth layer. Another interesting difference between group A and the other two groups is that children in the former group focus almost entirely on morphological elaboration. There are very few occurrences of the elaboration types originating from [S2+(S2)], and in three cases out of four – [S2+(S2)], [S2+(S2+(M))], and [S2+((S2+(M))+M)] – tokens of them exist in one utterance category only. In Figures 54 and 55 illustrating the inventories of elaboration types used by groups B and C respectively, the left- and right-hand branches are equally covered with dark grey, indicating that component elaborations originating from [S2+(S2)] are used in all the utterance categories as well. Thus, the complexity of elaborations in the inventory grows both downwards and horizontally, towards more deeply elaborated elements as well towards more elements to be elaborated.

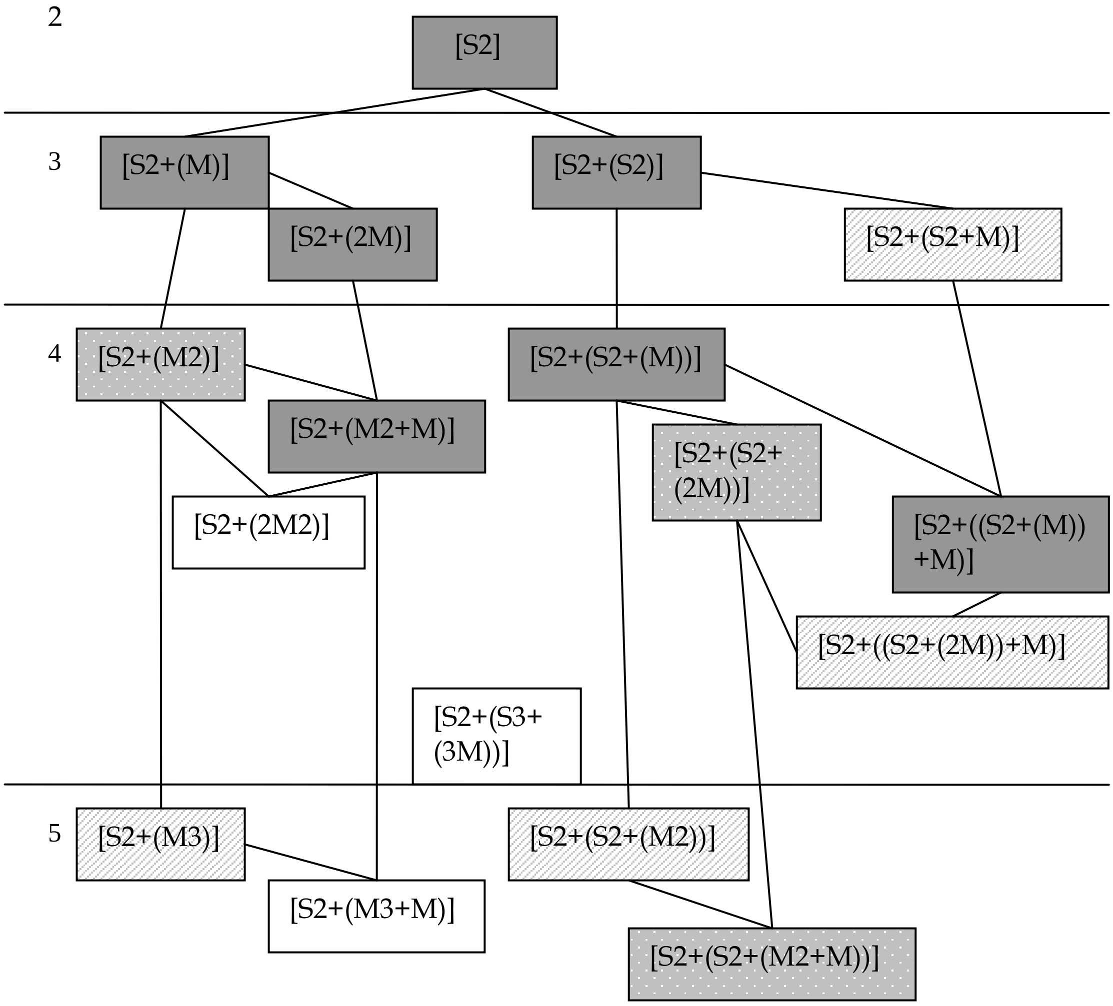


FIGURE 55 Occurrences of the [S2]-family elaboration types in the utterances of group C. White = no occurrences of the elaboration type; diagonal lines in pale gray = tokens in one utterance category (either in two-, three-, or four-component utterances); white dots on medium gray = tokens in two utterance categories (in two- and three- component, three- and four-component or two- and four-component utterances); dark grey = tokens in all three utterance categories.

The children in groups B and C have proceeded in both directions, whereas those in group A are only starting to discover the right-hand branches of the [S2]-family.

The third family of elaboration types, the [S3]-family, is the smallest of all in number of tokens. The different [S3]-based elaborations only account for approximately 1% of all elaborated components, and the number of tokens varies from six in group A, to eight in group C, and twelve in group B. Figures 56–58 illustrate the distribution of the [S3]-based elaboration types and the breadth of their use by groups A–C.

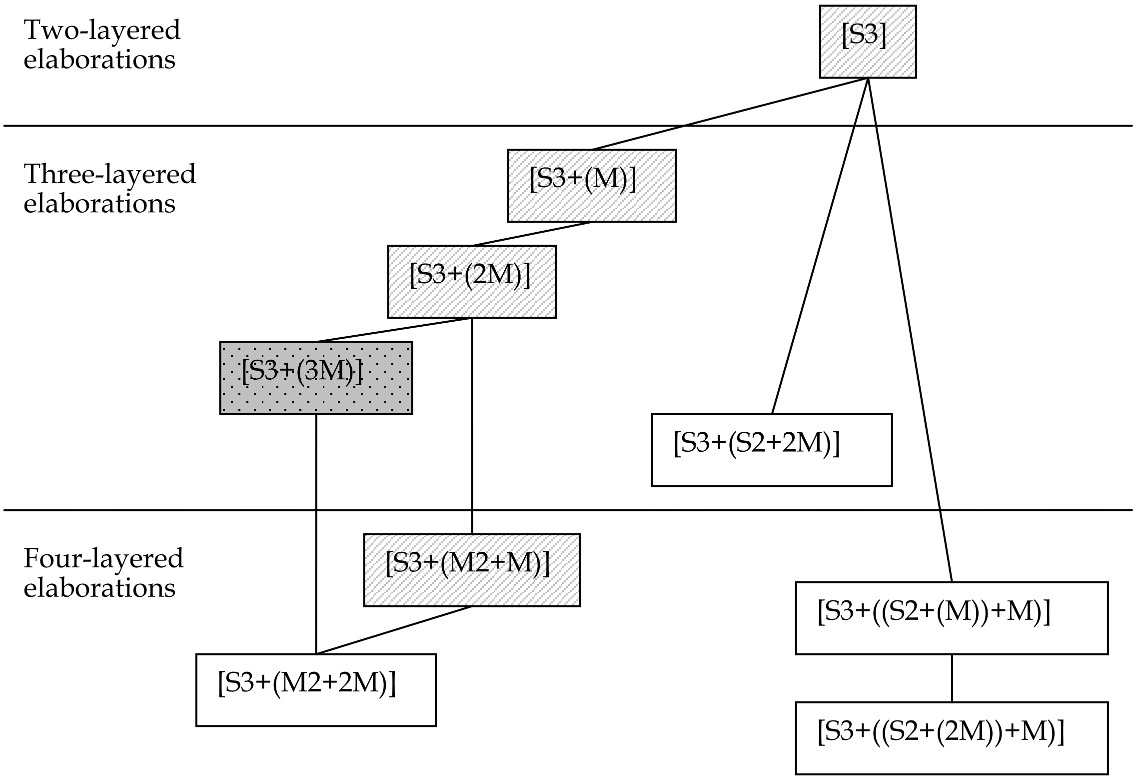


FIGURE 56 The [S3]-based elaboration types and their use in two-, three and four-component utterances by the children in group A.

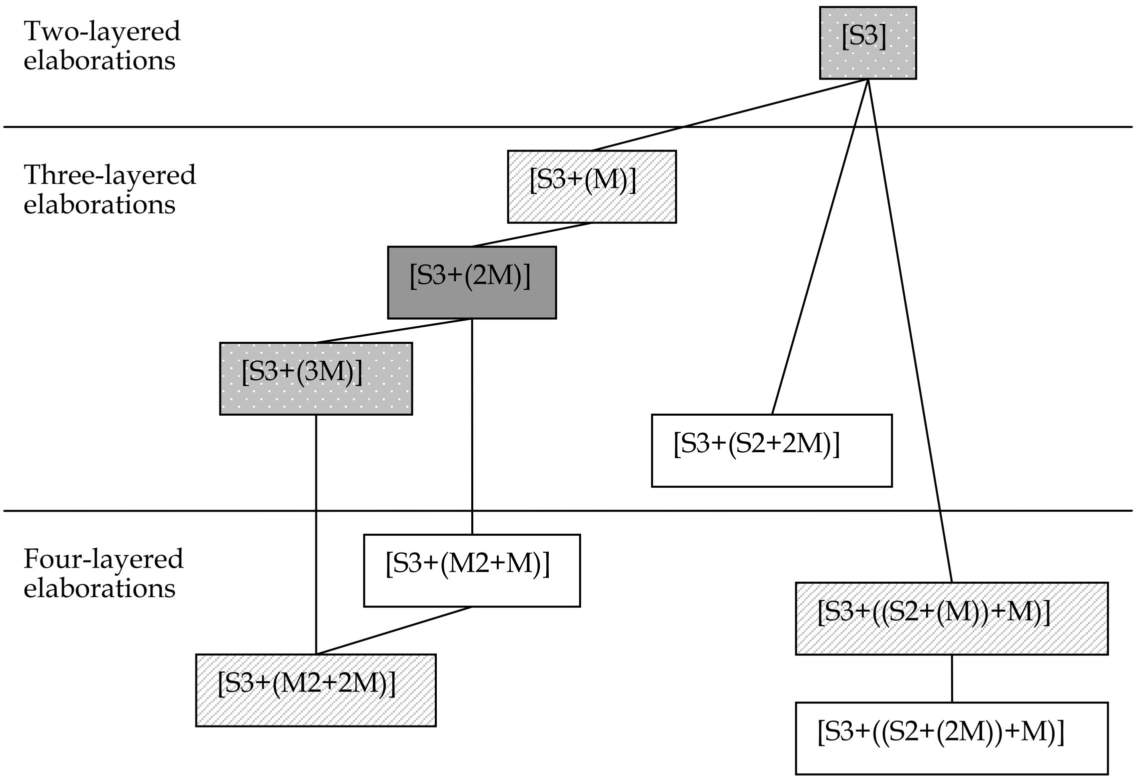


FIGURE 57 The [S3]-based elaboration types and their use in two-, three and four - component utterances by the children in group B.

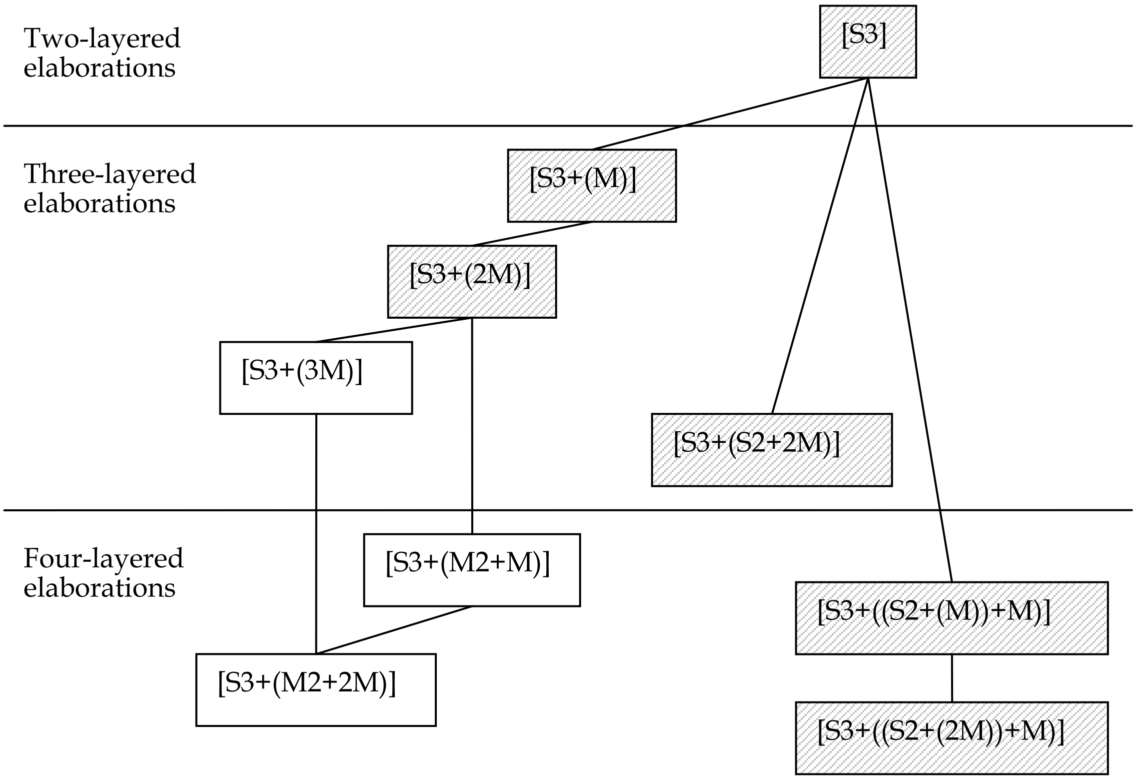


FIGURE 58 The [S3]-based elaboration types and their use in two-, three and four -component utterances by the children in group C.

This time, it is groups A and C that look similar, in that both of them have [S3]-based elaborations in only one utterance category. What differentiates these two is the quality of the [S3]-components. In group A, there are occurrences of such components where [S3] is elaborated only with morphological means. In group C, despite the narrowness of the use and infrequent occurrences, syntactic means of elaboration are well represented alongside morphological ones. It appears, from looking at the results for the [S2]- and [S3]-families, that growth of structural complexity in elaboration begins primarily with the use of morphological devices and only after that are syntactic devices added as well. In fact, group B gives further support to this idea with its use of [S3]-components. Qualitatively, the group B resembles group A with almost exclusively morphological elaboration of [S3], confirming the notion that morphological devices are the ones that children start with in elaborating the components. Although the number of [S3]-based elaboration tokens is very small in all three groups, this area is, however, the scene for innovations, and this makes these tokens significant despite their rarity.

To sum up, the group differences in the structural complexity of elaborated components are mostly concentrated in the marginal areas of the inventory of elaboration types. Typically, the differences are not a question of the existence or non-existence of an elaboration type, but a matter of breadth in use. The same elaboration type that occurs only in two-component utterances in group A may occur in all three utterance categories in groups B and C and thus be both more firmly established and more frequently used. There are three exceptions to these typical gradational differences between the groups, all concerning group A. These are the absence of [M3]-elaboration and five-layered elaborations originating from [S2], and the absence of the [S3]-based types with morphosyntactic elaboration. These features that are lacking in the performance of the children in group A are located in the peripheral regions of elaboration found in this study, and they represent two different dimensions. [M3]-elaboration and the five-layered types are the most deeply elaborated parts of their branches, and thus they represent development (only or mostly) in a vertical direction. Compared to them, [S3] and its further elaborations already represent the starting point for expansion in a horizontal direction. A component representing [S3]-elaboration consists of three words, ready to be elaborated vertically. Progress in structural complexity moves in both directions, resembling the circles in water, and moving away from the centre. It can be argued that group A is one circle behind the other groups, and the distance from the others is shown in the reklative wekness in the establishment of the components in general as well as in the lack of the more extreme elaboration types in the branches.

These conclusions concerning the inter-layer variation in structural complexity direct the study back to the part where the layered structure of whole utterances was investigated. When only the most deeply elaborated component from each utterance was considered (Section 7.2), in four-component utterances the groups showed very similar performance (see Table 17 and Figure 35). The same tendency for similarity between the groups was evident also when the focus was on how deeply elaborated components are combined in four-component utterances (Figure 41). Although these layer analyses suggest that all three groups produce utterances with similar structural complexity, the investigation of elaboration types gives a different picture. It was expected that groups B and C would have more complex components than group A.

In Table 34 the inventory of elaboration types found in four-component utterances for each group is presented. The table shows that the groups do differ in several ways although the previous layer analyses in Chapter 7 emphasized similarity between the groups in four-component utterances. First, groups B and C have more variation than group A in the elaboration types on each layer. Thus, children in these groups have more resources for producing utterances. Secondly, groups B and C have more elaboration types which use syntactic elaboration devices than group A. They have extended their structural resources to the [S3]-family on all layers. However, in the case of the [M]- and [S2]-families all groups seem to have similar inventories with only minor differences.

In a way, the four-component utterances represent a watershed of speech production in the data. They belong to the category of the three most frequently produced utterance types that altogether account for approximately 80% of all productions in this study. Utterances with more components are rather marginal in all groups.

TABLE 34 The elaboration types found in the four-component utterances of each group. The types are categorized according to the group and their layered structure. The elaboration types that are common to all three groups are highlighted in grey.

	Group A	Group B	Group C
Two-layered elaborations	[M]	[M]	[M]
	[S2]	[S2]	[S2]
		[S3]	[S3]
Three-layered elaborations	[M2]	[M2]	[M2]
	[S2+(M)]	[S2+(M)]	[S2+(M)]
	[S2+(2M)]	[S2+(2M)]	[S2+(2M)]
			[S2+(S2)]
		[S3+(2M)]	
		[S3+(3M)]	
Four-layered elaborations			[S2+(M2)]
	[S2+(M2+M)]	[S2+(M2+M)]	[S2+(M2+M)]
		[S2+(S2+(M))]	[S2+(S2+(M))]
	[S2+(S2+(2M))]	[S2+(S2+(2M))]	
		[S2+((S2+(M))+M)]	[S2+((S2+(M))+M)]
	[S3+(M2+M)]	[S3+(M2+M)]	
		[S3+((S2+(M))+M)]	[S3+((S2+(M))+M)]
Five-layered elaborations		[S2+(M3+M)]	

Thus a four-component utterance is an “upper limit “of ordinary conversational talk in this data. Layer analysis suggests that the four-component utterances are also the point where the general group differences vanish, but a more detailed analysis of elaboration types nevertheless reveals some distinctive features in the inner structures. There are reasons to believe that these differences reflect the routes that children use to develop structural complexity. This gives support to the idea that structural complexity is a multidimensional phenomenon and that progress towards more complex constructions takes place in several directions and on variouslevels at the same time.

8.4 Individual variation within groups A, B and C

8.4.1 Assumptions about the relations between IPSyn, MLU and Utterance Analysis

So far the groups that were originally separated on the basis of contradictory results from MLU and IPSyn have been analysed without paying attention to the individual variation within them. This was done to gain a general view of the factors that might prove to be distinctive in accounting for the different levels on the MLU and IPSyn scales. In other words, the aim was to find out how the MLU and IPSyn scores relate to the view of structural complexity adopted in the present study. Both similarities and differences have been found between the groups, but the reasons for the contradictory results from MLU and IPSyn have not been touched upon yet. Therefore, it is finally time to look at individual performance within the groups in order to get more information about what properties of language production MLU and IPSyn reflect.

Mean Length of Utterance takes account of each and every morpheme. This suggests that MLU should reflect growth in structural complexity as well, since complexity grows through elaboration of components with either suffixes or new independent elements. Moreover, MLU is calculated on the basis of whole utterances, and therefore it should somehow reflect the more complex combinations of components as well.

Being sensitive to each morpheme is, however, also a weak point of MLU. It does not differentiate the qualities of morphemes; thus, it puts a single-component utterance with four morphemes (e.g. [S2+(2M)]) into same category with an utterance consisting of four unelaborated components (=morphemes). The number of morpheme-rich utterances plays a role in the results as well because MLU is a mathematical average, and a large number of utterances with many morphemes will raise the average. Therefore, MLU responds to talkativeness as well as to repetition of the same structures consisting of several morphemes. Theoretically, it is possible that the same MLU value can be reached either with structurally sophisticated utterances or with talkative and repetitive performances. However, in practice the proportions in which these factors contribute to the results may vary substantially and they form a continuum sliding from purely structural properties to more pragmatic factors in the performance. MLU does not differentiate whether morphemes are of structural or pragmatic origin, and the contradictions between MLU values and IPSyn scores are due at least partly to this. Since Utterance Analysis focuses only on structural properties, a comparison between MLU and Utterance Analysis will give information on how much MLU actually reflects the structural properties of utterances.

The Index of Productive Syntax searches for specific structural properties in utterances, and thus connections to Utterance Analysis and morphosyntactic complexity seem to be more straightforward than in the case of MLU. IPSyn specifies each structure it is responding to and accepts maximally two occurrences of each structure. Thus, talkativeness and repetition will not be given credit. IPSyn pays attention to whole utterances and other multiword structures, particularly in the sentence structure section, while other structural properties (e.g. the use of suffixes) recieve attention in the nominal phrase and verb construction sections. However, these factors make the connection with Utterance Analysis somewhat fuzzy. IPSyn does not actually look at both the entirety and the details of the utterances at the same time, but separates them strictly into different parts of the analysis. There is, for example, a separate entry for a verb with a negative auxiliary (*ei mene* 'does not go') and another entry for the perfect tense (*on mennyt* 'has gone') as well as for first person (*menen* 'I go'), but there is nothing that tells whether these properties are actually produced in the same phrase (*en ole mennyt* 'I have not gone') or not. In other words, morphological properties are separated from syntactic ones, and the structure of an entire phrase or sentence is not considered in an integrated way. Separation of suffixes from the syntactic context causes the IPSyn to be better at making an inventory of structural details than in evaluating their use. Describing IPSyn as a morphosyntactic measuring g scale must be interpreted as depiction of a tool with separate morphological and syntactic sides. Due to the fragmentation of structures in IPSyn, it is difficult to assess whether Utterance Analysis and IPSyn reflect each other very well.

In the following sections the results from Utterance Analysis will be compared with the IPSyn scores and MLU values. Do children with a higher score on either of those scales produce structurally more complex components or component combinations? Are there any clear connections between the measuring scales and component structures?

8.4.2 The Index of Productive Syntax versus the structural complexity of components

8.4.2.1 Group A

In group A there are nine children, all within the relatively narrow MLU range of 3.688–4.388. At the same time their IPSyn scores vary from 42 to 68, implying that these children's productions are structurally quite different. When the IPSyn scores and the results for the component types are combined in Table 35, two things become obvious that are worth mentioning. First, Aleksi (31) and Leo (35) stand out from the rest, and second, their inventory of elaboration types does not match that of the rest of the children.

What Aleksi and Leo have in common is their performance in four-component utterances. They both use six different elaboration types in constructing components that occur in their four-component utterances, among them some four-layered types ([S2+(M2+M)], [S2+(S2+(2M)], and [S3+(M2+M)]) while the other children in the group have only one to three elaboration types in their four-component utterances and none of them is four-layered. The other children use their four-layered elaboration types almost exclusively in two-component utterances. Based on this, it seems that Aleksi and Leo, with the highest IPSyn scores of 66 and 68 points, do have structurally more complex components in their utterances than the other children in group A.

TABLE 35 The inventory of elaboration types and their use in children's two-, three- and four-component utterances by group A. The numerals 2, 3, and 4 refer to occurrences of each elaboration type in two-, three- and four-component utterances repsectively.

ID	**22**	**6**	**18**	**17**	**36**	**29**	**35**	**31**	**34**
Name	**Anna**	**Ronja**	**Saku**	**Tatu**	**Sampo**	**Elisa**	**Seppo**	**Aleksi**	**Leo**
IPSyn	**42**	**51**	**53**	**56**	**57**	**57**	**60**	**66**	**68**
MLU	3,888	4,1	3,962	4,063	3,987	4,388	3,688	4,2	4,388
Two-layered components									
[M]	234	234	234	234	234	234	234	234	234
[S2]	234	234	23	23		23	3	234	234
[S3]								2	
Three-layered components									
[M2]					23	234	23	34	34
[S2+(M)]	234	23	234	234	23	234	23	234	234
[S2+(2M)]	2	2	23		23	2	2	23	34
[S2+(S2)]		2							
[S3+(M)]		2							
[S3+(2M)]	3								
[S3+(3M)]			2						
Four-layered components									
[S2+(M2)]							23	23	
[S2+(M2+M)]									4
[S2+(2M2)]					3				
[S2+(S2+(M))]		2	2			2			2
[S2+(S2+(2M))]	2					2	3	4	
[S2+((S2+(M))+M)]				2			2		
[S2+(S3+(3M))]								3	
[S3+(M2+M)]								4	

Aleksi's and Leo's component inventory would seem to belong ar one end of a continuum where IPSyn scores reflect the quality of component elaboration, but as was mentioned earlier, the inventory of elaboration types does not sort the rest of the children into an order which reflects the IPSyn results. The same mismatch between IPSyn and the structural complexity of components continues in the case of component combinations (Appendix 5) except in four-component utterances. The same elaboration types that separated Aleksi and Leo from the other children in group A make their four-component utterances more complex than those of the others. Although the figures are low, they give some hints about emerging complexity from another point of view as well. Aleksi in particular has advanced in the number of elaborated components in his utterances. Five out of nine four-component utterances have three elaborated components and the component types in these utterances are among the most complex used by group A.

No clear relationship between component complexity and IPSyn scores can be found on the basis of the results for group A. The only contact between the two analyses is Aleksi's and Leo's performance.

TABLE 36 IPSyn scores in group A categorized in nominal phrase (NP), verb construction (V), and sentence structure (SS) sections.

ID	Name	NP	VP	SS	Total
22	Anna	8	15	19	42
6	Ronja	14	19	18	51
18	Saku	14	18	21	53
17	Tatu	13	22	21	56
29	Elisa	13	20	24	57
36	Sampo	15	23	19	57
35	Seppo	15	20	25	60
31	Aleksi	15	24	27	66
34	Leo	14	21	33	68

These two boys have the highest IPSyn scores and produce the most complex component structures in group A. This may be only coincidental, but nevertheless it gives some scope for speculation. A quick look at the detailed IPSyn scores in Table 36 shows that Aleksi and Leo have high points in both the verb construction (V) and sentence structure (SS) sections. Together these two sections may provide a favourable locus for complex component structures to emerge. In the V-section, there are many morphological forms (person inflection, tense inflection, mood inflection, passive form) as well as syntactic structures (negation expressions, infinitival constructions). When these are combined, they can result in complex components. Furthermore, if these complex V-structures are combined with multi-component items in SS-sections, the result is productions that stand out in the Utterance Analysis. Thus, the key here seems to be combinations. Examples 51–53 from Aleksi and Leo show, that they do combine details together in their utterances.

Example 51

V= [S2 + (S2 + (2M))], NP = [S3 + (M2 + M)], NP = [M]

1.		NP PRO:DEM		V			NP	NP N
2.	AUX: NEG			V	N	CONJ	N	ill
3.			V	V	pl		adess	
4.			neg	1Inf	adess			
	Ei	*nää*	*voim*	*mennä*	*vaatteilla*	*ja*	*mekolla*	*kylpyy.*
	Ei	nämä	voi	men-nä	vaatte-i-lla	ja	meko-lla	kylpy-yn.
	No	these	can	go-1INF	cloth-PL-ADESS	and	dress-ADESS	bath-ILL.

'These can't go and take a bath wearing clothes and a dress. ' (31 Aleksi; 1/193)

Example 52
V = [S2 + (2M)], NP = [M]

1.	ADVP ADV		V	NP PRO	NP PRO:DEM
2.		V	V		adess
3.		2s	1Inf		

Nyt *voit* *le3nt`äät* *tinä* *tällä.*
Nyt voi-t lentä-ä sinä tä-llä.
Now can-2S fly-1INF you this-ADESS.
'Now you can fly with this.' (34 Leo; 2/213)

Example 53
ADVP = [M], V = [S2 + (M)], NP = [M]

1.	ADVP ADV		NP PRO	V	NP PRO:DEM
2.	adess	AUX		V	partit
3.		1s			

Tääll+ *oom* *minä* *kaatamassa* *näitä.*
Tää-llä ole-n minä kaatamassa näi-tä.
Here-ADESS be-1s I cut these-PARTIT.
'Here I am, cutting these.' (34 Leo; 4/842)

Aleksi's utterance in Example 51 is worth 12 points in the IPSyn analysis. Included in the utterance are plural and adessive case in the NP-section (*vaatteilla* 'cloth + pl + adess'; N1, N2), as well as V+ I infinitive (*voi mennä* 'can go'; V12) and negative auxiliary + negative stem (*ei voi* 'can't'; V15) together (*ei voi mennä* 'can't go'; V16) in the V-section. In addition to that there is a co-ordinating conjunction (*ja* 'and'; S17) connecting two words (*vaatteilla ja mekolla* 'cloth+pl+adess and a dress+adess'; S19). On top of this, there is an illative case (*kylpyyn* 'to a bath'; N4), and two adverbs beside a verb (*vaatteilla ja mekolla, kylpyyn;* V10 for one adverb and V11 for the two adverbs). The whole utterance is a combination of more than two words (S1), with a subject (*nämä* 'these (people)') and a predicate (*ei voi mennä*; S2). Leo's two examples are worth 9 (Example 52) and 7 points (Example 53). Thus, our speculations about the importance of the combination of structural details proved to be right, and the component structure inventory, the knowledge of component combinations in different utterance categories together with the IPSyn sections and details in them were the clues that guided us to the right conclusions.

Examples 51–53 show another thing about IPSyn and the Utterance Analysis as well. They demonstrate concretely how powerful one utterance can be both in IPSyn and the Utterance Analysis. In Aleksi's case Example 51 gives 12 IPSyn points, which is 18% of his total IPSyn score of 66 points. For the Utterance Analysis, it provides two four-layered component types (V and NP-components), and a combination of components which is more complex than any produced by other children in the group. This might explain, at least partly the apparent contradictions between MLU and IPSyn, since one single utterance may affect the IPSyn score significantly whereas in MLU one utterance out of 80 is not enough to raise the average value noticeably.

8.4.2.2 Group C

In group C, the relation between MLU and IPSyn is similar to that for group A but on a different level. The MLU values range from 6.050 to 6.350 and the IPSyn scores from 62 to 84. The distribution of component types seems to be quite similar across children as shown in Table 37. Only Jaakko (11) is different from the others because he has fewer elaboration types. However, he concentrates on four-component utterances, producing 33 of them, while others in group C have 18–27 such utterances (see Table 12 in Chapter 6).

TABLE 37 The inventory of the elaboration types, and their occurrence in children's utterances in group C. The numerals 2, 3, and 4 refer to occurrences of each elaboration type in two-, three- and four-component utterances respectively.

ID **Name**	**11** **Jaakko**	**2** **Riikka**	**27** **Janna**	**7** **Taru**	**23** **Anniina**	**30** **Elina**	**19** **Tuomo**	**3** **Laura**	**1** **Tuija**
IPSYn	62	68	69	70	72	74	75	76	84
MLU	6.112	6.287	6.3	6.35	6.05	6.325	6.237	6.25	6.,25
2-layered components									
[M]	234	234	234	234	234	234	234	234	234
[S2]	34	3	34	234	234	34	234	34	34
[S3]	4				4				
3-layered components									
[M2]	4	234	34	34	234	234	234	3	234
[S2+(M)]	34	234	234	234	34	234	234	234	234
[S2+(2M)]	23	234	234	234	234	34	234	234	34
[S2+(S2)]					2	4	3	34	2
[S2+(S2+M)]		3							
[S3+(M)]			3						
[S3+(2M)]									2
[S3+(S2+2M)]					3				
4-layered components									
[M3]					3			3	
[S2+(M2)]		3			4		3		
[S2+(M2+M)]		23				234	234	4	
[S2+(S2+(M))]		2	3		2	3	34		24
[S2+(S2+(2M))]				2			3	2	
[S2+((S2+(M))+M)]			3	2				3	
[S2+((S2+(2M))+M)]		3				3			3
[S3+((S2+(M))+M)]	4			4		3			
[S3+((S2+(2M))+M)]					3		3		
5-layered components									
[S2+(S2+(M2))]		2							
[S2+(S2+(M2+M))]			23						
[S2+(M3)]								2	

Jaakko's pausity of elaboration types, his having the lowest IPSyn score, and his large number of utterances with four components all suggest that he is repeating some long utterance structures. This is a way to keep MLU high without providing much for IPSyn to register, since it counts maximally two different occurrences of each structural item. Such repetition also results in fewer different elaboration types. A look at Jaakko's four-component utterances supports this conclusion. Indeed, Jaakko has several repetitive features in his utterances. In his four-component utterances 22 out of 33 begin with the ADVP-component *sitten* 'then'. He also uses a two-word ADVP-component which consists of a locative adverb and an affirmative particle. This component has several variants such as *tuonne noin* 'to over there', *tässä näin* 'over here', and *tuossa noin* 'in over there'. Jaakko's four-component utterances mainly employ two distinctive sentence structures, first "*sitten*" + S + V + ADV and secondly "sitten" + V + O + ADV. In the former, the verb is usually a general multipurpose verb such as *menee* ´go´or *tulee* ´come', and in the latter the verbs are in the passive form. In the adverbial position, the combination of a locative adverb and an affirmative particle is very common. Such repetitive sub-parts insure high MLU but provide very little new material for the other methods of analysis. Examples 54 and 55 show two prototypical utterances from Jaakko, which include all the repetitive features discussed above.

Example 54
sitten + S + V + ADV

1.	ADVP ADV	NP N	V	ADVP	
2.				ADV:DEM	PTL
3.				ill	
	Hitte	*hattaat*	*menee*	*tähän*	*näi.*
	Sitten	rattaat	menee	tä-hän	näin.
	Then	go-cart	go	here-ILL	so.

'Then a cart goes over here.' (11 Jaakko; 21/999)

Example 55
sitten + O + V + ADV

1.	ADVP ADV	NP N	V	ADVP	
2.			pass	ADV:DEM	PTL
	Hittet	tämä	ajetaan	tuonnen	noi .
	Sitten	tämä	aje-taan	tuonne	noin.
	Then	this	drive-PASS	there	so .

'Then, this is driven over there.' (11 Jaakko; 28/692)

Tuija (1) differs from others remarkably in her IPSyn score but there is nothing exceptional in her inventory of elaboration types which would suggest that she is using structurally more complex components. Neither does the complexity of the elaborations grow alongside IPSyn scores for group C. Thus, the results for the distribution of elaboration types in group C support the results already presented for group A. The inventory of component combinations in group C (see Appendix 6) does not give any new clues which would suggest a way to connect the IPSyn results with the Utterance Analysis. IPSyn is an inventory of individual structural features (such as distinctive morphological forms) and entire linguistic productions (such as S+V+O) but a clear connection between these two is lacking. This is where Utterance Analysis makes its contribution, focusing on co-ordination between the elements both within and between components. IPSyn does not reflect the kind of structural complexity identified in this study. However, it is possible that if frequencies had been taken into account, there might have been more connections. This is because the same elaboration may vary in outcome, that is, it can be the structural basis for several items in IPSyn. Earlier it was claimed that, for example, [S2+(M)] and [S2+(2M)] are suitable bases for noun phrases as well as a verb construction (see Section 8.2.3). Especially in the latter case, they can be used in negation expressions, in other auxiliary constructions, in perfect tense, and in infinitive phrases. All these structures are separated in IPSyn, and therefore, a more frequent use of these elaboration types may provide more structures for IPSyn to count. Although such a connection is possible, it seems at the moment that it is MLU that is more closely connected to component structure.

There is yet another reason to suspect that MLU is the measure that most closely connects to component complexity. If the inventories of elaboration types in groups A (Table 35) and C (Table 37) are compared, group C clearly differs from group A. There are more different elaboration types, particularly more firmly established 3-layered types occurring in all utterance categories in group C. All the children in group C except Jaakko (11) have several 4-layered elaboration types which occur mostly in three-component utterances. Most children in group A have only one kind of four-layered elaboration, and this generally occurs only in two-component utterances. Two children in group C also use a five-layered elaboration type. In other words, the children in group C have a larger inventory of deeply elaborated types, and they use these types in larger utterances than children in group A. However, when the IPSyn scores and MLU values for these groups are compared, the differences between the groups are emphasized by MLU. There is a gap of 1.662 between the highest value in group A (Elisa 4.388) and the lowest value in group C (Anniina 6.050) whereas the IPSyn scores of the groups overlap (for group A the scores are 42–68 and for group C 62–84). In fact, the scores form a continuum proceeding in small steps from 51 to 76 points, and only the lowest score in group A (42 points) and the highest in group C (84 points) deviate from the general trend. The results from inventory of elaboration types showed a clear difference between the groups, and the gap of 1.662 in MLU is in concordance with that, while the IPSyn scores cross the gap without showing it up.

8.4.3 Mean Length of Utterance versus the structural complexity of components

Since we now have reasons to believe that MLU and complexity of components are connected, the relationship between the two will be investigated, this time in such a way that all 24 children in groups A, B, and C are involved[16]. This will create a firmer and wider basis for the comparison of MLU and component complexity.

In Table 38, an inventory of all the elaboration types used in the two-, three- and four-component utterances is introduced in a linear order of complexity. With the linear order of complexity an ordering is employd which is a combination of the absolute complexity discussed earlier and the linear notation. This is contrary to the principles for determining structural complexity that were discussed in previous chapters.

16 The original number of children in this study was 40. However, only utterances from 24 of them were selected for more detailed Utterance Analysis, because they had the most contradictory results in MLU and IPSyn. The purpose of the selection was to make sure that the study focuses on the problems arising from the differences between IPSyn and MLU as methods.

TABLE 38 The inventory of elaboration types and their occurrence in utterances for children A, B and C, ordered by MLU values. Numbers 2, 3, and 4 refer to occurrences of each elaboration type in two-, three- and four-component utterances respectively.

ID Name	**35 Seppo**	**22 Anna**	**18 Saku**	**36 Sampo**	**17 Tatu**	**6 Ronja**	**31 Aleksi**	**34 Leo**	**29 Elisa**	**13 Joel**	**8 Aino**	**15 Mika**
MLU	3,688	3,888	3,962	3,987	4,063	4,1	4,2	4,263	4,388	4,713	4,775	4,838
2-layered elaborations												
[M]	234	234	234	234	234	234	234	234	234	234	234	234
[S2]	23	234	23		23	234	234	234	23	234	23	234
[S3]							2				4	
3-layered elaborations												
[M2]	23			23			34	34	234	3	3	234
[S2+(M)]	23	234	234	23	234	23	234	23	234	23	23	234
[S2+(2M)]	2	2	23	23	3	2	23	34	2	23	23	23
[S2+(S2)]						2						
[S2+(S2+M)]												
[S3+(M)]			2			2						
[S3+(2M)]		3									2	4
[S3+(3M)]										3		
[S3+(S2+2M)]												
4-layered elaborations												
[M3]										23		
[S2+(M2)]	23						23			2	3	
[S2+(M2+M)]								4				
[S2+(2M2)]				3								
[S2+(S2+(M))]			2			2		2	2	3		3
[S2+(S2+(2M))]	3	2					4		2			
[S2+((S2+(M))+M)]	2		2		2						4	3
[S2+((S2+(2M))+M)]												
[S2+(S3+(3M))]							3					
[S3+(M2+M)]							4					
[S3+(M2+2M)]												
[S3+((S2+(M))+M)]												
[S3+((S2+(2M))+M)]												
5-layered elaborations												
[S2+(M3)]												
[S2+(M3+M)]												
[S2+(S2+(M2))]												
[S2+(S2+(M2+M))]												
Total	8	6	7	5	5	7	10	7	7	9	9	8

ID Name	24 Sanni	38 Pekka	10 Santeri	23Anniina	11 Jaakko	19 Tuomo	1 Tuija	3 Laura	2 Riikka	27 Janna	30 Elina	7 Taru
MLU	4,963	5,275	5,713	6,05	6,112	6,237	6,25	6,25	6,287	6,3	6,325	6,35
2-layered elaborations												
[M]	234	234	234	234	234	234	234	234	234	234	234	234
[S2]	34	3	34	234	34	234	34	34	3	34	34	234
[S3]				4	4							
3-layered elaborations												
[M2]	3	34	234	234	4	234	234	3	234	4	234	34
[S2+(M)]	234	234	234	34	34	234	234	234	234	234	234	234
[S2+(2M)]	34	3	23	234	23	234	34	234	234	23	34	234
[S2+(S2)]				2		3	2	34			4	
[S2+(S2+M)]	2								3	3		
[S3+(M)]										3		
[S3+(2M)]							2					
[S3+(3M)]					4							
[S3+(S2+2M)]				3								
4-layered elaborations												
[M3]				3				3				
[S2+(M2)]			3	4	3	3						
[S2+(M2+M)]						234		4	23		234	
[S2+(2M2)]												
[S2+(S2+(M))]	2	2		2		34	24		2		3	
[S2+(S2+(2M))]						3		2				2
[S2+((S2+(M))+M)]								3		3		2
[S2+((S2+(2M))+M)]									3		3	
[S2+(S3+(3M))]												
[S3+(M2+M)]												
[S3+(M2+2M)]	3											
[S3+((S2+(M))+M)]											3	4
[S3+((S2+(2M))+M)]				3		3	3					
5-layered elaborations												
[S2+(M3)]								2				
[S2+(M3+M)]			34									
[S2+(S2+(M2))]									2			
[S2+(S2+(M2+M))]										23		
Total	8	6	7	12	8	11	9	11	10	9	10	8

However, this compromise was made to make it possible to present potential trends in the growth in complexity in the simple form of a table. The elaboration types are therefore arranged in groups on the basis of their layered structure, and within these layer groups, types are presented in subgroups with similar kinds of elaboration, resembling the branches of each elaboration family introduced earlier. The children in groups A, B and C are ordered according to their MLU values.

The first thing to note is that the absolute number of the elaboration types per child does rise as the MLU value increases, ranging from five to twelve. However, the increase in types is not steady. For example, Seppo, Mika, Jaakko and Taru all have eight different elaboration types, but their MLU values range from the lowest (Seppo 3.688) to the highest (Taru 6.350) in the whole group of 24 children. This implies that the differences between children at the opposite ends of the MLU scale (or the IPSyn scale in turn) are not necessarily reflected in the specific structures they use but more likely in how these structures are used, combined, and spread through their language. In Chapters 6 and 7 concerned with the syntactic components and their layered structure, the most striking result was the similarities between the groups, rather than the differences, although it was the differences that were being sought in the first place. The similarities found here give further support[17].

Another noticeable factor in Table 38 is the differences in the four-layered elaboration types. The children whose MLU value is in the lower half of the scale, have few four-layered elaboration types in their speech, whereas those in the upper half have greater variation in the four-layered types, including even the most complex structures such as [S3+((S2+(2M))+M)]. Moreover, children with a lower MLU use their four-layered elaborations mostly in two-component utterances, but children with higher MLUs use them more often in three-component utterances. Individual variation is, however, substantial. Despite his MLU value of 4.200, Aleksi (31) uses four different four-layered elaborations, and they can be used in the three- and four-component utterances. According to his inventory of elaboration types, Aleksi belongs among the children with higher MLU. Contrary examples are Janna (27) and Jaakko (11), both having MLUs exceeding 6.000 while their elaboration inventories have only one four-layered type. In addition to progress with four-layered elaborations, some of the children at higher MLU levels are using five-layered elaboration types as well, although the variety of structures is still small.

The changes in the number of elaboration types (although not in a steady fashion), the addition of the four-layered elaboration types, and the extension to five-layered components are not the only signs of growing complexity in utterances. Table 38 shows that there is also something happening in the three-layered elaborations. In the lower half of the MLU scale, three-layered elaboration types are found in two- and three-component utterances but in the upper half of the scale the same elaborations are being used in four-component

17 However, it must be kept in mind that Table 38 lists only the elaboration types, and the tokens of each type are not included. If the tokens were considered, they might provide some evidence of differences between the children.

utterances as well. The spread of elaboration types reveals more advanced abilities in combining more complex structures in utterances with more components. It also reveals an increased variety of potential manifestations of a given structure that a child has brought into play. In other words, these results show a simultaneous advance in vertical and horizontal dimensions of structural complexity.

The inventory of elaboration types and the changes in it, which are in line with growth in MLU, suggest that MLU has after all some potential for registering progress in structural complexity. In a way, this is not surprising, because growth in complexity is defined as the addition of new morphological and syntactic elements into utterances, and these are exactly the same elements that are calculated in MLU, one by one. According to the inventory of elaboration types, growth in structural complexity is the result of a multidimensional penetration of structures. On the one hand it proceeds by adding new elaboration types as well as new layers. On the other hand, it involves the spreading of elaboration types that have already been produced on one or two occasions to utterances containing more components. Therefore, growth in complexity involves the acquisition of both new items and new ways to use items learned earlier.

The inventory of elaboration types discussed above concentrates mostly on the emergence of new components. To get more information about how the components are used in utterances and whether the changes in use are reflected in MLU, an examination of component combinations is needed. Inventories of elaboration combinations in two-, three- and four-component utterances are presented in Appendix 7.

The first table in Appendix 7 shows the elaboration types that are used in utterances including only one elaborated component. Examples 56–58 show the different kinds of utterances that are involved:

Example 56
A two-component utterance with one elaborated component (V = [S2 + (M)])

1.		NP PRO:DEM	V
2.	AUX:NEG		V
3.			neg

Ei	*tää*	*men`e.*
Ei	tämä	mene.
Not	this	go&NEG

'This one does not go.' (35 Seppo; 39/256)

Example 57
A three-component utterance with one elaborated component (NP = [S2+(2M)])

1.	NP N	V		NP
2.			PRO:DEM	N
3.			ill	ill

Issä *menee* *tähän* *tuolii.*
Isä menee tähän tuoliin.
Father go this-ILL chair-ILL.
'Farther goes to sit in this chair.' (38 Pekka; 10/1893)

Example 58
A four-component utterance with one elaborated component (ADVP = [M])

1.	ADVP ADV:DEM	NP N	V	NP N
2.	adess			

Täälä *kotona+* *om* *makuupus+si .*
Tää-llä kotona on makuupussi.
Here-ADESS at home is sleeping bag.
'There is a sleeping bag here at home.' (17 Tatu; 14/253)

The potential connections between MLU and component complexity are not shown in the number of elaboration types used in these utterances with only one elaborated component. The numbers range from four to nine different types, and they are distributed randomly with respect to growth in MLU. In the lower half of MLU values there is a clear tendency for more complex elaboration (especially with the four-layered elaboration types, but with some three-layered ones as well) only to appear in two-component utterances, while the simpler elaborations such as [M] seem not to have any restrictions on where they appear. In the upper half of the MLU scale the more complex elaboration types are used in three- and even four-component utterances. This is partly due to the fact that in group C, which covers approximately the highest third of the MLU scale, only 67 utterances of the total of 720 consist of 2 components (see Table 13, Chapter 6), and therefore there is little space to show variation and advancement in two-component utterances. However, the distribution of more complex elaboration types into three- and four-component utterances is a sign of an advance from a strong trade-off stage towards a level where complexity is no longer as conditional on the number of components as it used to be. Morover, the simpler elaboration types spread into four-component utterances. Therefore, the trend for elaboration to spread to utterances with more components in them does follow the rise in MLU.

Moving on to utterances where two components are elaborated brings us to the largest set of variations in elaboration types (Appendix 7, Table II). There are 36 different combinations of two elaborated components used in two-, three- and four-component utterances. The compositions consist mostly of pairs of two- and three-layered (13 combinations) or two- and four-layered elaborations (eight combinations). There are also combinations of two- and two-layered (three combinations), three- and three-layered (eight combinations) and four- and three-layered (two combinations) elaborations. A five-layered elaboration is involved in two combinations, once together with a 2-layered elaboration and once with a three-layered one. Examples 59–61 provide specimens of utterances with two elaborated components.

Example 59
A two-component utterance with both components elaborated (4+2 layers)
(V = [S2+(S2+(2M))]; NP = [M])

1.	NP N		V	
2.	gen	V	V	
3.			V	PTL
4.			1inf	clit

Tääv *voi* *ottaap* *poiski* .
Tämä-n voi otta-a pois-kin.
This-GEN can take-1INF away-clit
'This can be taken away, too.' (3 Laura; 18/642)

Example 60
A three-component utterance with two elaborated components (3+2 layers)
(V = [S2+(2M)]; NP = [M])

1.	NP PRO:DEM		V	NP
2.		V	V	all
3.		cond	1inf	

Te *vois* *soittaa* *mummulle* .
Se vo-isi soitta-a mummu-lle.
It van-VOND phone-1INF Grandma-ALL.
'It could phone Grandma.' (19 Tuomo; 16/269)

Example 61
A four-component utterance with two elaborated components (2+2 layers)
(V = [M]; NP = [M])

1.	N	NP PRO	V	NP N
2.			2s	partit

Äi *tinä* *ajat* *noosua* .
Äiti sinä aja-t norsu-a.
Mother you ride-2S elephant-PARTIT.
'Mother, you'll ride on an elephant.' (34 Leo; 16/594)

Tendencies that were already evident in utterances with only one elaborated component show up in two elaborated components as well. The number of different component combinations rises together with MLU, but not steadily (Table 39). The number of combinations of 3+3 layers and 4+2 layers is higher on the upper half of MLU scale (see Appendix 7, Table II). Usually these combinations appear in three-component utterances. Yet another trend seems to follow MLU value: there are several combinations of three- and two-layered elaboration types that become more widely established as MLU rises. In the lower half of the MLU scale all the children except Sampo (ID 36) produce utterances with the combination [S2+(M)]+[M], but usually only in two- and three-component utterances, whereas in the upper half of the MLU scale, almost all the children use this structure in four-component utterances as well. With the combination [S2+(M2)]+[M], which is another step towards more complex structures, the development is twofold. First, more children in the upper half of the MLU scale use this combination than in the lower half, and second, they all produce it in three-component utterances (some even in the four-component utterances). The children with a lower MLU level, however, concentrate on producing it in two- or three- component utterances. Similarly, a twofold development towards more complex utterances is seen in combinations of [M2]+[M] and [S2+(M)]+[S2].

TABLE 39 Number of different elaboration combinations in utterances with two elaborated components, ordered according to MLU values.

ID	Name	MLU	Combination types	ID	Name	MLU	Combination types
35	**Seppo**	3,688	8	**24**	**Sanni**	4,963	7
22	**Anna**	3,888	5	**38**	**Pekka**	5,275	7
18	**Saku**	3,962	5	**10**	**Santeri**	5,713	8
36	**Sampo**	3,987	3	**23**	**Anniina**	6,05	9
17	**Tatu**	4,063	4	**11**	**Jaakko**	6,112	9
6	**Ronja**	4,1	4	**19**	**Tuomo**	6,237	10
31	**Aleksi**	4,2	5	**1**	**Tuija**	6,25	11
34	**Leo**	4,263	8	**3**	**Laura**	6,25	14
29	**Elisa**	4,388	5	**2**	**Riikka**	6,287	12
13	**Joel**	4,713	8	**27**	**Janna**	6,3	5
8	**Aino**	4,775	9	**30**	**Elina**	6,325	12
15	**Mika**	4,838	10	**7**	**Taru**	6,35	11

The tendency towards producing more complex utterances as MLU rises, either by introducing new, more complex elaboration combinations into utterances or by spreading earlier acquired combinations into utterances with more and more components in them, is rather clearly shown in the inventory table (Appendix 7, Table II). However, individual differences and deviations in both directions from this general tendency are clear as well. Seppo (38) has the lowest MLU value of all, and yet his inventory of elaboration combinations has eight items, half of them belonging to the categories of 3+3 layers and 4+2 layers. If his simple combinations were more firmly established and appeared in all utterance categories, his component combinations would put him among the children with an MLU of 5.000–6.000. Sampo (36), Tatu (17), Ronja (6) and Janna (27) represent a deviation in another direction; they produce far fewer different elaboration combinations than expected on the basis of their MLU values. Sampo's, Tatu's, and Ronja's MLUs are rather low (3.987–4.100) but the three or four combinations they use is less than other children use with similar MLU level. All the other children with MLUs between 3.688 and 4.388 use five to eight elaboration combinations. Moreover, Sampo, Tatu and Ronja have only combinations with 2+2 and 3+2 layers while all other children on the lower half of the MLU scale have at least one elaboration combination consisting of 3+3 or 4+2 layers. Janna's (27) MLU is among the highest (6.300) but still her elaboration combinations are similar to those of Sampo, Tatu, and Ronja, differing only in that they have spread to a wider range of utterance categories in her performance. The total lack of more complex combinations makes her very different from other children with an MLU around 6.000. It seems that even though MLU and complexity of existing elaboration combinations have much in common and they tend to move in tandem there are factors that Utterance Analysis brings out and MLU does not. MLU does not discriminate talkative or repetitive children from children who expand their utterances with new grammatical features. On the other hand, the results from Utterance Analysis reveal children who exploit only a few grammatical means in their productions and therefore lag in structural complexity despite higher MLU levels.

Combinations of two elaborated components form the most versatile collection of structures in this data. For three elaborated components there is much less versatility (see Appendix 7, Table III). Another difference from the case with one or two elaborated components is that some children do not have this kind of combination at all. Ronja (6) and Elisa (29) do not have a single three- or four-component utterance containing three elaborated components, nor do they have any combinations with five-layered elaboration. For these structures the number of combinations does not seem to follow a rise in MLU values. Instead, the number of componenta stays constantly at a level of approximately two to three. However, there is a sharp peak where MLU ranges from 6.237 to 6.287. Tuomo (ID 19) and Tuija (ID 1) have seven different combinations and Riikka (ID 2) has eight. These children are also among those who have examples of more complex combinations, such as [S2+(2M)]+[M2]+[M] and [S2+(S2+(M))]+[S2+(M)]+[M]. In addition, Tuomo, Tuija, and Laura (3) produce at least two elaboration combinations in both three- and four-component utterances, in other words, they have more evidence of firmly established structures than the others.

The last potential elaboration category is elaboration of four components in an utterance. Here we find no more than three different elaboration combinations, and only three children out of 24 produced examples of them (see Appendix 7, Table IV). Ronja (ID 6) and Laura (ID 1) have one combination and Taru (ID 7) has two different combinations in her productions. Since these combination types are so rare they do not provide any evidence of a relationship between MLU and the structural complexity of utterances.

On the basis of these results it seems that MLU and the structural complexity of components and component combinations are related to one another but only under certain conditions. Basically such connections are to be expected because MLU counts morphemes, and morphemes – suffixes and stems – are a means of expanding structural complexity as well. However, talkativeness and repetitiveness are factors that constantly affect MLU, and may mask a deficiency in structural complexity.

The results show that the relationship between MLU and the complexity of component combinations is most clearly seen in cases where two components are elaborated in two-, three- and four-component utterances. In the other categories the connection was not yet so evident. In utterances where two components are elaborated the elaboration is deeper in layers than in utterances with more elaborated components. The variety of combinations is also much larger. The number of elaborated components in these utterances is still so low that it does not substantially restrict the ways that the children use in elaboration. In other words, there is no evidence of a trade-off effect. The children can show their structural repertoire at its widest, and the differences in complexity between them come out clearly. Such a result is comparable with mathematical average values like MLU. The second factor is that many of the combinations of two elaborated components appear in three-component utterances. Because the proportions of three-component utterances range from 42.4% (group A) to 50.3 % (group C) there is a lot of material to be elaborated in this way. Therefore three-component utterances really do affect MLU, especially when they contain elaboration. The third factor that makes it possible for a relationship between MLU and the complexity of utterances to become evident in utterances with two elaborated components is at least partly a consequence of the preceding factor. When there is a lot of material in three-component utterances containing two elaborated components, it is a mathematical fact that it will show in a mean value. This is something that the other kinds of combinations cannot bring about. If only one component is elaborated, its effect is too small to be seen in the mean value and a possible connection between MLU and structural complexity is left uncovered. In the case of utterances with three elaborated components there is so much less material that performance in them cannot show up in MLU values either, whereas Utterance Analysis can handle all complexity levels in these utterances too.

8.5 Summary

In this chapter the focus was on the elaborated components and their use as parts of utterances of different lengths. They were examined from three different angles. First, before it was possible to compare the productions of individual children or groups of children in terms of complexity, the elaboration types had to be evaluated with respect to their structural complexity and in relation to each other. After that it was possible to compare how the groups performed within each elaboration layer and how the elaboration types were used in utterances. In other words, group performances were evaluated in terms of structural complexity in detail, within layers, elaboration type by elaboration type. The third angle focused on the potential connections between MLU, IPSyn and the complexity of elaboration types as well as elaboration combinations. The fact that the groups were already found to be different from each other in the investigation of layers suggested a relationship of some kind, but at that point there were only speculations about how MLU and IPSyn might reflect the properties of structural complexity. Now, once these properties were formally identified it was possible to find out whether the speculations could be substantiated.

In determining a scale of growing complexity of elaboration types in children's productions, the baseline for the evaluation procedure was that a strict linear order of complexity is not possible, since it would require the comparison of features that are not comparable. To avoid comparing the complexity of, for example, inflection [M] with the complexity of forming a phrase with two syntactic elements [S2], the elaboration types were arranged according to near relations into a net of families and branches rather than a linear continuum. This made it possible to see how elaboration types of the same family and branch developed from each other and how much variation existed within each elaboration layer. The discussion resulted in a netlike construction which describes the relationships between morphosyntactic structures.

After creating the netlike framework for discussing complexity differences, the group performances were evaluated. Group comparison revealed several interesting results about growth in complexity. First, the differences are relatively small if we look at the presence of elaboration types alone. The distinctions are most evident in four- and five-layered elaborations, but on the other hand, these types have very low frequencies in all groups. The situation changes in an interesting way, when the occurrence of elaboration is considered together with the size of utterances. In groups B and C the elaboration types are put to wider use than in group A.

When the spread of component types is examined qualitatively there is another interesting result. Group A is similar to the other groups as far as morphological elaboration is concerned, but as soon as the examination moves to elaboration types involving both syntactic and morphological elaboration, group A seems to lag behind the other groups. This is most obvious in case of the [S3]-family of elaboration. Therefore the investigation of elaboration types in connection with morphosyntactic complexity suggests that there are several directions for growth in complexity. One direction is the addition of four- and five-layered elaborations, which represent new innovations in the repertoire. They are not yet frequent but there are more of them in groups B and C than in A, implying that these groups are already able to make use of these kinds of complex structures. Another direction for growing complexity affords substantially greater scope, involving components which include more than one elaboration device and/or more than two elements to be elaborated. All this suggests that morphological elaboration is the primary way to get started in making more complex productions and that syntactic elaboration comes a bit later, followed by combined morphological and syntactic elaboration.

It is difficult to find points of connection between IPSyn scores and the results from Utterance Analysis. These two analyses consider structures from different points of view. In IPSyn a structure is a unit with a particular grammatical function, whereas in Utterance Analysis a structure is purely a structural entity. In IPSyn, utterances are broken down into functional subunits, whereas in Utterance Analysis the components are kept together and analysed for their structural details. The same structure may be used for several functions, like [M] in all the different nominal cases, and in verb inflection, as well as in clitics. On the other hand, IPSyn does not distinguish whether [M]-elaboration appears with other structural features to form more complex units, whereas in Utterance Analysis the combination of structural features is the most important property.

Structural complexity and MLU seem to have more in common than structural complexity and IPSyn. However, this is more evident under particular conditions. Structural complexity grows through combining free or bound elements in a component. The number of elements is the feature that MLU registers. But to be able to reflect growth in complexity, there must be enough instances to affect MLU, because one or two tokens are not enough to change an average which is calculated from 80 or even more utterances. The optimal circumstances are reached in three-component utterances with two elaborated components. Approximately 50 % of all utterances consist of three components so this provides a large body of material, and when two of the components are elaborated it will affect MLU. Probably two-component utterances with both components elaborated also strengthen the effect, but by themselves they are not sufficient to affect MLU values.

Despite the tendency suggesting that MLU and structural complexity are connected with one another, there were several children who clearly deviated from this tendency. Such variation suggests that component complexity is not the only factor that affects MLU. MLU does not separate repetition or talkativeness from complex innovations. It is well known that it is possible to reach similar MLU values by different means. However, there has not yet been any investigation of how large a proportion of children actually use means other than increase in complexity to achieve growth in MLU. According to the results of this study, the number of talkative and repetitive children is not substantial. In that sense, MLU might be an even more reliable measure than has been recognized. However, it is worth remembering that the connection between MLU and structural complexity was not clear in all contexts.

9 STRUCTURAL COMPLEXITY AND GENETIC RISK OF DYSLEXIA

In recent decades problems in phonological processing and dyslexia have been connected with each other in studies on reading difficulties. Along with phonological deficit, poor syntactic skills have been proposed as precursors of dyslexia. Whether poorer syntactic skills are a consequence of phonological deficit or a more independent symptom of dyslexia is an issue of constant debate. In this chapter the focus is on the relation between the structural complexity of expressive language and the genetic risk of dyslexia of the eleven risk children whose data was surveyed using Utterance Analysis.

9.1 The purpose of the analysis

One of the main goals of the present study was to examine the potential morphosyntactic differences in speech production between children with and without a high genetic risk of dyslexia. In the field of dyslexia studies, the poor phonological processing is now generally accepted as being highly likely to be responsible for causing reading difficulties. Moreover, syntactic skills have been found to vary between dyslectics and normal readers, but there are divergent views on what is the role of syntactic knowledge in dyslexia. There are some who regard low syntactic skills as an individual deficit characteristic of some dyslectic people, whereas others consider them to be a consequence of phonological deficit or a metalinguistic problem (for a review, see Leikin & Assayag-Bouskila 2004).

Scarborough (1990b, 1991) found that problems in syntactic production skills were evident between 2½ and 4 years of age. The results concerned sentence comprehension (Northwestern Syntax Screening Test, NSST), grammatical complexity (IPSyn) and utterance length (MLU). (Scarborough 1991.) In the present study MLU and IPSyn did not show clear differences between the groups, although a tendency in control group for a better performance on both scales was clear. However, it must be remembered that Scarborough's study differs from the present study in that it was retrospective research, and thus the children were divided into groups according to their diagnosed reading status and familial background, and not depending on a familial risk of dyslexia alone. Therefore, the risk group of the present study may include both potential dyslexic children from dyslexic families (the poorest performers in Scarborough's study) and potential normal readers from dyslexic families (the best performers in Scarborough's study). This may have an impact on the results and equalize the two groups in the present study.

In this chapter the possible differences in syntactic complexity between the risk and control children are re-evaluated. In the previous chapter it was concluded that IPSyn does not react to the same factors that are considered to be properties of structural complexity in the present study. MLU, for its part, has more capacity for tracing those properties but only in certain conditions. Thus, outside the scope of IPSyn and MLU there are aspects of structural complexity that have not yet been examined with respect to dyslexia using Utterance Analysis.

Because Utterance Analysis concerns elaboration as well as combination of components, this survey will focus on those kinds of syntactic components which risk and control children use, and whether their elaboration type inventories differ in frequency, quality or some other aspect. In addition, possible differences in component combinations are searched for.

TABLE 40 The children representing the risk and control groups in Utterance Analysis. The MLU values and IPSyn scores as well as the group means are presented. The original mean values in parenthesis refer to results of the original risk and control groups consisting of 20 children (see Chapter 4).

Control group					**Risk group**			
ID	**Name**	**MLU**	**IPSyn**		**ID**	**Name**	**MLU**	**IPSyn**
1	Tuija	6.250	84		22	Anna	3.888	42
2	Riikka	6.287	68		23	Anniina	6.050	72
3	Laura	6.250	76		24	Sanni	4.963	66
6	Ronja	4.100	51		27	Janna	6.300	69
7	Taru	6.350	70		29	Elisa	4.388	57
8	Aino	4.775	61		30	Elina	6.325	74
10	Santeri	5.713	65		31	Aleksi	4.200	66
11	Jaakko	6.112	62		34	Leo	4.263	68
13	Joel	4.713	64		35	Seppo	3.688	60
15	Mika	4.838	70		36	Sampo	3.987	57
17	Tatu	4.063	56		38	Pekka	5.275	65
18	Saku	3.962	53					
19	Tuomo	6.237	75					
Mean		5.358	65.8		**Mean**		4.848	63.3
(Original CG		5.09	63.8)		(Original RG		4.40	57.5)

Since the Utterance Analysis was conducted only on the data from the children in groups A, B and C, the same data sample is used now as well, with the children now separated into the original risk and control groups. The focus of the study continues to be on two-, three- and four-component utterances. The division into groups is presented in Table 40.

Table 40 shows that in this sub-sample the 13 control children still have better average values than the 11 risk children, but now the difference is narrower, especially in IPSyn points. Moreover the values are now slightly higher than those for the original groups of 20 children. The fact that the groups are more equal in this subsample is due to the criteria which required a similar performance on either of the scales for children to be assigned to subgroups A, B and C. This does not harm the forthcoming analysis and group comparison. On the contrary, in my opinion, in being so similar the initial state gives another opportunity to test the relationship between MLU, IPSyn and Utterance Analysis.

In the following sections I will first concentrate on an inventory of elaboration types used by risk and control children. After that the focus will be on how the children combine these elaborated and unelaborated components into utterances. Group comparisons will be made in order to uncover any differences in the structural complexity of utterances.

9.2 Elaboration types in the risk and control children

To find out whether there are differences between the risk and control children in their elaboration type inventories the results from the earlier analysis were rearranged so as to divide the children into risk and control groups. The inventories for both groups are presented in Table 41, and a detailed, child-specific description of inventories is available in Appendix 8.

The control children produce 24 different elaboration types, and the risk children produce 25. In both groups, the favourite types are the same two- and three-layered elaborations [M], [S2], [M2], [S2+(M)], and [S2+(2M)]. Even in the number of users of each elaboration type the groups are similar. The group differences almost exclusively involve situations where a component has only one user in one group as opposed to none in the other. I interpret this as evidence that these components are infrequent in general, and do not represent a crucial difference between risk and control groups, especially because it is not always a question of control children having and risk children not having a given elaboration type. In the four-layered elaborations, for example, it is the risk children that show greater versatility of types.

In Table 41 the elaboration types are arranged in linear order of complexity, which approximates the order of complexity within branches and structural families (although the branches are now listed in linear order instead of being side by side as in the original complexity networks described in

Chapter 7). However, the linear order also shows that both groups have elaboration types of the same complexity levels. Therefore, it can be claimed that risk children and control children cannot be distinguished from each other quantitatively or qualitatively on the basis of the elaboration type inventory.

TABLE 41 The inventory of different elaboration types used by the control (CG) and risk children (RG). The figures refer to the number of children using the elaboration type in each group (controls N = 13, risks N = 11).

		CG	RG
Two-layered elaborations			
	[M]	13	11
	[S2]	13	10
	[S3]	2	2
Three-layered elaborations			
	[M2]	10	10
	[S2+(M)]	13	11
	[S2+(2M)]	13	11
	[S2+(S2)]	4	2
	[S2+(S2+M)]	1	2
	[S3+(M)]	2	1
	[S3+(2M)]	3	1
	[S3+(3M)]	2	
	[S3+(S2+2M)]		1
Four-layered elaborations			
	[M3]	2	1
	[S2+(M2)]	5	3
	[S2+(M2+M)]	3	2
	[S2+(2M2)]		1
	[S2+(S2+(M))]	7	6
	[S2+(S2+(2M))]	3	4
	[S2+((S2+(M))+M)]	6	2
	[S2+((S2+(2M))+M)]	1	1
	[S2+(S3+(3M))]		1
	[S3+(M2+M)]		1
	[S3+(M2+2M)]		1
	[S3+((S2+(M))+M)]	1	1
	[S3+((S2+(2M))+M)]	2	1
Five-layered elaborations			
	[S2+(M3)]	1	
	[S2+(M3+M)]	1	
	[S2+(S2+(M2))]	1	
	[S2+(S2+(M2+M))]	1	1
Total number of types		**24**	**25**

A closer look at individual children (Appendix 8) does not give much evidence of differences between the control and risk groups, either. In the number of elaboration types per child, the control children have a somewhat higher mean of 8.4 types whereas the average number for risk children is 7.9. In both groups the median value is eight, and the number of elaboration types ranges from five to eleven for the control children and from five to twelve for the risk children. In addition, the spread of elaboration types into different utterance categories (two-, three and four-component utterances) seems to follow similar patterns for both groups. Of course, individual differences are present in both groups, but in general the groups do have the same tools for building utterances.

9.3 The use of elaboration types in utterances

9.3.1 The elaboration of components in utterances – a general view

The survey of the elaboration types used by the risk and control children suggests that the groups have similar tools and therefore similar opportunities to create utterances of varying degrees of structural complexity. However, more important than the inventory of types is the way they are used in creating utterances. Before going into details, the data consisting of two-, three and four-component utterances is surveyed in a more general way. The purpose of the survey is to find out how the elaboration of components is distributed in the utterances of risk and control children, and whether there are any quantitative differences in elaboration at group level.

The data was categorized according to two variables, the number of components in an utterance (two, three and four components) and the number of elaborated components (zero, one, two, three and four elaborated components). Thus, the utterances were divided into twelve categories (two-component utterances with no elaborated components, two-component utterances with one elaborated component etc.). There are a total of 768 utterances for the risk group and 914 utterances for control group. The distribution of elaboration in utterances is presented in Figures 59 and 60 and in Table 42.

Figure 59 and Table 42 show that the risk children clearly prefer to elaborate only one component in all three utterance categories. By contrast the control children prefer utterances with two elaborated components in three- and four-component utterances. In two-component utterances the preference for utterances including only one elaborated component is not as strong as it is fin the case of the risk children. Another noteworthy point shown by Figure 59 is the proportion of unelaborated utterances. In all the three utterance categories the risk group children produce approximately twice as large a proportion of these as the control children. The group differences are most evident in two-component utterances, but they start balancing out when the number of components in utterances increases.

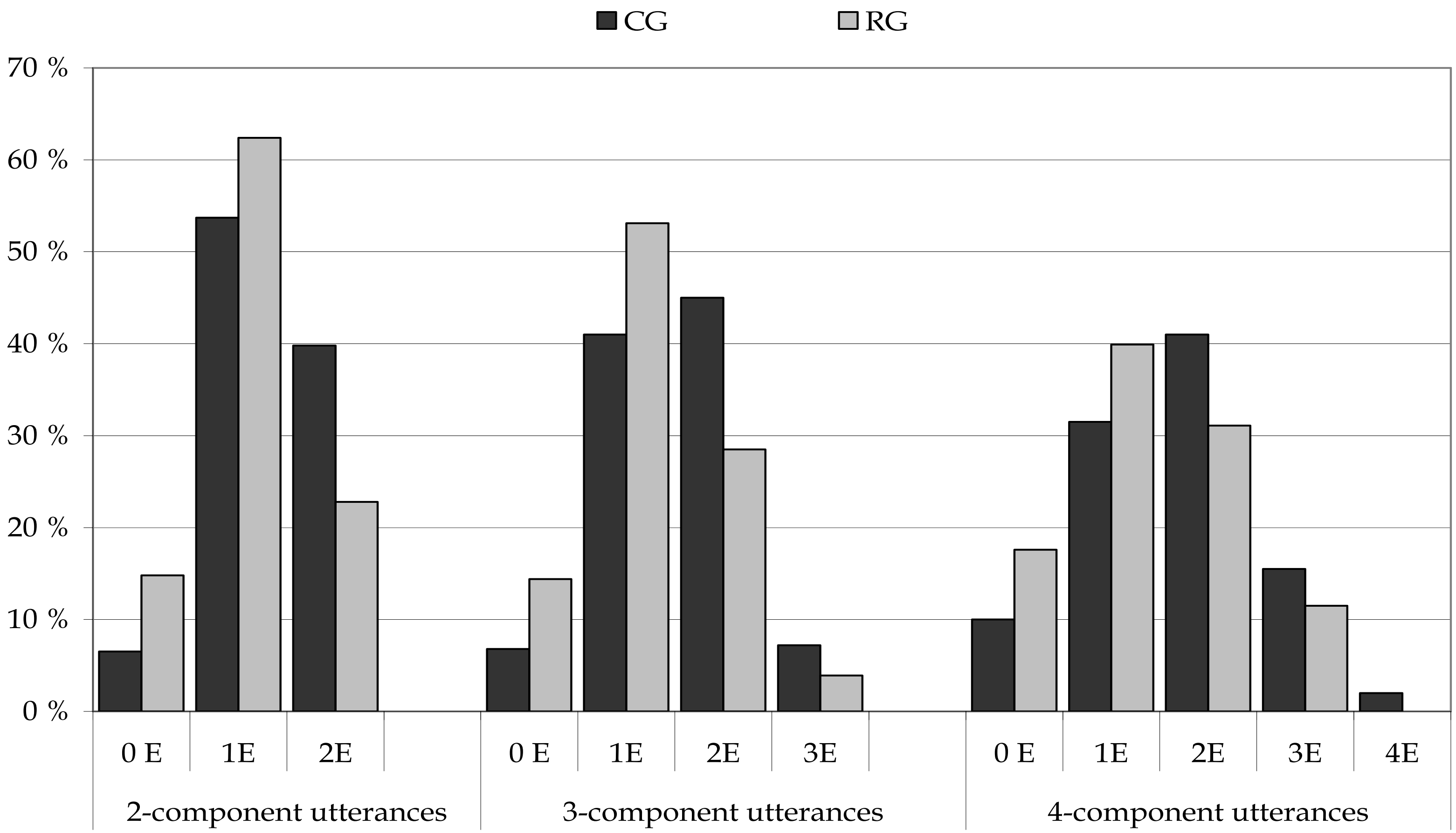


FIGURE 59 The elaboration of components in the two-, three-, and four-component utterances produced by children in control group (CG) and risk group (RG). The figure shows the proportions (%) of utterances with a given number of elaborated components in them produced by each group. (0E = no elaborated components in an utterance; 1E = 1 elaborated component; 2E = 2 elaborated components; 3E = 3 elaborated components; 4E = 4 elaborated components in an utterance.) The same results are presented in Table 42.

TABLE 42 The absolute numbers and proportions (%) of utterances (U) with a given number of elaborated components in them. The results for the risk group and the control group are presented in separate tables. These results are shown in graphic form in Figures 59 (each utterance category separately) and 60 (total proportions of each elaboration category).

Risk group											
Components	**Elaborated components per utterance**										**Total**
per	**0**		**1**		**2**		**3**		**4**		
utterance	**U**	%	**U**	%	**U**	%	**U**	%	**U**	%	**U**
2	28	14.8 %	118	62.4 %	43	22.8 %	*	*	*	*	189
3	62	14.4 %	229	53.1 %	123	28.5 %	17	3.9 %	*	*	431
4	26	17.6 %	59	39.9 %	46	31.1 %	17	11.5 %	0	0.0 %	148
Total	**116**	**15.1 %**	**406**	**52.9 %**	**212**	**27.6 %**	**34**	**4.4 %**	**0**	**0.0 %**	768
Control group											
Components	**Elaborated components per utterance**										**Total**
per	**0**		**1**		**2**		**3**		**4**		
utterance	**U**	%	**U**	%	**U**	%	**U**	%	**U**	%	**U**
2	14	6.5 %	116	53.7 %	86	39.8 %	*	*	*	*	216
3	34	6.8 %	204	41.0 %	224	45.0 %	36	7.2 %	*	*	498
4	20	10.0 %	63	31.5 %	82	41.0 %	31	15.5 %	4	2.0 %	200
Total	**68**	**7.4 %**	**383**	**41.9 %**	**392**	**42.9 %**	**67**	**7.3 %**	**4**	**0.4 %**	914

When the results of the three utterance categories are combined in Table 42, the preference of the risk group for less elaborated utterances is emphasized. Table 42 and Figure 60 show once again the fact evident already in Figure 59 that the risk children prefer utterances with only one elaborated component**.** The combined proportion of 52.9% putting together all three utterance categories shows this even more clearly than the bars in Figure 59 separating utterance categories from each other. Similarly, the evenly divided predominance of one and two elaborated components in the control group with proportions of 41.9% and 42.9% respectively is shown up in Table 42 and Figure 60.

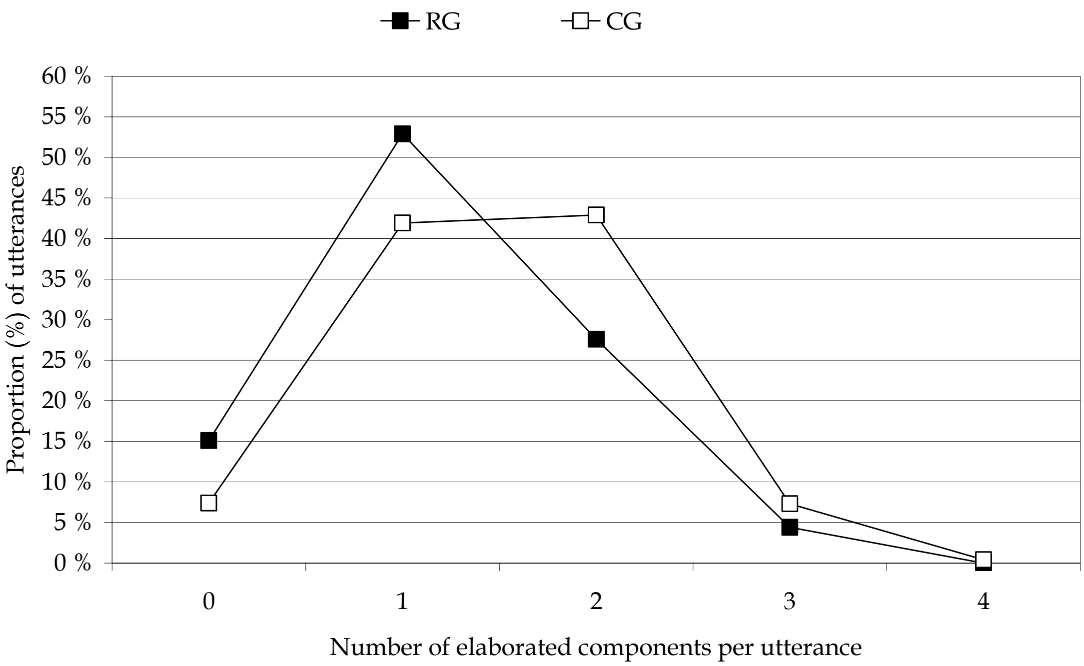


FIGURE 60 Proportions of utterances including 0, 1, 2, or 4 elaborated components for risk and control group. The total results are presented also in Table 42.

Despite the evident similarities in MLU, IPSyn, and even in inventory of elaboration types, the risk group and the control group seem to have at least one clear difference in how elaborated components are used in utterances. For the risk children using more than one elaborated component in an utterance is not so easy. The control children, by contrast, have taken a step forward and use more elaboration and combinations of elaborated components in their utterances. Although this difference is quantitative it might turn out to have qualitative properties as well. To find out whether there are also qualitative aspects in this difference, each elaboration category will now be surveyed separately.

9.3.2 Utterances including one elaborated component

Our examination of elaboration use focuses now on utterances where a single elaborated component is combined with one, two or three unelaborated ones.

Previously, it was found that this type of utterance accounts for a substantial proportion of two-, three- and four-component utterances in both groups and that especially the risk children prefer them. Altogether 406 (52.9%) out of the total of 768 utterances produced by the risk children and 383 (41.9%) out of the 914 utterances by control children belong to this category. The results for the individual children in both groups are presented in Table I in Appendix 9. Table 43 below provides a summary of the results shown in Appendix 9, with the addition of the numbers of utterances.

TABLE 43 Survey of elaboration types used by risk and control children in two-, three and four-component utterances (U) containing only one elaborated component.

Risk group					
	Types	**Range**	**Median**	**Mean**	**Number of U**
2-layered elaboration	2	1–2	2	1.8	262
3-layered elaboration	7	3–4	3	3.3	125
4-layered elaboration	8	0–3	1	1.3	17
5-layered elaboration	1	0–1	0	0.1	2
Total	18	5–9	7	6.5	406
Control group					
	Types	**Range**	**Median**	**Mean**	**Number of U**
2-layered elaboration	3	1–3	2	1.9	226
3-layered elaboration	7	1–4	3	2.9	135
4-layered elaboration	5	0–2	1	0.9	21
5-layered elaboration	1	0–1	0	0.1	1
Total	16	4–8	6	5.7	383

The survey reveals that the risk children use a slightly wider range of elaboration types in their utterances with only one elaborated component. There are on average 6.5 different elaboration types in the utterances of risk children whereas for control children the average is 5.7 types. Another noteworthy result is that all the risk children perform very coherently as a group with respect to two- and three-layered elaboration types. There are one or two two-layered elaboration types and three or four three-layered ones in every risk child's utterances, whereas the variation ranges for the control children are from one to three and from one to four respectively. In four-layered elaboration types, the risk children clearly perform better than the control children when the number of elaboration types is considered. In the risk group three out of eleven children employ only one elaboration type, but in the control group there are as many as eight children behaving similarly. In addition to this, four risk children have two types and one child as many as three types of four-layered elaborations in their utterances but in the control group, the largest number of types is two and only two children achieve it. In other words, the risk children use a wider variety of elaboration types and they perform more coherently as a group than the control children. The difference suggested by the average numbers of elaboration types is mostly due to the uniformity of the risk children in two- and three-layered elaboration types as well as to their more advanced performance in four-layered elaborations. In addition to their advanced performance in quantity, the risk children have qualitatively more complex 4-layered elaboration types in their productions as well.

On several occasions a trade-off effect has been discussed, inspired by specific results. Here, too, there is occasion for explanations that originate from this theoretical view. To oversimplify, four-layered elaborations are more complex than three-layered ones and three-layered are more complex than two-layered ones. A pure trade-off effect would thus show in a decrease in the number of components in an utterance accompanying an increase in the number of layers in the elaborated components. Therefore, a two-component utterance should include more complex elaborations than a four-component utterance. A quick look at Table I in Appendix 9 presenting the results for each individual child shows that this is more or less the case in the children's utterances. This is shown most clearly in the control children whose MLU is less than 6.000. The same tendency is evident in the risk children as well but not as clearly as in the controls. The four-layered elaboration types are present in two-component utterances, three-layered ones have commonly spread into three-component utterances as well, and two-layered elaboration types cover the whole range of utterances. When MLU is greater than 6.000 the tendency is no longer as clear, although it is still noticeable.

Another variety of trade-off effect, within layers, is not clearly present. It would seem logical that the simplest elaboration types within a layer would have a wider range of occurrences and that the most complex components from the same layer would occur only in utterances with fewer components. Signs of this within-layer trade-off effect can occasionally be seen. For example Ronja (ID 6) has [S2+(M)] elaboration in both two- and three-layered utterances and Jaakko (ID 11) has the same elaboration type in three- and four-layered utterances, but when they move a step forward in complexity into [S2+(2M)]-elaboration, Ronja has it only in two-component utterances and Jaakko in three-component ones.

It seems that the risk children perform better than the control children in utterances containing only one elaborated component. The risk children have a wider variety of component types, the children form a coherent group without major individual differences, and they occasionally use more complex elaboration types, especially in the category of four-layered elaborations. The trade-off effect is shown mostly in the control children with lower MLUs, and it is stronger between utterances of distinct sizes than within layers of elaboration. In other words, a four-layered elaboration type is more likely to occur in two-component utterances than in three- or four-component ones, but there is no such clear and straightforward evidence of the same phenomenon being involved when it is question of elaboration types on the same layer but with different degrees of complexity.

9.3.3 Utterances containing two elaborated components

Utterances containing two elaborated components form the other quantitatively important category of children's utterances. In the risk group, there are 212 such utterances (27.6%) and in the control group, 392 utterances (42.9%) in this category. Whether there are qualitative distinctions as well will be examined next. An inventory of combinations of two elaborated components is available in Table II in Appendix 9, categorized by groups and individual children. Table 44 below, gives a summary of the details presented in Appendix 9, supplemented with the numbers of utterances in each elaboration category for both groups.

TABLE 44 Survey of combinations of two elaborated components used by risk and control children in two-, three and four-component utterances (U).

Risk group					
	Types	**Range**	**Median**	**Mean**	**Number of U**
2+2-layered elaboration	3	1–3	2	1.9	126
3+2-layered elaboration	9	2–5	3	3.1	66
3+3-layered elaboration	4	0–3	1	0.8	10
4+2-layered elaboration	6	0–3	1	0.9	10
4+3-layered elaboration	0				
5+2-layered elaboration	0				
5+3-layered elaboration	0				
Total	22	3–12	7	6.7	212
Control group					
	Types	**Range**	**Median**	**Mean**	**Number of U**
2+2-layered elaboration	3	2–3	2	2.2	189
3+2-layered elaboration	9	2–6	4	3.8	160
3+3-layered elaboration	7	0–4	2	1.5	22
4+2-layered elaboration	8	0–3	1	1.2	16
4+3-layered elaboration	1	0–1	0	0.2	2
5+2-layered elaboration	2	0–1	0	0.2	2
5+3-layered elaboration	1	0–1	0	0.1	1
Total	31	4–14	9	8.8	392

The key ratios compiled in Table 44 show that the children in the risk group perform relatively poorly. They use fewer combinations than the children in the control group. Thus, in addition to fewer utterances in this category in general, the risk children also have a narrower inventory of variation in elaborating them.

The inventory of combinations in Appendix 9 provides more evidence of qualitative differences, especially in the more complex utterances of the control group. If we look at the inventory list alone, without the results from individual children, it is easy to see that elaboration combinations usually consist of a more complex and a simpler component. In the majority of cases the simpler component is either an inflected word [M] or two uninflected words [S2], and only occasionally two words with inflection in either or both of them ([S2+(M)], [S2+(2M)]) or a multiply inflected word ([M2]). However, the more complex component varies more widely. Group differences are related to elaboration of both the complex and the simple component. First, in the productions of the control children there are occurrences of combinations of four- and three-layered, five- and two-layered, as well as five- and three-layered components. None of these exist in the productions of the risk children. It is the performance of the control children that creates the complexity scale of elaboration combinations, and they cover the whole scale. The risk children do not attain such deep layers, and within each layer they stick to simpler elaboration types.

The same phenomenon is repeated when the other part of the combination is concerned, although not as clearly as in case of the complex halves. In both groups, and in the majority of combinations, the other part is an inflected word [M] or a combination of two uninflected words [S2]. However, the control children also produce six different combinations where the latter part is a further elaborated [S2]-component, but in the risk group, there are only two such combinations ([S2+(M)]+[S2+(M)] and [S2+(2M)]+[S2+(M)]) and one of them accounts for four out of five occurrences.

In utterances including only one elaborated component the risk group is very homogenous: there were three elaboration types that all children produced in their utterances. In utterances with two elaborated components the situation is reversed. Now it is the control group that shows more uniform performance (in terms of combinations produced by all the children). There are three combinations of two elaborated components ([M]+[M], [S2]+[M], and [S2+(M)]+[M]) that all control children produce, whereas the risk children have no combinations shared by all.

9.3.4 Utterances with three or four elaborated components

The presence of three elaborated components in an utterance is in many ways a dividing line. First, it separates the risk and control groups in a more distinct way. It also draws a dividing line within the groups more clearly than the other combination categories. The inventory of combinations is now smaller that it was with two elaborated components, and a common feature in all of the combinations is that there is always at least one two-layered elaboration type involved. Thus, there are no utterances where all the elaborated components contain a relatively high degree of structural complexity.

The group difference is easily shown in terms of some key ratios presented in Table 45 (the detailed information is shown in Appendix 9, Table III). The control children have on average 3.4 different combinations of three elaborated components in their utterances, the number of combinations varying from zero (ID 6 Ronja) up to eight (ID 2 Riikka). In the risk group the corresponding figures are 2.0, and from zero to four. Table III in Appendix 9 shows that there are three advanced control children, Tuomo (ID 19), Tuija (ID 1), and Riika (ID 2), who are mainly responsible for the group difference, and who at the same time split the control group in two. These children have from seven to eight combinations which is substantially more than any of the others have. Previously, the number of different elaboration combinations occurring in utterances has grown smoothly within a group, but here there is a clear subgroup of children which differs from the others.

The group difference is not only quantitative. In fact, the qualitative difference is more important when the focus is on structural complexity. There are seven control children who produce combinations of 3+3+2 layers, 4+2+2 layers, and 4+3+2 layers. In the risk group, only Aleksi (ID 31) and Janna (ID 27) do this. Aleksi even has a combination of 4+4+2. All the other risk children have combinations of 2+2+2 layers or 3+2+2 layers only. Thus, the level of complexity in utterances in the risk group is lower than in the control group.

TABLE 45 Survey of combinations of three elaborated components used by risk and control children in two-, three- and four-component utterances (U).

Risk group					
	Types	**Range**	**Median**	**Mean**	**Number of U**
2+2+2-layered elaboration	2	0–2	1	1.2	23
3+2+2-layered elaboration	3	0–1	0	0.5	6
3+3+2-layered elaboration	2	0–1	0	0.2	3
4+2+2-layered elaboration	0				0
4+3+2-layered elaboration	1	0–1	0	0.1	1
4+4+2-layered elaboration	1	0–1	0	0.1	1
Total	9	0–4	2	2.0	34
Control group					
	Types	**Range**	**Median**	**Mean**	**Number of U**
2+2+2-layered elaboration	3	0–3	1	1.2	31
3+2+2-layered elaboration	4	0–3	1	1.1	20
3+3+2-layered elaboration	6	0–3	0	0.8	11
4+2+2-layered elaboration	2	0–1	0	0.2	2
4+3+2-layered elaboration	3	0–1	0	0.2	3
4+4+2-layered elaboration	0				
Total	18	0–8	3	3.4	67

It was proposed previously that the use of three elaborated components in utterances forms a category that makes divisions within the groups as well as between the groups. In the control group there is a subgroup of three children with a larger variety of combinations. Qualitatively, a division into subgroups might result in three groups, according to the layered structure in these combinations. First there is a subgroup of five children (ID 18 Saku, ID 17 Tatu, ID 13 Joel, ID 8 Aino, and ID 10 Santeri) who have 2+2+2-layered and 3+2+2-layered combinations in their utterances. The next subgroup is formed by children who reach the level of 3+3+2-layered combinations (ID 15 Mika, ID 11 Jaakko, ID 2 Riikka, and ID 7 Taru). Finally, there are three children, Tuomo (ID 19), Tuija (ID 1), and Laura (ID 3), who have expanded their inventory of combinations to 4+3+2-layered ones. In the risk group there is no clear quantitative division into subgroups, but a qualitative division is clear. Aleksi (ID 31) and Janna (ID 27) have an inventory of combinations stretching across the whole range of combinations, while the other children produce 2+2+2-layered and 3+2+2-layered combinations only.

Combinations of four elaborated components (see Table VI in Appendix 9) are very rare. The risk children do not produce such combinations at all. Only three different combination types appear and all of them occur in the utterances of three control children: Ronja (ID 6), Laura (ID 3) and Taru (ID 7). Interestingly, Ronja produced no combinations of three elaborated components, while Taru had only one. However, they both belong to the small group which produces utterances with four elaborated components. This indicates that growth in f structural complexity is not straightforward. Signs of this were already evident in the general component inventory (Section 8.2.6). Some structures that could have been placed between two other structures in structural complexity were totally lacking in the data. Although, linguistic structures can be ranked in an order of complexity, this is not the same order in which structures are acquired. Moreover, absolute structural complexity is not necessarily experienced as complex by a language acquirer or user. (Kauppinen 1998, 149–150; Kusters & Muysken 2001.) In Dahl's (2004: 39–40) terms, complexity and difficulty should be kept separate.

9.4 The possible contribution of excluded children in the comparison of risk and control children

Originally this study involved 40 subjects, 20 risk children and 20 control children. According to the combined MLU and IPSyn results three subgroups including altogether 24 children were selected to represent points were MLU and IPSyn showed contradictory results. At that point seven control children and nine risk children were excluded and Utterance Analysis was not conducted on their data. The impact of these children's data on the comparison between the risk and control groups can thus be only a matter of speculation. These speculations are based on their MLUs and IPSyn scores as well as the Utterance Analysis results from those children in subgroups A, B and C with matching MLU and IPSyn results. The groups for comparison are presented in Table 46.

TABLE 46 The children excluded from Utterance Analysis, matched with those who are closest to them in MLU and IPSyn results in subgroups A, B and C. Children are compared within the risk and control group.

Risk group								
Included in UA					**Excluded from UA**			
ID	**Name**	**MLU**	**IPSyn**		**ID**	**Name**	**MLU**	**IPSyn**
35	Seppo	3.688	60		26	Tiina	1.587	26
22	Anna	3.888	42		25	Siiri	1.650	30
36	Sampo	3.987	57		40	Risto	2.329	30
					32	Juho	2.450	41
					33	Jukka	2.775	42
					37	Henri	3.638	39
23	Anniina	6.050	72		21	Paula	5.400	77
27	Janna	6.300	69		39	Juuso	6.963	83
30	Elina	6.325	74		28	Jenni	7.862	86
Control group								
Included in UA					**Excluded from UA**			
ID	**Name**	**MLU**	**IPSyn**		**ID**	**Name**	**MLU**	**IPSyn**
18	Saku	3.962	53		5	Liisa	1.233	10
17	Tatu	4.063	56		12	Lauri	2.614	47
6	Ronja	4.100	51		9	Lassi	3.175	51
11	Jaakko	6.112	62		14	Patrik	5.50	78
19	Tuomo	6.237	75		20	Tuomas	5.862	76
1	Tuija	6.250	84		16	Kyösti	6.70	73
3	Laura	6.250	76		4	Henna	7.162	86
2	Riikka	6.287	68					
7	Taru	6.350	70					

In the risk group the first group of children excluded (Tiina, Siiri, Risto, Juho, Jukka, and Henri) are on a lower MLU level than those included in the Utterance Analysis (Seppo, Anna and Sampo). Only Henri is very close in this respect to the children compared. An MLU beneath 2.000 (Tiina and Siiri) implies that there is not much scope for elaboration or multi-componential utterances. These children would not increase the number of elaboration types and elaboration combinations although they may occasionally use an inflected word, that is, a two-layered elaboration. Their contribution may also be very small in terms of new elaboration types, because many of their utterances consist of one component only and these are excluded from the analysis. Children with an MLU between 2.000 and 3.000 (Risto, Juho and Jukka) do not have much chance to use elaboration either. However, Juho's and Jukka's IPSyn scores reveal that they already use morphosyntactic devices in their utterances at the same level as Anna. In the previous comparison between risk and control children, the risk children were found to be a homogenous group in utterances with only one elaborated component. If these six children were added to the comparison, they would probably break the homogeneity and widen the range in number of elaboration types used. They would also bring down the group average. Their contribution in elaboration combinations would also show as widening ranges and a decrease in average values.

The next pair of subgroups to be compared also belongs to the risk group. Anniina, Janna and Elina represented the best performing children in the risk group, when their use of elaborated components was evaluated. Paula, Juuso and Jenni have higher IPsyn scores and Juuso and Jenni also higher MLUs than the risk children included in the Utterance Analysis. It is reasonable to think that Paula, Juuso and Jenni would reach at least the same level of elaboration as Anniina, Janna end Elina, and strengthen the group results especially in respect of the use of elaboration combinations. They may also use some types of elaboration combinations that are not yet used by other risk group children and might thus increase the number of different types. There is also a possibility that they might increase the group averages, but the combined effect with the six children discussed earlier could cause this impact to vanish. However, it can be assumed that if all these nine children were involved in the group comparison as well the number of elaboration types and elaboration combinations used would have been greater, although the ranges in the results might also have been larger. The average values for the group would probably stay at the same level as they are now, because the two extremes would balance each other out.

In the control group Saku, Tatu and Ronja did not deviate from the other controls when utterances with only one elaborated component were concerned. However, in elaboration combinations they all performed clearly more poorly than the other controls. These three children form the comparison group for Liisa, Lassi and Lauri, who are all clearly behind Saku, Tatu and Ronja in MLU, and Liisa also in IPSyn scores. According to her MLU and IPSyn results Liisa has very few if any elaborations and if she has they probably occur in single-component utterances which are excluded from the analysis. Her contribution would be to bring down the group results in all possible respects. Lauri's and Lassi's IPSyn score imply elaboration, but their MLUs suggest that elaboration is not yet multilayered and that elaborated component combinations are either very rare or do not exist at all. Lauri's and Lassi's contribution to the control group is likely to be in utterances with only one elaborated component and mostly in two-layered elaboration types.

To balance Liisa's, Lauri's and Lassi's contribution there are four more children in the risk group, Henna, Patrik, Kyösti and Tuomas, who have high results in both MLU and IPSyn. It is likely that they would show their abilities especially in combining elaborated components and even bring some new types and combinations into the inventory. This assumption is based on the performance of Jaakko, Tuomo, Tuija, Laura, Riikka and Taru who are mostly responsible for the good results in elaborated combinations for the control group. Thus, Henna, Patrik, Kyösti and Tuomas would strengthen the performance of the control group in combinations, and the difference between the risk and control groups would probably increase.

9.5 Summary

In this chapter I have returned to the original division into groups of subjects to search for possible differences between children with a high genetic risk of dyslexia and their controls. Previously a comparison has been made using MLU and IPSyn, and the groups consisted of twenty children each. Both scales showed small differences between the groups in favour of the control group. In addition to these small group differences, it was noticed that there is huge individual variation within both groups.

In this chapter a comparison between risk and control children was made by concentrating on the Utterance Analysis conducted on the two-, three- and four-component utterances produced by thirteen control children and eleven risk children. On the MLU and IPSyn scales, these subgroups were found to be more equal than the original risk and control groups consisting of twenty children each. This makes the comparison even more interesting, since the similarity in terms of the available quantitative measures shows that to this extent the groups are very similar. Therefore any differences in structural complexity are not just a side effect of differences in utterance length or in inventory of structures. When other variables are held constant the focus can be entirely on structural complexity.

The two main results of the survey emphasize on one hand similarities between the groups but differences on the other hand. First, the risk and control children do not differ from each other when the overall inventory of elaborated components is considered. In both groups the inventory is of equal extent and it covers the same range on the complexity scale. However, the inventory of individual component types is not the most important issue in this study. Rather, it is the use of components. The emphasis on component use leads to the other main result of the survey. In spite of having similar inventories of component types, the risk and control children differ remarkably in the ways they combine elaborated components within utterances. The risk children seem to focus on utterances including only one elaborated component. They seem to be more advanced than the control children in this utterance category. However, the situation clearly changes in utterances including two or more elaborated components. Here it is the control children that show more advanced performance. They have more different combination types and it is the members of the control group that are responsible for the most complex combinations of elaborated components. In addition, the control group seems to be more homogenous as a group.

There are at least three different views of how syntactic abilities and dyslexia are related to each other. According to the first view, children's linguistic skills are delayed or they have structural deficiencies in their language system, and either of these can cause the lack of basic syntactic abilities. A syntactic deficiency is seen as characteristic of dyslectic people. In this study, no evidence of lack of basic syntactic abilities was found. Instead, the risk children performed extremely well in the case of utterances including only a single elaborated component combined with unelaborated ones. Moreover, they had as large, as versatile and as complex a variety of syntactic component types as the control children. Therefore, the abilities and skills seem to be on the same level in the both groups.

The second view argues that syntactic deficiency is not characteristic of dyslexia. This view emphasizes the role of a phonological deficit in dyslexia, and sees syntactic problems as a consequence of limitations in short term memory caused by a basic difficulty in generating phonological codes. Dyslectic people produce syntactically accurate utterances within the limits of their working memory. (For a review, see Leikin & Assayag-Bouskila 2004.) As a matter of fact, in this study children in both groups produce morphosyntactically accurate utterances but risk children combine elaborated components less than control children. The preference for fewer elaborations could be explained in terms of limitations in short term memory but attesting this assumption would need further investigation.

The third view agrees with the previous one in that it does not assume any absence of basic syntactic skills. According to this view, syntactic inferiority in dyslexic children can be explained by poor ability to use their syntactic knowledge proficiently, and the problem is metalinguistic by nature. It is argued that the ability to apply knowledge does not significantly affect speech but may become more evident in less natural linguistic activities like reading. (Deutch & Bentin 1996.) In the present study MLU and IPSyn did not reveal major differences between the groups, and this supports the idea that speech is more or less intact in dyslectic children. Only a specific examination might reveal that utterance formulation is affected. A careful investigation of every syntactic component and especially the combinations of elaborated components in utterances showed that it is the complexity of whole utterances that separates the groups, not individual components as resources. IPSyn does not reflect this kind of difference, since it is basically an inventory of resources. Why MLU does not reflect these differences either is explained by the fact that MLU is an average number of morphemes, and even though less elaborated components usually have fewer morphemes, they do not bulk sufficiently large to result in a clear difference in MLU value. Moreover, utterances with only one elaborated component may function as a balancing force. Risk children performed better in this category, and they probably compensate with these utterances including one elaborated component for their poorer performance in utterances with several elaborated components. Therefore, the difference between risk and control children remains very small when measured by MLU and IPSyn.

10 CONCLUSIONS

In the present study structural complexity, the acquisition of Finnish morphosyntax and genetic risk of dyslexia were brought together. Although they are in themselves quite distinct fields of enquiry, there are certain connecting factors that strongly justify the way they are linked in this study. In the field of dyslexia studies there is a growing interest in finding evidence of precursors of dyslexia in linguistic areas other than phonology. Syntactic abilities are a strong candidate for this, since during the last decade several studies have suggested that dyslectic people differ from others especially in their processing of syntactic structures. In these studies the differences have been related especially to syntactic complexity, although the concept of complexity has been left without any further specification. In cases of dyslexia it is very important to be able to make the diagnosis as early as possible and for this reason the manifestation of a possible deficiency in syntactic abilities has been studied during the early phases of language acquisition. Against this background the spontaneous speech of children with a high genetic risk of dyslexia has been studied here with special attention to the morphosyntactically complex structures of Finnish. However, in addition to pursuing this original purpose the present study is at the same time a survey of how Finnish children advance in their use of morphosyntax. The concept of complexity introduces methodological factors into the picture and raises the question of what kind of complexity is highlighted in each of the methods used in this study.

Possible differences between the risk and control children were sought out using three different methods which all reflect an absolute approach to complexity, although each of them highlights different facets of structural complexity. Mean Length of Utterance (MLU) and the Index of Productive Syntax (IPSyn) are traditional methods used in language acquisition studies. Both methods indicated some differences between the risk and control children. In neither case was it a question of major differences. The children in both groups performed at the same level but in the risk group there were more children at the lower ends of the scales and thus kept the group average at a somewhat lower level than that of the control group. Neither scale of measurements could highlight any specific area of morphosyntax which might be responsible for the differences, although in the verb section of IPSyn the difference was slightly clearer than in other areas.

The third method used in looking for possible differences between the risk and control children was Utterance Analysis. It was created specifically for the present study and is in line with the view of multidimensional structural complexity introduced in the present study as well. Utterance Analysis divides every utterance into its syntactic components, and also specifies each elaboration within a component. Using Utterance Analysis two major findings were made concerning the risk and control groups. First, both risk and control children have similar basic skills and resources for morphosyntax. Both groups performed equally well in the elaboration type inventory, indicating that children in both groups have an equal number of morphosyntactic component structures to use in utterances, and they are at the same level of complexity. The second major finding, however, differed remarkably from the first. It appeared that the vast majority of utterances that the risk children produce include only a single elaborated component, whereas the control children combine several elaborated components in their utterances. Thus, despite the similar resources, the control children used them much more effectively.

In a sense, this is good news for dyslectic people. It strongly suggests that there is no particular syntactic deficit affecting their basic syntactic skills. The risk children produced utterances that are as accurate as those produced by the control children, and they have syntactic resources as large and complex as the controls have. Producing fewer elaborated components in their utterances does not prevent them from effective communication.

In terms of language acquisition, particularly the acquisition of morphosyntax, the present study highlights two properties of development. One is similarity and the other is spreading. Subgroups A, B and C were selected specifically to test the methods and to shed light on how the development of morphosyntactic complexity proceeds. The results from MLU and IPSyn suggested that the three subgroups differ from each other remarkably. However, Utterance Analysis did not confirm a view of large differences between the groups. Rather, it highlighted similarities. It turned out that children in each of the groups produced similar proportions of noun phrases (approximately 42% of components in each group), verb constructions (approximately 30%), adverb phrases (approximately 19%) and other syntactic components. They produced similar utterance patterns in terms of syntactic components. For example, approximately 80 % of all two-component utterances in each group followed the patterns of V + NP, V + ADVP and NP + ADVP. Moreover, when the number of syntactic components in utterances was examined, it turned out that two-, three- and four-component utterances account for almost 90% of the data and that three-component utterances predominate in every group with proportions ranging from 42.4% in group A to 50.3% in group C. In addition to similarities in the basic structure of utterances it was found that the three groups resemble each other more and more closely as the number of components in utterances grows: four-component utterances have a similar layer structure as well as similar syntactic components. In fact the differences between the groups are found in the two-component utterances. The results seem to support a trade-off effect: overall structural complexity is balanced within the utterances.

The similarity in the basic structures was surprising because MLU and IPSyn had given a very different picture of the situation. However, the situation in which the data was recorded may have affected its quality. In a play session where new toys are introduced to a child there is a lot of occasion to use similar utterances for naming, expressing locality and so forth. These expressions very often happen to be three-component utterances which contain noun phrases, verb constructions and adverbs. Moreover, they appear in a child's expressive language very early. This may have affected the results. However, the discovery of so many similarities also suggested looking at the growth of structural complexity from other directions as well. If utterances are not expanded by means of syntactic components, then the components themselves must expand.

Developmental growth in structural complexity actually spreads in several directions. The concept of complexity adopted in this study is a multidimensional view where linguistic units consist of components and layers, and complexity may grow either horizontally by adding new components, or vertically by adding elaboration to components. Vertical expansion may occur in two ways: either more components are elaborated or elaboration goes deeper. The comparison of groups A, B and C showed that elaboration mostly begins with morphological devices. Simple utterances are short and only one component is elaborated, usually very moderately. Next elaboration expands to a second component in an utterance, and only after that are the components elaborated more deeply. Children in group A represent the first phase, group B the second phase and group C the last phase. However, this does not happen in all kinds of utterances at the same time. As noted earlier, four-component utterances are very alike in all the groups. The expansion of elaboration starts with shorter utterances. Thus slow expansion and spreading is simultaneously in progress within components, individual utterances and utterance categories.

Since there is progress going on simultaneously in several directions, growth in complexity is not a straightforward process which is easily illustrated in a linear fashion. This is true for the complexity relations of individual components as well. In specifying the complexity of individual elaboration types, a net-like model was adopted. It describes the relations of an elaboration type to its nearest "relatives", but it does not force all components into an absolute order of complexity, nor is it necessary to determine the complexity relations of two distinct elaboration devices (such as inflection and syntactic expansion). The way children bring these component types into use in their language indicates non-straightforward progress as well. Children do not necessarily proceed in straight line within a component family. This is consistent with views claiming that the order of acquisition does not follow the order of linguistic complexity. More likely it is relative complexity that coincides with acquisition order.

Returning to the individual components in utterances, the expansion of elaboration and the morphological and syntactic devices used in expansion do not exclude the views that learning is item-based. In fact, the similarities found in the structures as well as the non-straightforward progress in component types in the complexity net in a way support it. Similarities in basic structures may imply that larger units that are acquired as entire phrases are in constant use in all groups. The same functions are repeated in each and every play session in the data, and thus the same expressions are used over and over again, although the more advanced children can already modify the basic structures in several ways. They have become frame structures which probably include several slots which can be filled in. Rather than the learning of a new abstract structure, the absence of linear progress through the complexity net can suggest the acquisition of a new (possibly complex) expression which happens to contain such a structure.

Growth in structural complexity has now been described as spreading in several directions. How do the methods of analysis used in the present study reveal spreading? Obviously Utterance Analysis reflects both the different dimensions and simultaneous progress. The other two methods are, however, designed to focus only on one aspect. MLU grows when enough utterances grow longer. It is not sensitive to one or two advanced utterances since it is a compromise over the whole data set. It cannot pinpoint the actual point where progress occurs either, because it is programmed to register only those units that qualify as morphemes.

The Index of Productive Syntax is sensitive to resources. Growth of complexity in terms of IPSyn reflects a new building block ready to be used in utterances. IPSyn registers each of the building blocks, but it does not give information about how these blocks are used in utterances or combined with other blocks. This is why IPSyn did not reveal the difference between the risk and control children. It does not go beyond the level of resources.

Although MLU and IPSyn are both inadequate as measures of the multidimensional properties of structural complexity, they do help us see the general outlines when they are used together. In the present study five hypotheses were made on the basis of combined results from MLU and IPSyn concerning the general lines of development and growth in complexity. The analysis conducted using Utterance Analysis proved these hypotheses to be right in general terms.

In the present study an absolute approach to complexity was adapted. All the methods used followed this approach and concentrated on the language produced rather than the language user. However, each of the methods gave a different impression of the data, either from a one-dimensional perspective or a more multidimensional one. In the case of a relative approach to complexity it has been emphasized that the conditions under which complexity is examined must be unambiguously determined in terms of for example the language user and the language used. In my opinion, in the absolute approach to complexity the conditions should be made equally clear. This study shows that "absolute" complexity is as relative as "relative" complexity. Therefore, careful attention must be paid to specifying the underlying concepts whenever complexity is discussed.

TIIVISTELMÄ

Tämä tutkimus käsittelee suomalaislasten spontaanien puhetuotosten morfosyntaktista kompleksisuutta kielenomaksumisprosessin alkuvaiheessa. Tutkimuksessa rakenteellista kompleksisuutta valotetaan morfosyntaktisen kehityksen, kehityksellisen dysleksian ja kielen kompleksisuuden arviointiin yleisesti käytettyjen menetelmien näkökulmasta.

Kielenomaksumistutkimuksessa kompleksisuus on hyvin keskeisellä sijalla, onhan kehityksen usein ajateltu tarkoittavan systeemin laajenemista, monimutkaistumista tai hienojakoistumista. Morfosyntaktinen kehityskin voitaisiin siis yksinkertaistaen määritellä morfologisten keinojen käytön lisääntymiseksi, syntaktisten kuvioiden laajentumiseksi ja morfologian ja syntaksi yhteistyön monipuolistumiseksi. Tätä prosessia on pyritty havainnoimaan erilaisilla menetelmillä, joista ilmausten keskipituus (MLU) lienee kaikkein tunnetuin ja käytetyin. Erilaiset mittarit kuitenkin piilottavat lapsen puheen keskiarvojen tai pistemäärien taakse ja näin yhtymäkohta itse kieleen hämärtyy tai katoaa kokonaan. Tulokset ovat siis numeroita, joiden informaatioarvo sellaisenaan on sangen pieni. Kehityksellinen dysleksia puolestaan yhdistyy rakenteelliseen kompleksisuuteen juuri syntaksin kautta. Jo pitkään on dysleksiaan yhdistetty fonologisen prosessoinnin ongelmat mutta myöhempi tutkimus on paljastanut, että myös kielen syntaksissa on eroja dyslektikkojen ja heidän verrokkiensa välillä. Merkitsevät erot on löydetty juuri spontaanin puheen syntaktisesta kompleksisuudesta, ja niiden toteamisessa on menetelminä käytetty muun muassa ilmausten keskipituutta ja produktiivisen syntaksin indeksiä (IPSyn). Kuitenkaan nämä tulokset eivät kerro sen tarkemmin, mitkä syntaktiset piirteet ryhmiä erottavat.

Siitä huolimatta, että kielenomaksumistutkimuksessa kompleksisuus on hyvin usein mukana joko eksplisiittisesti tai implisiittisesti, se on käsitteenä saanut kovin vähän huomiota osakseen ja se on jätetty määrittelemättä ikään kuin itsestäänselvyytenä. Yksi tämän tutkimuksen tavoitteista olikin pohtia ensin rakenteellisen kompleksisuuden olemusta ja vasta sitten selvittää, miten tämän näkemyksen mukainen kompleksisuus toteutuu lasten spontaanissa puheessa. Tässä tutkimuksessa lähtökohdaksi otetaan absoluuttinen kompleksisuus, tässä tapauksessa kielen morfosyntaksin muodostama rakenteellinen systeemi. Suhteellinen kompleksisuus eli kielenkäyttäjän kokemus kielen vaikeudesta jätetään siis kokonaan tutkimuksen ulkopuolelle.

Absoluuttiseen kompleksisuuteen keskittyminen edellyttää sen pohtimista, mikä tekee yhdestä kielellisestä rakenteesta kompleksimman kuin toinen. Tässä tutkimuksessa absoluuttista kompleksisuutta katsotaan moniulotteisena ominaisuutena, joka koostuu sekä horisontaalisesta että vertikaalisesta dimensiosta (komponenttien sisäinen rakenne). Horisontaalinen ulottuvuus tarkoittaa ilmauksen sisältämiä komponentteja (esim. substantiivilauseke, verbirakenne, adpositiolauseke, interjektio), joita katsotaan keskenään samanarvoisina kokonaisuuksina. Keskustelukielen ilmaus ei noudattele läheskään aina täydellisen lau-

seen rakennetta, joten lauserakenteen hierarkinen kuvaus on tässä hylätty. Vertikaalinen dimensio puolestaan tarkoittaa komponenttien sisäistä koostumusta. Se kuvataan kerroksellisena rakenteena, joka saadaan aikaan morfologisin keinoin taivuttamalla sanoja ja syntaktisin keinoin erilaisilla täydennyksillä ja laajennuksilla. Ilmauksen morfosyntaktista kompleksisuutta on siis mahdollisuus kasvattaa sekä lisäämällä ilmauksen komponenttimäärää että elaboroimalla komponentteja. Tällä lähestymistavalla halutaan tietoisesti välttää kielen liiallinen yksinkertaistus ja lineaaristaminen. Samalla kuitenkin rakenteiden kompleksisuusjärjestyksen määrittäminen käy hankalaksi, koska eri dimensioiden ja eri elaborointikeinojen välinen vertailu ei ole mahdollista. Moniulotteisen tarkastelun avulla saadaan rakenneltua pikemminkin morfosyntaktisten rakenteiden verkosto, jossa suoraviivaista kompleksisuusjärjestystä kaikkien rakenteiden kesken ei ole edes tarpeen luoda.

Tutkimuksen aineisto tallennettiin vapaassa leikkitilanteessa lasten kotona, kun koehenkilöt olivat kahden vuoden ja kuuden kuukauden ikäisiä (vaihteluväli 2;5.20–2;6.20). Tutkimuksessa oli mukana kaikkiaan 40 keskisuomalaista lasta, joista 20:llä on geneettinen dysleksiariski. Aineisto litteroitiin ja koodattiin morfologisesti käyttäen CHILDES-systeemiä, ja lopullista analysointia varten seulottiin jokaiselta lapselta 80 eniten morfeemeja sisältänyttä ilmausta. Tällä pyrittiin varmistamaan se, että kaikilta koehenkilöiltä saatiin mahdollisimman paljon sellaista kielenainesta, jossa morfosyntaktiset piirteet tulivat esille. Aineiston analyysimenetelminä hyödynnettiin ilmausten keskipituutta (Mean Length of Utterance, MLU), produktiivisen syntaksin indeksin suomalaisversiota (Index of Productive Syntax, IPSyn) sekä tätä tutkimusta varten kehiteltyä ilmausanalyysia (Utterance Analysis, UA).

Tutkimuksessa pyrittiin vastaamaan kolmeen kysymykseen:

1. Kuinka morfosyntaktinen kompleksisuus rakentuu lasten spontaanissa puheessa?
2. Eroavatko dysleksiariskilapset ja heidän verrokkinsa toisistaan, kun tarkastelun kohteeksi otetaan morfosyntaktisten rakenteiden kompleksisuus?
3. Miten tavanomaiset kompleksisuusmittarit kuten MLU ja IPSyn kuvastavat rakenteellista kompleksisuutta?

Morfosyntaktisen kompleksisuuden rakentumisen selvittelemisessä riski- ja verrokkiryhmäjakoa ei otettu lainkaan huomioon. Lasten saamia MLU-arvoja ja IPSyn-pisteitä verrattiin toisiinsa. Kummankin mittarin tuloksissa vaihteluväli oli hyvin suuri (MLU 1,233–7,862; IPSyn 10–86) eli lasten tuotosten morfeemimäärät vaihtelivat kovasti samoin kuin se, kuinka paljon erilaisia rakenteita ja muotoja lapset tuottivat. Tämän tarkemmin numerot eivät kuitenkaan pysty lasten tuotosten kieltä luonnehtimaan. MLU-arvot ja IPSyn-pisteet korreloivat keskenään vahvasti ($r = .92$), mikä on havaittu myös aiemmissa tutkimuksissa. Vahvasta positiivisesta korrelaatiosta huolimatta tulosten välillä oli ristiriitaisuuksia: toinen mittari osoitti tuotosten olevan kompleksisuudeltaan hyvin samanlaisia kun taas toinen mittari tunnisti niissä isojakin eroja. Lapsilla, joiden MLU-arvot olivat 3,688–4,388, IPSyn-pisteet vaihtelivat 42 ja 68 välillä. Samalla tavoin keskenään lähes samanlaisia MLU-arvoja 6,050–6,350 vastasi IPSyn-pisteiden vaihteluväli 62–84 pistettä. Näitä kahta ristiriitaryhmää kompensoi kolmas, jossa mittareiden osat olivat vaihtuneet: rajusti vaihtelevien MLU-arvojen (3,688–6,350) parina oli 60 ja 70 pisteen vaihteluvälillä pysyttelevät IP-Syn-pisteet. Näihin kolmeen ristiriitakohtaan vertailussa sijoittuneista lapsista muodostettiin alaryhmät A (MLU 3,688–4,388; IPSyn 42–68), B (MLU 3,688–6,350; IPSyn 60–70) ja C (MLU 6,050–6,350; IPSyn 62–84), joiden tuotokset analysoitiin ilmausanalyysin avulla. MLU:n ja IPSYnin tulosten ja erityisesti niiden välisten ristiriitaisuuksien pohjalta muodostettiin viisi hypoteesia siitä, miten morfosyntaktinen kompleksisuus kehittyy lasten tuotoksissa:

Hypoteesi 1
Pieni morfeemimäärä yhdistettynä vähäisiin rakenneresursseihin (osa ryhmän A lapsista) reaalistuu ilmauksina, joissa elaborointia on vain vähän ja samoja elaborointikeinoja käytetään ilmauksesta toiseen. Rakenteiden kerroksellisuus toteutuu vasta hyvin maltillisesti. Syntaktisia komponentteja on vain vähän.

Hypoteesi 2
Pieni morfeemimäärä yhdistettynä kohtalaiseen määrään rakenneresursseja (osa ryhmän A lapsista) reaalistuu ilmauksina, joissa käytetään vain vähän elaborointia, mutta keinot ilmauksesta toiseen vaihtelevat. Elaborointikeinoja ei yhdistellä samassa ilmauksessa, ja syntaktisia komponentteja on edelleen vain vähän.

Hypoteesi 3
Kun suuri morfeemimäärä ja runsaat rakenneresurssit yhdistyvät (ryhmä C) syntyy lauseita, joissa on pitkälle elaboroituja komponentteja. Morfosyntaktinen kompleksisuus on kasvanut sekä horisontaalisesti että vertikaalisesti, ja omaksuttuja resursseja yhdistellään samassa ilmauksessa.

Hypoteesi 4
Runsaasti vaihteleva morfeemimäärä ilmauksessa yhdessä vakiintuneen resurssimäärän kanssa (ryhmä B) implikoi, että rakenteellinen kompleksisuus kasvaa erityisesti yhdistelemällä yhä enemmän morfosyntaktisia resursseja samassa ilmauksessa.

Hypoteesi 5
Rakenteellinen kompleksisuus kasvaa asteittain, ja se näkyy vuorotellen morfeemimäärän kasvuna (MLU) ilmauksessa ja morfosyntaktisten resurssien lisäänymisenä (IPSyn). Ensin opitaan ilmauksen peruselaborointitavat. Tämä näkyy lyhyinä ilmauksina ja vaihtelevana rakenneresurssivarastona. Seuraavaksi on näitä resursseja opittava käyttämään tehokkaasti. Tällöin resursseja yhdistellään ja kerroksellisuus komponenteissa kasvaa, mikä näkyy isompana ilmausten keskipituutena. Seuraavaksi on vuorossa taas uusien rakenteiden omaksuminen. Ilmausten pituus ei välttämättä kasva, vaikka resursseja tuleekin lisää ja ilmausten kompleksisuus kasvaa sekä vertikaalisesti että horisontaalisesti.

Hypoteesien paikkansapitävyyttä selvitettiin ilmausanalyysin avulla, ja siinä keskityttiin syntaktisten komponenttien määrään, niiden komponenttien elaborointiin sekä erilaisten komponenttien yhdistelemiseen ilmauksessa.

Syntaktisten komponenttien määrässä kaikki kolme ryhmää osoittautuivat melko samanlaisiksi. Kaikissa ryhmissä valtaosa ilmauksista oli 2-, 3- ja 4-komponenttisia, A-ryhmässä painottuen 2- ja 3-komponenttisiin ja vastaavasti C-ryhmässä 3- ja 4-komponenttisiin. Syntaktisten komponenttien määrä lisääntyy siis hyvin maltillisesti. Laadullisesti tarkasteltuna eri ryhmissä tuotetut ilmaukset edustavat samoja komponenttityyppejä ja -yhdistelmiä: noin 90 % kaikista komponenteista edustaa nominilausekkeita, verbirakenteita ja adverbilausekkeita. Koska komponenttien määrissä ja tyypeissä ei näyttäisi olevan suuria eroja ryhmien A, B ja C välillä, näyttäisi siltä, että ilmausten perusrakenne luodaan jo varhain. Komponenttikombinaatiot luovat siis pohjan, jolle elaboroinnin myötään tuoma morfosyntaktinen kompleksisuus voi kasvaa. Rakenteellisen kompleksisuuden kannalta on tärkeää, millä tavoin kunkin komponentin elaborointipotentiaalia käytetään hyödyksi. Tässä erityisesti nominilausekkeet ja verbirakenteet nousevat tärkeiksi, koska niissä on runsaasti mahdollisuuksia sekä syntaktiseen että morfologiseen elaborointiin.

Seuraavaksi ilmausanalyysista seulottiin, kuinka monikerroksisiksi komponentteja oli elaboroitu ja miten kerroksellisuutta yhdisteltiin ilmauksissa. Kartoituksen kohteena olivat vain 2-, 3- ja 4-kompnenttiset ilmaukset, jotka yksinään kattoivat n. 90 % kaikista ilmauksista niin A-, B- kuin C-ryhmässäkin. Hypoteeseissa oletettiin, että A-ryhmän tuotoksissa elaborointia on lähinnä vain yhdessä komponentissa ilmausta kohden eikä se vielä ulotu kovin moneen kerrokseen. Tulokset tukevat tätä hypoteesia, sillä yli puolet A-ryhmän lasten ilmauksista on sellaisia, joissa pisimmälle elaboroitu komponentti on 2-kerroksinen ja suurimmassa osassa 2- ja 3-komponenttisia ilmauksia vain yksi komponenteista on elaboroitu. B-ryhmässä elaboroiminen lisääntyi, aivan kuten Hypoteesissa 4 ennustettiin. Elaboroimisen lisääminen tapahtui ennen muuta siten, että yhä useampia ilmauksen komponentteja elaboroitiin. C-ryhmässä erityisesti 2-komponenttiset ilmaukset antoivat todisteita yhä syvemmälle menevästä elaboroinnista. Kerroksellisuuden analysointi toi esille todisteita myös rakenteellisista kompromisseista ja vaihtokaupoista: mitä enemmän ilmauksissa on komponentteja, sitä yksinkertaisemmin niitä elaboroidaan. Samoin tulokset osoittavat, että MLUn ja IPSynin perusteella toisistaan poikkeaviksi todetut ryhmät A, B ja C osoittavat sitä enemmän samanlaisia piirteitä mitä enemmän ilmauksissa on komponentteja. Tämä näkyy pisimmälle elaboroiduissa komponenteissa, erilaisten komponenttien yhdistelemisessä ja vielä erityyppisten komponenttien elaboroinnissa.

Yhteisten piirteiden löytyminen ei välttämättä johdu pelkästään samanlaisista resursseista tai taidoista. Osasyynä tähän on todennäköisesti myös tilanne, jossa aineisto on tallennettu. Tilanteen samanlaisuus jokaisen lapsen kanssa heijastunee myös tilanteessa käytettyyn kieleen. Keskeisiä ilmaistavia asioita olivat esimerkiksi lelujen nimet (*Tää on x.*) sekä eksistentiaalisuus ja lokaatio. Näihin käytetyt kielelliset keinot ja ilmausrakenteet ovat samanlaisia jokaisessa ryhmässä, ja se saattaa selittää, miksi 3-komponenttiset ilmaukset ovat niin suosittuja ja miksi kaikissa ryhmissä käytetään NP + V + ADVP -ilmausrakennetta. Spontaani kieli on aina kontekstinsa ja funktionsa värittämää.

Pelkkää kerroksellisuutta pidemmälle menevä syntaktisten komponenttien koostumuksen analyysi pureutui niihin keinoihin, joilla komponentteja oli elaboroitu. Jälleen kerran ensimmäisenä esille tuli ryhmien välinen samanlaisuus: ryhmät eivät juuri eronneet toisistaan, kun tarkasteltiin kunkin ryhmän erilaisten elaborointityyppien inventaaria. Vain 4-ja 5-kerroksissa elaborointityypeissä A-ryhmä jäi B- ja C-ryhmistä jälkeen. Elaborointityypien käyttö sen sijaan osoitti selvemmän eron, sillä ne olivat B- ja C-ryhmällä laajemmassa käytössä kuin A-ryhmällä. Toinen mielenkiintoinen tulos oli se, että A-ryhmä käytti morfologista elaborointia tasavertaisesti B- ja C-ryhmän tapaan, mutta morfologisten ja syntaktisten keinojen yhdistämiseen samassa elaborointityypissä A-ryhmä ei vielä yltänyt samalla tavoin kuin B- ja C-ryhmä.

Morfosyntaktisen kompleksisuuden kasvaminen voi tapahtua monilla eri rintamilla samanaikaisesti. B- ja C-ryhmän ensimmäiset 4- ja 5-kerroksiset elaborointityypit olivat kokonaan uusia tasoja repertoaarissa. Toisaalta morfologisten ja syntaktisten keinojen yhteiskäyttö oli osoitus elaboroinnin laajentumisesta keinojen yhdistelemisen kautta. Elaboroinnin eteneminen tapahtui eri keinojen näkökulmasta siten, että ensin se oli vain morfologista kunnes rinnalle tuli myös syntaktisesti elaboroituja komponentteja. Vasta tämän jälkeen keinoja yhdistelevät elaborointityypit pääsevät mukaan ilmauksiin. Jos kompleksisuuden kehitystä katsotaan etenemissuuntien näkökulmasta, oli elaborointi ensin mukana vain yhdessä ilmauksen komponentissa. Tämän jälkeen elaborointi eteni horisontaalisessa suussa toiseenkin komponenttiin, edelleen kerroksellisuuden kannalta varsin maltillisena. Vasta tämän jälkeen vertikaalinen, kunkin komponentit syvempi elaborointi pääsi kunnolla vauhtiin ja komponentit laajenivat myös yhä usemapikerroksisiksi.

Tutkimuskysymyksistä toinen käsitteli dysleksian ja morfosyntaktisen kompleksisuuden yhteyksiä. Aiemmissa tutkimuksissa on ehdotettu kolmea erilaista vaihtoehtoa, joilla dyslektikoilla todetut syntaktisten taitojen erot verrokkeihin nähden voitaisiin selittää. Ensimmäisen selitysvaihtoehdon mukaan dyslektikkojen lingvistiset taidot ovat joka viivästyneet tai heillä on rakenteellisia poikkeamia kielellisessä systeemissään. Näistä kumpikin voi aiheuttaa dyslektikoille ominaisiksi katsottuja syntaktisten perustaitojen puutteita. Tässä tutkimuksessa ei löytynyt mitään, mikä viittaisi syntaktisten perustaitojen puutteisiin dysleksiariskilapsilla. Riskilasten elaborointityyppi-inventaari osoittautui yhtä laajaksi ja monipuoliseksi sekä laadultaan yhtä kompleksiksi kuin verrokkilasten vastaava.

Toinen näkemys, joka yhdistää syntaktiset taidot ja dysleksian, painottaa fonologian osuutta dysleksiassa. Sen mukaan syntaktiset ongelmat eivät ole primääri dysleksian ominaisuus vaan seurausta fonologisista poikkeamista ja lyhytkestoisen muistin rajoituksista. On mahdollista että lyhytkestoisen muistin rajoitukset voisivat selittää sen, että dysleksiariskilasten tuotoksissa painopiste oli vain yhden komponentin elaboroimisessa eikä elaborointi laajentunut ilmauksen muihin komponentteihin samaan tapaan kuin verrokkiryhmässä. Kuitenkin tämän varmistaminen edellyttäisi lisätutkimusta.

Kolmas näkemys dysleksian ja syntaktisten taitojen suhteesta ei myöskään edellytä syntaktisten perustaitojen puutetta vaan tähdentää sitä, että dyslektikot eivät pysty käyttämään näitä taitojaan hyväkseen yhtä tehokkaasti kuin verrokit. Lähinnä metalingvistiseksi luonnehdittu ongelma tulisi esille juuri lukemisessa kirjoittamisessa, joita ei katsota yhtä luonteviksi lingvistisiksi toiminnoiksi kuin puhuminen. Kun tässä tutkimuksessa verrattiin dysleksiariskilapsia ja verrokkeja toisiinsa MLU-arvojen ja IPSyn-pisteiden avulla, ryhmien välillä ei löytynyt merkitsevää eroa. Vasta morfosyntaktisen elaboroinnin tarkka analyysi paljasti eroja. IPSyn ei päässyt käsiksi näihin eroihin, koska se rekisteröi erilaiset rakenteet ilman koko ilmauksen kontekstia ja erilaisten elaborointityyppien suhteenhan riskilapset olivat aivan tasavertaisia verrokkien kanssa. Myöskään MLU ei pystynyt kunnolla reagoimaan riskien ja verrokkien eroihin, koska se on puhtaasti matemaattinen keskiarvo ilmausten morfeemimäärästä. Koska ero kulminoitui vain monikomponenttisiin elaboroituihin ilmauksiin, se ei näy keskiarvossa tarpeeksi saadakseen aikaan merkitseviä ryhmäeroja. Tämä tutkimus siis tukee sitä oletusta, että dyslektikoilla ei ole primääriä syntaktisten perustaitojen vajetta vaan heidän käytössään on samanlaiset resurssit kuin verrokeillakin. Pienempi elaboroitujen komponenttien määrä ilmauksessa ei estä tehokasta ja tarkoituksenmukaista kommunikointia.

Kolmas tutkimuskysymys oli luonteeltaan metodologinen ja keskittyi selvittämään sitä, mitä MLU ja IPSyn oikeastaan mittaavat ja mitä tekemistä näillä mittaustuloksilla on tässä tutkimuksessa määritellyn morfosyntaktisen kompleksisuuden kanssa. Sekä MLU että IPSyn sopivat absoluuttisen kompleksisuuden viitekehykseen mutta ne käsittävät rakenteellisen kompleksisuuden eri tavoin. Vahvasta keskinäisestä korrelaatiosta huolimatta ne siis eivät mittaa samaa asiaa. MLU reagoi vain morfeemimääriin piittaamatta siitä, minkälaisista morfeemeista on kyse tai minkälaisia kokonaisuuksia morfeemeista muodostuu ilmauksen sisällä. Se näkee ilmaukset puhtaasti lineaarisina yksikköjonoina. Uudet innovaatiot, käyttöön tulevat rakenteet, rakenteiden yhdistelmät tai aiemmin käytettyjen rakenteiden laajentuminen eivät välttämättä näy MLU-arvossa, vaikka niillä saattaisi hyvinkin olla merkitystä morfosyntaktisen kompleksisuuden kannalta. IPSynin näkökulma rakenteelliseen kompleksisuuteen on aivan toinen: se pilkkoo jokaisen ilmauksen osiinsa, analysoi ja nimeää nämä osat sekä jakaa niille pisteet. IPSynin kannalta tärkeää on tehdä inventaario siitä, minkälaisia rakenteellisia resursseja kielenkäyttäjällä tuotoksen perusteella on. Sen sijaan IPSyn ei kiinnitä mitään huomiota siihen, miten näitä resursseja käytetään ja yhdistellään toisiinsa. Niinpä se ei pysty erottamaan juuri rakenteet oppinutta lasta toisesta, joka rakentaa samoista resursseista yhdistellen ja laajentaen komplekseja ilmauksia.

IPSyn arvioi kielisysteemiä ja sen laajuutta erilaisten morfosyntaktisten rakenteiden määrän perusteella kun taas MLU arvioi konkreettista kielentuotosta ja sen laajuutta morfeemimäärän perusteella. Yksittäisen ilmauksen näkökulmasta nämä mittarit tavallaan toteuttavat hajota ja hallitse -politiikkaa, jossa MLU hallitsee ilmauksen kokonaisuutta, vaikka ilmauksen rakenteellinen sisältö jääkin siltä analysoimatta. IPSyn taas pilkkoo ilmauksen rakenneyksiköihin niin, että ilmauksen kokonaisuus ei hahmotu. Kumpikaan mittari ei pysty saamaan esille ilmausten kerroksellista rakennetta ilmausanalyysin tavoin ja näin morfosyntaktisen kompleksisuuden moniulotteinen luonne latistuu ja häviää. Hyödyttömiä mittarit eivät kuitenkaan ole, sillä yhdessä käytettynä ja toisiinsa suhteutettuina ne antavat vahvoja viitteitä siitä, miten olemassa olevasta rakennemäärästä voidaan koota tietynmittaisia ilmauksia. Samalla ne alkavat paljastaa, miltä tulosten taakse piilotettu kieli rakenteeltaan näyttää. IPSyn- ja MLU-tulosten perusteella muodostettujen hypoteesien 1–5 osoittautuminen pääpiirteissään tosiksi on osoitus tästä kahden mittarin sopivuudesta hedelmälliseen yhteiskäyttöön.

Tässä tutkimuksessa tarkastelun kohteena oli rakenteellinen, erityisesti morfosyntaktinen kompleksisuus absoluuttisen kompleksisuuden viitekehyksessä. Tämä näkemys korostaa sitä, että kompleksisuutta arvioidaan kielen rakenteen ominaisuutena mahdollisimman objektiivisesti. Kielenkäyttäjän kokemukset rakenteista eivät siis vaikuta siihen, arvioidaanko ne yksinkertaisiksi vai komplekseiksi. Kaikki kolme käytettyä analysointimenetelmää edustavat absoluuttista näkemystä mutta silti niiden tuottamat tulokset ovat erilaisia. Niistä kukin painottaa eri aspektia kielestä ja myös hahmottaa kielen eri tavoin: MLUssa kieli on lineaarinen tuotos, IPSynissä se kulminoituu rakennevarastoon ja ilmausanalyysissä kieli näyttäytyy moniulotteisena ilmiönä. Tässä mielessä absoluuttinen kompleksisuus on yhtä suhteellista kuin kielenkäyttäjän kokemuksiin pohjautuva suhteellinen kompleksisuuskin. Lopputulokseen vaikuttavat siis ainakin lingvistinen taustateoria, analysoitava kokonaisuus (esim. ilmaus tai lause) ja analysointiyksikkö (morfeemi tai rakenne) sekä se, onko analysoinnin varsinaisena kohteena tuotos vai sen taustalla vaikuttava systeemi. Morfosyntaktinen kompleksisuus ei siis missään tapauksessa ole itsestäänselvyys eikä se myöskään ole absoluuttista sanan kirjaimellisessa merkityksessä. Onkin hyvin tärkeää, että kompleksisuus ja sen arviointiin vaikuttavat tekijät määritellään hyvin tarkasti aina, kun kompleksisuutta tutkitaan.

REFERENCES

Allen, S. E. M. & Crago, M. B. 1996. Early passive acquisition in Inuktitut. Journal of Child Language 23, 129–55.

Altenberg, B. 1990. Speech as linear composition. In G. Caie, K. Haastrup, A. L. Jakobsen, J. E. Nielsen, J. Sevaldsen, H. Specht & A. Zettersten (eds.), Proceedings from the Fourth Nordic Conference for English Studies, Vol 1. Department of English, University of Copenhagen, 133–143.

Bernstein Ratner, N. 1995. Language complexity and stuttering in children. Topics in Language Disorders 15 (3), 32–47.

Bernstein Ratner, N. & Sih, C. C. 1987. Effects of gradual increases in sentence length and complexity on children's disfluency, Journal of Speech and Hearing Disorders 52, 278–287.

Bishop, D. V. M. & Snowling, M.J. 2004. Developmental dyslexia and specific language impairment: same or different? Psychological Bulletin 130, 858–886.

Blake, J., Quartaro, G. & Onorati, S. 1993. Evaluating quantitative measure of grammatical complexity in spontaneous speech samples. Journal of Child Language 20, 139–152.

Bock, J. K. 1982. Toward a cognitive psychology of syntax: information processing contributions to sentence formulation. Psychological Review, 89, 1–47.

Bowerman, M. 1973. Early syntactic development. A cross-linguistic study with special reference to Finnish. Cambridge studies in linguistics 11. London: Cambridge University Press.

Bowerman, M. 1996. The origins of children's spatial semantic categories: cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (eds.), Rethinking linguistic relativity. Cambridge: Cambridge University Press, 145–176.

Brown, R. 1973/1976. A first language. The early stages. Harmondsworth: Penguin Education.

Catts, H. W. 1989a. Defining dyslexia as a developmental language disorder. Annals of Dyslexia 39, 50–64.

Catts, H. W. 1989b. Speech production deficits in developmental dyslexia. Journal of Speech and Hearing Disorders 54, 422–428.

Catts, H. W. 1991. Early identification of dyslexia: Evidence from a follow-up study of speech-language impaired children. Annals of Dyslexia 41, 163–177.

Cheung, H. & Kemper, S. 1992. Competing complexity metrics and adult's production of complex sentences. Applied Psycholinguistics 13, 53–76.

Crystal, D. 1987. Towards a "bucket" theory of language disability: Taking account of interaction between linguistic levels. Clinical Linguistics & Phonetics 1, 7–22.

Dahl, Ö. 2004. The growth and maintenance of linguistic complexity. Philadelphia, PA: John Benjamins.

Deutch, A. & Bentin, S. 1996. Attention factors mediating syntactic deficiency in reading-disabled children. Journal of Experimental Child Psychology 63, 386–415.

Dromi, E. & Berman, R. 1982. A morphemic measure of early language development: data from modern Hebrew. Journal of Child Language 9, 403–424.

Dromi, E. , Leonard, L. B. & Shteiman, M. 1993. The grammatical morphology of Hebrew-speaking children with specific language impairment: Some competing hypothesis. Journal of Speech and Hearing Research 36, 760–771.

Gallagher, A., Frith, U & Snowling, M. J. 2000. Precursors of Literacy delay among children at genetic risk of dyslexia. Journal of Child Psychology and Psychiatry 41, 203–213.

Gammerman, A. & Vovk, V. 1999. Kolmogorov complexity: Sources, theory and applications. The Computer Journal 42, 252–255.

Gell-Mann, M. 1994. The Quark and the Jaguar. Adventures in the Simple and the Complex. London: Little Brown.

Grela, B. G. & Leonard, L. B. 2000. The influence of argument-structure complexity on the use of auxiliary verbs by children with SLI. Journal of Speech, Language and Hearing Research 43, 1115–1125.

Hadley, P. 1998. Early verb-related vulnerability among children with specific language impairment. Journal of Speech, Language and Hearing Research 41, 1384–1397.

Hakulinen, L. 1979. Suomen kielen rakenne ja kehitys. Fourth, revised edition. Helsinki: Otava.

Hickey, T. 1991. Mean length of utterance and the acquisition of Irish. Journal of Child Language 18, 553–569.

ISK 2004 = Hakulinen, A. , Vilkuna, M., Korhonen R., Koivisto V., Heinonen T. R. & Alho, I. (toim.) 2004. Iso suomen kielioppi. SKST 950. Helsinki: Suomalaisen Kirjallisuuden Seura.

Joanisse, M. F., Manis, F. R., Keating, P. & Seidenberg, M. S. 2000. Language deficits in dyslexic children: Speech perception, phonology, and morphology. Journal of Experimental Child Psychology 77, 30–60.

Johnston, J. R. 2001. An alternate MLU calculation: Magnitude and variability of effects. Journal of Speech, Language, and Hearing Research 44, 156–164.

Kangassalo, R. 1995. Mastering the question. The acquisition of interrogative clauses by Finnish-speaking children. Umeå Studies in Humanities 126.

Karlsson, F. 1983. Suomen kielen äänne- ja muotorakenne. Porvoo: WSOY.

Kauppinen, A. 1996. The Italian imperfetto compared to the Finnish conditional verb form – Evidence from child language. Journal of Pragmatics 26, 109–136.

Kauppinen, A. 1998. Puhekuviot, tilanteen ja rakenteen liitto. Tutkimus kielen omaksumisesta ja suomen konditionaalista. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kemper, S., Rice, K. & Chen, Y.-J. 1995. Complexity metrics and growth curves for measuring grammatical development from five to ten. First Language 15, 151–166.

Kielinen, M. 1997. Monikon 3. persoonan inkongruenssin tarkastelua. S. Makkonen & H. Mantila (eds.), Pohjoissuomalaisen puhekielen sosiolingvistinen variaatio. University of Oulu. Publications of the Department of Finnish, Saami and Logopedics 8, 33–44.

Klee, T. & Fitzgerald M. D. 1985. The relation between grammatical development and mean length of utterance in morphemes. Journal of Child Language 12, 251–269.

Klee, T., Schaffer, M., May, S., Membrino, I. & Mogey, K. 1989. A comparison of the age–MLU relation in normal and specifically language-impaired children. Journal of Speech and Hearing Disorders 54, 226–233.

Kusters, W. 2003. Linguistic complexity. The influence of social change on verbal inflection. LOT Dissertation Series 77. Utrecht: LOT.

Kusters, W. & Muysken, P. 2001. The complexities of arguing about complexity. Linguistic Typology 5, 182–185.

Laalo, K. 1994. Kaksitavuvaihe lapsen kielen kehityksessä. Virittäjä 98, 430–448.

Laalo, K. 1995. Skeemakongruenssi. Morfologisten skeemojen kongruenssia lapsenkielessä. Virittäjä 99, 153–172.

Laalo, K. 1997. Alkukatsaus lapsenkielen esi- ja varhaismorfologiaan. Lyhyen ja pitkän loppuvokaalin sekä vartalonvaihteluiden ikonisuutta. Virittäjä 101, 186–207.

Laalo, K. 1998. Välikatsaus lapsenkielen varhaismorfologiaan. Virittäjä 102, 361–385.

Laalo, K. 2002: Acquisition of case in Finnish: a preliminary overview. In M. D Voiekova & W. U. Dressler (eds.), Pre- and protomorphology: Early phases of morphological development in nouns and verbs. Munich: Lincom Europa, 83–103.

Lahey, M., Liebergott, J., Chesnick, M., Menyuk, P. & Adams, J. 1992. Variability in children's use of grammatical morphemes. Applied psycholinguistics 13, 373–398.

Langacker, R. W. 1987. Foundations of cognitive grammar. Volume I: Theoretical prerequisites. Stanford: Stanford University Press.

Langacker, R. W. 1988a. An overview of cognitive grammar. In B. Rudzka-Ostyn (ed.), Topics in cognitive linguistics. Current issues in linguistic theory 50. Amsterdam: John Benjamins, 3–48.

Langacker, R. W. 1988b. A usage-based model. In B. Rudzka-Ostyn (ed.), Topics in cognitive linguistics. Current issues in linguistic theory 50. Amsterdam: John Benjamins, 127–161.

Larjavaara, M. 2001a. Tämä, tuo vai se? Kielikello 2001 (4), 17–22.

Larjavaara, M. 2001b. Määräinen artikkeli – suomessa? Kielikello 2001 (4), 25–27.

Laury, R. 1996. Sen kategoriasta – onko suomessa jo artikkeli. Virittäjä 100, 162–181.

Laury, R. 1997. Demonstratives in interaction. The emergence of a definite article in Finnish. Amsterdam: John Benjamins.

Leikin, M. & Assayag-Bouskila O. 2004. Expression of syntactic complexity in sentence comprehension: A comparison between dyslexic ans regular readers. Reading and Writing: An Interdisciplinary Journal 17, 801–821.

Leonard, C. M., Lombardino, L. J., Walsh, K., Eckert, M. A., Mockler, J. L., Rowe, L. A., Williams, S & DeBose C. B. 2002. Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children. Journal of Communication Disorders 35, 501–531.

Leong, C. K. & Parkinson, M. E. 1995. Processing of English morphological structure by poor readers, In C. K. Leong and R. M. Joshi (eds.), Developmental and Acquired Dyslexia. Dordrecht: Kluwer, 237–261.

Lieko, A. 1992. The development of complex sentences: a case study of Finnish. Helsinki: Suomalaisen Kirjallisuuden Seura.

Lieven , E., Behrens, H. Speares, J. & Tomasello, M. 2003. Early syntactic creativity: a usage-based approach. Journal of Child Language 30, 333–370.

Linell, P. 1998. Approaching dialogue. Talk, interaction and contexts in dialogical perspectives. Amsterdam: John Benjamins.

Locke, J. L., Hodgson, J., Macaruso, P., Roberts, J., Lambrecht-Smith, S. & Guttentag, C. 1997. The development of developmental dyslexia. In C. Hulme and M. Snowling (eds.), Dyslexia: Biology, Cognition and Intervention. London: Whurr Publishers, 72–96.

Logan, K. J. & LaSalle, L. R. 1999. Grammatical characteristics of children's conversational utterances that contain disfluency clusters. Journal of Speech, Language, and Hearing Research 42, 80–91.

Lyytinen, H., Leinonen, S., Nikula, M., Aro, M. & Leiwo, M. 1995. In search of the core features of dyslexia: observations concerning dyslexia in the highly orthographically regular Finnish language. In V. W. Berninger (ed.) The varieties of orthographic knowledge II: Relationships to phonology, reading, and writing. Dordrecht: Kluwer Academic Publishers, 177–204.

Lyytinen, P. 1978. The acquisition of Finnish morphology in early childhood. Jyväskylä Studies in Education, Psychology and Social Research 37.

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.

MacWhinney, B. 1991. The CHILDES project: tools for analyzing talk. Hillsdale, NJ: Lawrence Erlbaum Associates.

Manninen, T. 1997. Monikon 1. persoonan verbimuoto. In S. Makkonen & H. Mantila (eds.), Pohjoissuomalaisen puhekielen sosiolingvistinen variaatio. University of Oulu. Publications of the department of Finnish, Saami and logopedics 8, 24–32.

McWhorter, J. H. 2001. The world's simplest grammars are creole grammars. Linguistic Typology 5, 125–166.

Mielikäinen, A. 1981. Nominin- ja verbintaivutuksen ongelmia nykypuhekielessä. In A. Mielikäinen (ed.), Nykysuomalaisen puhekielen murros.

Jyväskylän osatutkimus. Raportti 3. Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja 26, 67–100.
Mielikäinen, A. 1984. Monikon 3. persoona kongruenssi puhekielessä. Virittäjä 88, 162–175.
Mielikäinen, A. 1991. Murteiden murros. Levikkikarttoja puhekielen piirteistä. Jyväskylän yliopiston suomen kielen laitoksen julkaisuja 36.
Miestamo M. 2006. On the feasibility of complexity metrics. In K. Kerge & M-M Sepper (eds.), FinEst Linguistics. Proceedings of the Annual Finnish and Estonian Conference of Linguistics. Tallinn, May 6–7, 2004. Publications of the Department of Estonian of Tallinn University 8, 11–26.
Miller, J. F. 1981. Assessing language production in children. Experimental procedures. London: Edward Arnold.
Nation, K., Clarke, P., Marshall. C. M. & Durand, M. 2004. Hidden language impairments in children: parallels between poor reading comprehension and specific language impairment. Journal of Speech, Language, and Hearing Research 47, 199–211.
Niemi, J. & Niemi, S. 1985. Suomenkielisen lapsen morfosyntaksin ja sanaston kehityksestä. Tapaustutkimus. Virittäjä 89, 152–171.
Nieminen, L. & Torvelainen, P. 2003. Produktiivisen syntaksin indeksi – suomenkielinen versio [Index of productive syntax – a Finnish version]. Puhe ja kieli 23, 119–132.
Oetting, J. B., Cantrell, J. P. & Horohov, J. E. 1999. A study of specific language impairment (SLI) in the context of non-standard dialect. Clinical Linguistics & Phonetics 13, 25–44.
Peters, A. 1983. The units of language acquisition. Cambridge University Press.
Peters, A. 1997. Language typology, prosody, and the acquisition of grammatical morphemes. In D. I. Slobin (ed.), The crosslinguistic study of language acquisition. Volume 5: Expanding the context. Mahwah, NJ: Lawrence Erlbaum Associates, 135–197.
Pinker, S. 1984. Language learnability and language development. Cambridge, MA: Harvard University Press.
Rescorla, L., Bascome, A., Lampard, J. & Feeny, N. 2001. Conversational patterns in late talkers at age 3. Applied Psycholinguistics 22, 235–251.
Rescorla, L., Dahlsgaar, K. & Roberts, J. 2000. Late-talking toddlers: MLU and IPSyn outcomes at 3;0 and 4;0. Journal of Child Language 27, 643–664.
Richardson, U. 1998. Familial dyslexia and sound duration in the quantity distinctions of Finnish infants and adults. Studia Philologica Jyväskyläensia 44.
Rollins, P. R. , Snow, C. E. & Willet, J. B. 1996. Predictors of MLU: semantic and morphological developments. First Language 16, 243–259.
Rondal, J. A., Ghiotto, M., Bredart, S. & Bachelet, J-F. 1987. Age-relation; reliability and grammatical validity of measures of utterance length. Journal of Child Language 14, 433–446.
Räisänen, A. 1975. Havaintoja lastenkielestä. Virittäjä 79, 251–266.
Savijärvi, I. 1977. Itämerensuomalaisten kielten kieltoverbi. I Suomi. [The negative verb in the Baltic- Finnic languages. 1. Finnish] Suomalaisen

Kirjallisuuden Seuran Toimituksia 333. Publications of the University of Joensuu, Series A No 8.

Scarborough, H. S. 1990a. Index of productive syntax. Applied Psycholinguistics 11, 1–22.

Scarborough, H. S. 1990b. Very early language deficits in dyslexic children. Child Development 61, 1728–1743.

Scarborough, H. S. 1991. Early syntactic development of dyslectic children. Annals of Dyslexia 41, 207–220.

Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E. & Sudhalter, V. 1991. The relation of utterance length to grammatical complexity in normal and language-disordered groups. Applied Psycholinguistics 12, 23–45.

Scarborough, H. S., Wyckoff, J. & Davidson, R. 1986. A reconsideration of the relation between age and mean length utterance. Journal of Speech and Hearing Research 29, 394–399.

Schnell de Acedo, B. 1994. Early morphological development: The acquisition of articles in Spanish. In J. Sokolov & C. Snow (Eds.) Handbook of research in language development using CHILDES. Hillsdale, NJ: Lawrence Erlbaum Associates, 210–253.

Silverman, S. W. & Bernstein Ratner, N. 1997. Syntactic complexity, fluency, and accuracy of sentence imitation in adolescents. Journal of Speech, Language, and Hearing Research 40, 95–107.

Smith, C. S 1988. Factors of linguistic complexity and performance. In A. Davison & G. M. Green (eds.), Linguistic complexity and text comprehension. Readability issues reconsidered. Hillsdale, NJ: Lawrence Erlbaum Associates, 247–279.

Smith, C. S. & van Kleeck, A. 1986. Linguistic complexity and performance. Journal of Child Language 13, 389–408.

Snowling, M., Bishop, D. V. M & Stothard, S. E. 2000. Is preschool language impairment a risk factor for dyslexia in adolescence? Journal of Child Psychology and Psychiatry and Allied Disciplines 41, 587–600.

Thordardottir, E. T. & Weismer, S. E. 1998. Mean length of utterance and other language sample measures in early Icelandic. First language 18, 1–32.

Toivainen, J. 1980. Inflectional affixes used by Finnish-speaking children aged 1–3 years. Suomalaisen kirjallisuuden seuran toimituksia 359. Helsinki: SKS.

Toivainen, J. 1990. Acquisition of Finnish as a first language: general and particular themes. Publications of the Department of Finnish and general linguistics of the University of Turku 35.

Tomasello, M. 2000a. Do young children have adult syntactic competence? Cognition, 74, 209–253.

Tomasello, M. 2000b. The item-based nature of children's syntactic development. Trends in Cognitive Sciences 4, 156–163.

Tomasello, M. 2000c. First steps toward a usage-based theory of language acquisition. Cognitive Linguistics 11, 61–82.

Tomasello, M. 2003. Constructing a language. A usage-based theory of language acquisition. Cambridge, MA: Harvard University Press.
Tomblin, J. B., Spencer, L, Flock, S., Tyler, R. & Gantz, B. 1999. A comparison of language achievement in children with cochlear implants and children using hearing aids. Journal of Speech, Language and Hearing Research 42, 497–511.
Vilkuna, M. 1996. Suomen lauseopin perusteet. Kotimaisten Kielten Tutkimuskeskuksen julkaisuja 90. Helsinki: Edita.
Wray, A. & Perkins, M. R. 2000. The functions of formulaic language: and integrated model. Language & Communication 20, 1–28.
Yaruss, J. S. 1999. Utterance length, syntactic complexity, and childhood stuttering. Journal of Speech, Language, and Hearing Research 42, 329–345.

APPENDIX 1 Codes used in morphological coding

CODES FOR PARTS OF SPEECH			
adj	adjective		
adj:pro	pronominal adjective		
adv	adverb		
conj	conjunction		
interj	interjection		
n	noun		
n:prop	proper noun		
num	numeral		
ptl	particle		
pro	personal pronoun		
pro:dem	demonstrative pronoun		
pro:indef	indefinite pronoun		
pro:int	interogative pronoun		
v	verb		
v:aux	auxiliary verb		
v:neg	negative auxiliary		
CODES FOR GRAMMATICAL MORPHEMES			
1s	1st person singular	**gen**	genitive
2s	2nd person singular	**ill**	illative
3s	3rd person singular	**imp**	imperative
1p	1st person plural	**iness**	inessive
2p	2nd person plural	**inf**	infinitive
3p	3rd person plural	**neg**	negative form
abl	ablative	**part**	participle
acc	accusative	**partit**	partitive
adess	adessive	**pass**	passive
all	allative	**past**	past tense
clit	clitic	**pl**	plural
cond	conditional mood	**poss**	possessive
cp	comparative	**post**	postposition
elat	elative	**pp**	past participle
ess	essive	**transl**	translative
THE BASIC SCHEMES FOR CODING OF WORDS ON %mor LINE			
part-of-speech | stem & fusional suffix			
part-of-speech | stem - suffix			

APPENDIX 2 The Index of Productive Syntax, the Finnish version

Child____________
Utterances________

Nominal Phrases

Item	Description	Examples	P
N1	PLURAL		
N2	INESSIVE/ADESSIVE		
N3	ELATIVE/ABLATIVE		
N4	ILLATIVE/ALLATIVE		
N5	PARTITIVE		
N6	GENITIVE		
N7	N/ADV + a clitic		
N8	$_{NP}$[W + W]$_{NP}$		
N9	$_{NP}$[W + W + W]$_{NP}$		
N10	Other complex nominal structure		

Total points ______/20

Verb Constructions

Item	Description	Examples	P
V1	SG 1P / PLU 1P		
V2	SG 2P / PLU 2P		
V3	PASSIVE VOICE		
V4	PAST TENSE		
V5	PERFECT TENSE		

V6	IMPERATIVE		
V7	CONDITIONAL		
V8	III INF		
V9	V + CLITIC		
V10	V + ADV		
V11	V + ADV +ADV		
V12	V + I INF		
V13	V + III INF		
V14	V + I INF + III INF		
V15	NEGATIVE AUX + V		
V16	NEGATIVE AUX+ V + I INF		
V17	Other complex verb structure		

Total points ______/34

Sentence Structure

Item	Description	Example	p
S1	W + W		
S2	S + V		
S3	[ATTR + S] + V		
S4	S + COP + N		
S5	V + OBJ		

S6	V + [ATTR + OBJ]		
S7	S + V + OBJ		
S8	V + OBJ + ADV		
S9	V + OBJ + ADV + ADV		
S10	S + V + ADV		
S11	S + V + ADV + ADV		
S12	S + V + [ATTR + ADV]		
S13	S + V + OBJ + ADV		
S14	S + V + OBJ + ADV + ADV		
S 15	INF-phrase in object position		
S 16	Adposition structure		
S 17	COORD CONJ		
S 18	SUBORD CONJ		
S 19	W+ COORD CONJ + W		
S 20	CL + COORD CONJ + CL		
S21	SUBORD CL		
S22	Other complex sentence structure		

Total points ______/44

Nominal phrases	/20
Verb structures	/34
Sentence structures	/44
TOTAL POINTS	/98

The basic principles in the Finnish version of IPSyn

A clause, an utterance or a phrase can be used as an example of the same structure only once. Thus, the phrase *tässä punaisessa autossa* 'in this red car' gives only one example of inessive case, even though there are three different words inflected in inessive. Another utterance with inessive case is needed to get maximum points.

Noun phrases

N1 PLURAL

The plural forms of all other cases than the partitive case are accepted. Words in plural partitive are typically the first plural forms that Finnish children produce, and they are not yet productive forms (e.g. Laalo 2002). The plural pronouns (*me* 'we', *te* 'you', *he* 'they', *nämä* 'these', *nuo* 'those', *ne* 'they' etc.) are not accepted as examples of plural forms. Also nouns used only in plural (the so-called plurale tantum words; *housut* 'trousers', *sakset* 'scissors') are ignored. A child is credited with two points if plural forms of two different words are used.

N2 INESSIVE/ADESSIVE

Both singular and plural forms are accepted. Demonstrative adverbs (e.g. *täällä* 'in here') are ignored. There are three possibilities to get two points: a) two different words inflected in inessive (or adessive respectively), b) the same word in both inessive and adessive form, and c) different words in inessive and adessive form.

N3 ELATIVE/ABLATIVE

Both singular and plural forms in elative and ablative are accepted. Demonstrative adverbs (e.g. *täältä* 'from here') are ignored.

N4 ILLATIVE/ALLATIVE

Both singular and plural forms are accepted.

N5 PARTITIVE

Both singular and plural forms are accepted. The partitive forms of mass nouns are ignored, because they are used in the same way as nominative forms of other nouns, as well as interrogative pronoun *ketä* 'who+partitive', which has a different stem from the corresponding nominative *kuka* 'who'. For maximum points two different words inflected partitive are required.

N6 GENITIVE

Both singular and plural forms are accepted. The interrogative pronouns *kenen* 'whose' and *keiden* 'whose+plural' are ignored because the stem is different from that in the basic form. Personal pronouns in genitive form (*minun/mun* 'my, mine' etc.) will give maximally one point. Words other than pronouns are needed to get two points.

N7 N/ADV + CLIT

All suffix particles, including the interrogative suffix -kO, are accepted. There are two alternatives to get maximum points from this item: a) the same suffix particle in two different words or b) two different suffix particles attached to the same word.

N8 $_{NP}[W + W]_{NP}$

Adjective + noun, genitive attribute + noun, indefinite or demonstrative pronoun + noun, quantifier + noun etc are accepted as examples of this structure. Two different occurrences of any of the structures mentioned above or two occurrences of the same structure in different words is required for two points in this item.

N9 $_{NP}[W + W + W]_{NP}$

See instructions in two-word NP (N8).

N10 Other complex nominal phrases

The comparative form of an adjective (except *parempi* 'better'), use of the possessive suffix, a four-word NP, translative and essive forms of nominals, two-word adjective phrases (*pienen pieni* 'tiny, minute') and adverb phrases (*hyvin myöhään* ‘very late’) are accepted as examples of other structures worth points.

Verb phrases

Both the affirmative and negative forms are accepted. The items V16 and V17 are reserved for negative expressions.

V1 SG 1P /PLU 1P

The passive forms are accepted as representatives of 1P PLU if they are clearly used as personal forms and not in passive meaning. In spoken Finnish, the *-n* marking the SG 1P is very often omitted. If this is the case, utterance is accepted as a representative of 1P SG if the verb is used together with pronoun *minä* ‘I’, or if consonant gradation is present (in verbs where consonant gradation is part of inflection).

V2 SG 2P/PLU 2P

Imperative forms in singular are not accepted as representatives of 2P.

V3 PASSIVE

All passive forms are accepted, including those referring to 1P PLU.

V4 PAST TENSE

All the colloquial variants of past tense forms are accepted. The only exception is the word *loppu* ‘stopped, finished, ran out’ because it is used both in past tense ftunction of the verb and as an adverb.

V5 PERFECT TENSE

All verbs in the perfect tense form are accepted without restrictions.

V6 IMPERATIVE

Katso/kato ‘look’ is not accepted, because it is also used as a particle (Hakulinen & Seppänen 1992).

V7 CONDITIONAL

All verbs in the conditional mode are accepted.

V8 III INFINITIVE

To get the maximum points of 2, both the inessive and illative forms must be found in a child’s utterances.

V9 V + CLITIC

All particle clitics are accepted. See the instructions in N7.

V10 V + ADV

Any adverbial expressing time, location, reason, manner, instrument etc. are accepted.

V11 V + ADV + ADV

See the instructions in V10.

V12 V + I INFINITIVE

Any verb together with another verb in the 1st infinitive form is accepted.

V13 V + III INFINITIVE

Any verb together with another verb in the 3rd infinitive form is accepted.

V14 V + I INFINITIVE + III INFINITIVE

Any verb structure consisting of any verb, a verb in 1st infinitive and a verb in 3rd infinitive are accepted.

V15 NEGATIVE AUXILIARY + V

Any combination of the negative auxiliary and a verb in the negative form is accepted.

V16 NEGATIVE AUXILIARY + V + I INFINITIVE

Any combination of the negative auxiliary, a verb in the negative form and a verb in the 1st infinitive form is accepted.

V17 Other complex verb constructions

The pluperfect, any inflected participle, NEG AUX + V + I INF + III INF, the 3P PLU in standard Finnish form etc. are accepted as an example of the other complex verb structures.

III Sentence structures

S1 W + W
Any combinations of two words are accepted.

S2 S + V
Any combinations of a subject and a predicate are accepted. Verb forms consisting of the stem and a personal suffix are considered as a combination of a subject and a predicate.

S3 [ATTR + S] + V
All structures including the combinations of the sentence constituents specified in the item description are accepted in the items S3–S15.

S4 S + Cop + N

S5 V + OBJ

S6 V + [ATTR + OBJ]

S7 S + V + OBJ

S8 V + OBJ + ADV

S9 V + OBJ + ADV + ADV

S10 S + V + ADV

S11 S + V + ADV + ADV

S12 S + V + [ATTR + ADV]

S13 S + V + OBJ + ADV

S14 S + V + OBJ + ADV + ADV

S15 INFINITIVE PHRASE IN OBJECT POSITION

S16 ADPOSITION PHRASE
The constructions consisting of a preposition + N or N + postposition are accepted. In Finnish, the nominal word is inflected usually either in genitive or partitive case in these constructions, but in children's utterances also constructions without the accurate inflection are accepted.

S17 COORD CONJ
To get maximum points, two different coordinative conjunctions must be used.

S18 SUBORD CONJ
To get maximum points, two different subordinate conjunctions must be used.

S19 W + COORD + W
Two words or two phrases are connected with a coordinative conjunction.

S20 CL + COORD + CL
Two clauses are connected with a coordinative conjunction.

S21 SUBORD CL
A subordinate clause is accepted and credited with a point in case it occurs together with a main clause.

S22 Other complex sentence structures
Constructions such as an infinitive construction including an object in the object position, an adverbial consisting of the III infinitive form and an object etc. are accepted as examples of the other complex sentence structures.

APPENDIX 3 The codes used in Utterance Analysis chart

A	adjective
AD	adpositional particle (postpositions and preposition)
ADV:DEM	demonstrative adverb
ADV:INT	interrogative adverb
ADVP	adverb phrase
AP	adjective phrase
AUX	auxiliary
AUX:NEG	negative auxiliary
AUX:V:NEG	negative form of an auxiliary (other than AUX:NEG)
CONJ	conjunction
INFP	infinitive phrase
INTERJ	interjection
N	noun
N(P)	proper noun
NP	noun phrase
NUM	numeral
PP	adposition phrase
PRO	personal pronoun
PRO: INDEF	indefinite pronoun
PRO:ADJ	adjectival pronoun
PRO:DEM	demonstrative pronoun
PRO:INT	interrogative pronoun
PTL	particle
V	verb / verb structure
V:NEG	negative form of a verb
V:NES	necessity verb

1p, 2p	1st and 2nd person plural
1s, 2s	1st and 2nd person singular
abl	ablative case
adess	adessive case
adv	adverb
all	allative case
clit	clitic particle
cond	conditional
elat	elative case
gen	genitive case
ill	illative case
imp	imperative
iness	inessive case
inf	infinitive
part	participle
partit	partitive case
pass	passive voice
past	past tense
pl	plural

APPENDIX 4 The elaboration types and tokens in two-, three- and four-component utterances

TABLE I Tokens of [M]-family elaboration types in 2-, 3-, and 4-component utterances

		2-comp U	3-comp U	4-comp U		Total
Group A						
2-layered						
	[M]	155	231	79		465
3-layered						
	[M2]	10	17	5		32
4-layered						
	[M3]	0	0	0		0
Group B		**2-comp U**	**3-comp U**	**4-comp U**		**Total**
2-layered						
	[M]	115	452	209		776
3-layered						
	[M2]	7	35	14		56
4-layered						
	[M3]	1	1	0		2
Group C		**2-comp U**	**3-comp U**	**4-comp U**		**Total**
2-layered						
	[M]	40	325	228		593
3-layered						
	[M2]	9	31	15		55
4-layered						
	[M3]	0	2	0		2

(dark grey)	elaboration type is used in all three utterance categories
(medium grey)	elaboration type used in two out of three utterance categories
(light grey)	elaboration type used in only one utterance category
(white)	elaboration type is not used

TABLE II Tokens of [S2]-family elaboration types in 2-, 3-, and 4-component utterances

Group A					
	2-comp U	**3-comp U**	**4-comp U**		**Total**
2-layered					
[S2]	17	20	4		41
3-layered					
[S2+(M)]	41	37	7		85
[S2+(2M)]	15	7	1		23
[S2+(S2)]	1	0	0		1
[S2+(S2+M)]	0	0	0		0
4-layered					
[S2+(S2+(M))]	4	0	0		4
[S2+(S2+(2M))]	2	1	1		4
[S2+(S3+(3M))]	0	1	0		1
[S2+(M2)]	6	2	0		8
[S2+(M2+M)]	0	0	1		1
[S2+(2M2)]	0	1	0		1
[S2+((S2+(M))+M)]	2	0	0		2
[S2+((S2+(2M))+M)]	0	0	0		0
5-layered					
[S2+(M3)]	0	0	0		0
[S2+(M3+M)]	0	0	0		0
[S2+(S2+(M2))]	0	0	0		0
[S2+(S2+(M2+M))]	0	0	0		0
Group B					
	2-comp U	**3-comp U**	**4-comp U**		**Total**
2-layered					
[S2]	24	63	30		117
[S3]	1	0	3		4
3-layered					
[S2+(M)]	43	103	30		176
[S2+(2M)]	21	47	9		77
[S2+(S2)]	0	0	0		0
[S2+(S2+M)]	1	1	0		2
4-layered					
[S2+(S2+(M))]	5	3	0		8
[S2+(S2+(2M))]	1	1	1		3
[S2+(S2+(3M))]	0	1	0		1
[S2+(M2)]	3	13	1		17
[S2+(M2+M)]	2	1	3		6
[S2+(2M2)]	0	0	0		0
[S2+((S2+(M))+M)]	4	2	1		7

TABLE II continues					
[S2+((S2+(2M))+M)]	0	2	0		2
5-layered					
[S2+(M3)]	0	0	0		0
[S2+(M3+M)]	0	1	1		2
[S2+(S2+(M2))]	1	0	0		1
[S2+(S2+(M2+M))]	1	1	0		2
<u>Group C</u>					
	2-comp U	**3-comp U**	**4-comp U**		**Total**
2-layered					
[S2]	4	56	32		92
[S3]	0	0	3		3
3-layered					
[S2+(M)]	24	98	33		155
[S2+(2M)]	15	58	21		94
[S2+(S2)]	2	2	1		0
[S2+(S2+M)]	0	1	0		1
4-layered					
[S2+(S2+(M))]	5	3	3		11
[S2+(S2+(2M))]	2	1	0		3
[S2+(S2+(3M))]	0	0	0		0
[S2+(M2)]	0	5	1		6
[S2+(M2+M)]	4	3	2		9
[S2+(2M2)]	0	0	0		0
[S2+((S2+(M))+M)]	3	4	1		8
[S2+((S2+(2M))+M)]	0	5	0		5
5-layered					
[S2+(M3)]	1	0	0		1
[S2+(M3+M)]	0	0	0		0
[S2+(S2+(M2))]	1	0	0		1
[S2+(S2+(M2+M))]	1	1	0		2

(dark grey)	elaboration type is used in all three utterance categories
(medium grey)	elaboration type used in two out of three utterance categories
(light grey)	elaboration type used in only one utterance category
(white)	elaboration type is not used

TABLE III Tokens of [S3]-family elaboration types in 2-, 3-, and 4-component utterances

Group A					
	2-comp U	**3-comp U**	**4-comp U**		**Total**
2-layered					
[S3]	1	0	0		1
3-layered					
[S3+(M)]	1	0	0		1
[S3+(2M)]	0	1	0		1
[S3+(3M)]	1	1	0		2
[S3+(S2+2M)]0					
4-layered					
[S3+(M2+M)]	0	0	1		1
[S3+(M2+2M)]	0	0	0		0
[S3+((S2+(M))+M)]	0	0	0		0
[S3+((S2+(2M))+M)]	0	0	0		0
Group B					
	2-comp U	**3-comp U**	**4-comp U**		**Total**
2-layered					
[S3]	1	0	0		1
3-layered					
[S3+(M)]	0	1	0		1
[S3+(2M)]	1	1	1		3
[S3+(3M)]	0	1	1		2
[S3+(S2+2M)]	0	0	0		0
4-layered					
[S3+(M2+M)]	0	0	0		0
[S3+(M2+2M)]	0	1	0		1
[S3+((S2+(M))+M)]	0	0	1		1
[S3+((S2+(2M))+M)]	0	0	0		0

TABLE III continues						
Group C						
		2-comp U	**3-comp U**	**4-comp U**		**Total**
2-layered						
	[S3]	1	0	0		1
3-layered						
	[S3+(M)]	0	1	0		1
	[S3+(2M)]	0	1	0		1
	[S3+(3M)]	0	0	0		0
	[S3+(S2+2M)]	0	1	0		1
4-layered						
	[S3+(M2+M)]	0	0	0		0
	[S3+(M2+2M)]	0	0	0		0
	[S3+((S2+(M))+M)]	0	0	1		1
	[S3+((S2+(2M))+M)]	0	1	0		1

(dark grey)	elaboration type is used in all three utterance categories
(medium grey)	elaboration type used in two out of three utterance categories
(light grey)	elaboration type used in only one utterance category
(white)	elaboration type is not used

APPENDIX 5 The elaboration combinations in utterances produced by individual children in group A (ordered by IPSyn score)

TABLE I 2-component utterances in group A						
ID		**22**	**6**	**18**	**17**	**36**
Name		**Anna**	**Ronja**	**Saku**	**Tatu**	**Sampo**
IPSyn		**42**	**51**	**53**	**56**	**57**
Two elaborated components		**3**	**7**	**11**	**3**	**3**
[M]+	[M]	1	4	6	3	1
[M2]+	[M]					1
[S2]+	[M]	1	1	2		
[S2+(M)]+	[M]	1	1	1		
[S2+(2M)]+	[M]		1			1
[S2+(2M)]+	[S2+(M)]			1		
[S2+(S2+(M))]+	[M]					
[S2+((S2+(M))+M)]	[M]					
[S3]+	[S2+(M)]					
[S3+(3M)]+	[M]			1		
One elaborated component		**18**	**25**	**25**	**15**	**26**
[M]		3	18	13	7	19
[M2]						4
[S2]			2	2	5	
[S2+(M)]		12	1	5	2	3
[S2+(M2)]						
[S2+(2M)]		2	1	4		
[S2+(S2)]			1			
[S2+(S2+(M))]			1	1		
[S2+(S2+(2M))]		1				
[S2+((S2+(M))+M)]					1	
[S3+(M)]			1			
Unelaborated two-component U		**0**	**1**	**2**	**4**	**5**
Total number of two-component U		**21**	**33**	**38**	**22**	**34**

TABLE I continues						
ID		**29**	**35**	**31**	**34**	
Name		**Elisa**	**Seppo**	**Aleksi**	**Leo**	**Total**
IPSyn		**57**	**60**	**66**	**68**	
Two elaborated components		**5**	**6**	**2**	**5**	**45**
[M]+	[M]	1	2		2	20
[M2]+	[M]	2				3
[S2]+	[M]	1	1	1	1	8
[S2+(M)]+	[M]	1	1		1	6
[S2+(2M)]+	[M]					2
[S2+(2M)]+	[S2+(M)]		1			2
[S2+(S2+(M))]+	[M]				1	1
[S2+((S2+(M))+M)]	[M]		1			1
[S3]+	[S2+(M)]			1		1
[S3+(3M)]+	[M]					1
One elaborated component		**21**	**14**	**9**	**13**	**166**
[M]		8	8	4	13	93
[M2]			1			5
[S2]		1		2		12
[S2+(M)]		9	3	1		36
[S2+(M2)]			1	1		2
[S2+(2M)]		1	1	1		10
[S2+(S2)]						1
[S2+(S2+(M))]		1				3
[S2+(S2+(2M))]		1				2
[S2+((S2+(M))+M)]						1
[S3+(M)]						1
Unelaborated two-component U		**1**	**6**	**8**	**8**	**35**
Total number of two-component U		**27**	**26**	**19**	**26**	**246**

TABLE II Three-component utterances in group A							
ID			**22**	**6**	**18**	**17**	**36**
Name			**Anna**	**Ronja**	**Saku**	**Tatu**	**Sampo**
IPSyn			**42**	**51**	**53**	**56**	**57**
Three elaborated components			**1**	**0**	**1**	**1**	**3**
[M]	[M]	[M]	1		1		2
[M2]	[M]	[M]					1
[S2]	[M]	[M]					
[S2+(M)]	[M]	[M]				1	
Two elaborated components			**9**	**6**	**5**	**10**	**8**
[M]	[M]		5	4	2	5	7
[M2]	[M]						1
[S2]	[M]		1	2	2	1	
[S2]	[M2]						
[S2+(M)]	[S2]		1				
[S2+(M)]	[M]		1		1	3	
[S2+(M)]	[S2+(M)]		1				
[S2+(2M)]	[M]					1	
[S2+(M2)]	[M]						
[S2+(S2+(2M))]	[M]						
One elaborated component			**18**	**19**	**9**	**19**	**15**
[M]			11	16	7	16	8
[M2]							3
[S2]					1	2	
[S2+(M)]			6	3		1	1
[S2+(2M)]					1		2
[S2+(2M2)]							1
[S2+(S3+(3M))]							
[S3+(2M)]			1				
Unelaborated 3-component U			**24**	**6**	**4**	**11**	**2**
Total number of 3-component U			**52**	**31**	**19**	**41**	**28**

TABLE II continues							
ID			**29**	**35**	**31**	**34**	
Name			**Elisa**	**Seppo**	**Aleksi**	**Leo**	**Total**
IPSyn			**57**	**60**	**66**	**68**	
Three elaborated components			**0**	**2**	**0**	**0**	**8**
[M]	[M]	[M]		1			5
[M2]	[M]	[M]					1
[S2]	[M]	[M]		1			1
[S2+(M)]	[M]	[M]					1
Two elaborated components			**4**	**12**	**8**	**9**	**71**
[M]	[M]			7	2	4	36
[M2]	[M]		1	2		1	5
[S2]	[M]		2		2		10
[S2]	[M2]					1	1
[S2+(M)]	[S2]						1
[S2+(M)]	[M]		1	1	2	2	11
[S2+(M)]	[S2+(M)]				1		2
[S2+(2M)]	[M]					1	2
[S2+(M2)]	[M]			1	1		2
[S2+(S2+(2M))]	[M]			1			1
One elaborated component			**30**	**8**	**20**	**15**	**153**
[M]			21	8	14	10	111
[M2]			2		1	2	8
[S2]			3		1		7
[S2+(M)]			4		1	2	18
[S2+(2M)]					2	1	6
[S2+(2M2)]							1
[S2+(S3+(3M))]					1		1
[S3+(2M)]							1
Unelaborated 3-component U			**7**	**4**	**7**	**8**	**73**
Total number of 3-component U			**41**	**26**	**35**	**32**	**305**

TABLE III Four-component utterances in group A								
ID				**22**	**6**	**18**	**17**	**36**
Name				**Anna**	**Ronja**	**Saku**	**Tatu**	**Sampo**
IPSyn				**42**	**51**	**53**	**56**	**57**
Four elaborated components				**0**	**1**	**0**	**0**	**0**
[M]	[M]	[M]	[M]		1			
Three elaborated components				**0**	**0**	**1**	**0**	**2**
[M]	[M]	[M]						2
[S2]	[M]	[M]						
[S2+(M)]	[M]	[M]				1		
[S2+(M)]	[M2]	[M]						
[S2+(S2+(2M))]	[S3+(M2+M)]	[M]						
Two elaborated components				**1**	**2**	**1**	**3**	**1**
[M]	[M]			1	1	1	2	1
[S2]	[M]				1			
[S2+(M)]	[M]						1	
[S2+(M)]	[M2]							
[S2+(2M)]	[M]							
[S2+(M2+M)]	[M]							
One elaborated component				**3**	**5**	**0**	**3**	**1**
[M]				1	5		3	1
[M2]								
[S2]				1				
[S2+(M)]				1				
Unelaborated 4-component U				**2**	**1**	**1**	**2**	**0**
Total number of 4-component U				**6**	**9**	**3**	**8**	**4**
ID				**29**	**35**	**31**	**34**	
Name				**Elisa**	**Seppo**	**Aleksi**	**Leo**	**Total**
IPSyn				**57**	**60**	**66**	**68**	
Four elaborated components				**0**	**0**	**0**	**0**	**1**
[M]	[M]	[M]	[M]					1

TABLE III continues							
Three elaborated components			**0**	**3**	**5**	**2**	**13**
[M]	[M]	[M]		3	1	1	7
[S2]	[M]	[M]			1		1
[S2+(M)]	[M]	[M]				1	2
[S2+(M)]	[M2]	[M]			2		2
[S2+(S2+(2M))]	[S3+(M2+M)]	[M]			1		1
Two elaborated components			**2**	**2**	**0**	**5**	**17**
[M]	[M]		1	2		3	12
[S2]	[M]						1
[S2+(M)]	[M]						1
[S2+(M)]	[M2]		1				1
[S2+(2M)]	[M]					1	1
[S2+(M2+M)]	[M]					1	1
One elaborated component			**4**	**1**	**2**	**3**	**22**
[M]			3	1	2	1	17
[M2]			1			1	2
[S2]						1	2
[S2+(M)]							1
Unelaborated 4-component U			**2**	**0**	**2**	**2**	**12**
Total number of 4-component U			**8**	**6**	**9**	**12**	**65**

APPENDIX 6 The elaboration combinations in utterances produced by individual children in group C (ordered by IPSyn score)

TABLE I Two-component utterances in group C						
ID		**11**	**2**	**27**	**7**	**23**
Name		**Jaakko**	**Riikka**	**Janna**	**Taru**	**Anniina**
IPSyn		**62**	**68**	**69**	**70**	**72**
Two elaborated components		**1**	**6**	**3**	**7**	**2**
[M]	[M]					
[S2]	[S2]					1
[M2]	[M]		1			
[M2]	[S2]					
[S2+(M)]	[M]		3	1	2	
[S2+(M)]	[S2+(M)]					
[S2+(2M)]	[M]		1	2	2	
[S2+(2M)]	[S2]				1	
[S2+(2M)]	[S2+(M)]					
[S2+(2M)]	[S2+(2M)]	1				
[S2+(S2)]	[M]					1
[S2+(S2)]	[S2+(M)]					
[S3+(2M)]	[M]					
[S2+(M2+M)]	[M]					
[S2+(S2+(M))]	[M]					
[S2+(S2+(M2))]	[M]		1			
[S2+(S2+(2M))]	[M]				1	
[S2+((S2+(M))+M)]	[M]				1	
[S2+(M3)]	[S2+(M)]					
One elaborated component		**1**	**6**	**3**	**0**	**5**
[M]		1				
[M2]			1			3
[S2+(M)]				1		
[S2+(2M)]			2	1		1
[S3+(2M)]						
[S2+(M2+M)]			2			
[S2+(S2+(M))]			1			1
[S2+(S2+(M2+M))]				1		
Unelaborated 2-component U		**0**	**0**	**0**	**0**	**0**
Total number of 2-component U		**2**	**12**	**6**	**7**	**7**

TABLE I continues

ID		**30**	**19**	**3**	**1**	
Name		**Elina**	**Tuomo**	**Laura**	**Tuija**	**Total**
IPSyn		**74**	**75**	**76**	**84**	
Two elaborated components		**5**	**3**	**7**	**10**	**44**
[M]	[M]	2		1	2	5
[S2]	[S2]					1
[M2]	[M]	1			1	3
[M2]	[S2]		1			1
[S2+(M)]	[M]	1	1	3	3	14
[S2+(M)]	[S2+(M)]	1				1
[S2+(2M)]	[M]					5
[S2+(2M)]	[S2]					1
[S2+(2M)]	[S2+(M)]			1		1
[S2+(2M)]	[S2+(2M)]					1
[S2+(S2)]	[M]					1
[S2+(S2)]	[S2+(M)]				1	1
[S3+(2M)]	[M]				1	1
[S2+(M2+M)]	[M]		1			1
[S2+(S2+(M))]	[M]				2	2
[S2+(S2+(M2))]	[M]					1
[S2+(S2+(2M))]	[M]			1		2
[S2+((S2+(M))+M)]	[M]					1
[S2+(M3)]	[S2+(M)]			1		1
One elaborated component		**2**	**3**	**1**	**2**	**23**
[M]						1
[M2]		1				5
[S2+(M)]			1			2
[S2+(2M)]			2	1		7
[S3+(2M)]					1	1
[S2+(M2+M)]		1				3
[S2+(S2+(M))]					1	3
[S2+(S2+(M2+M))]						1
Unelaborated 2-component U		**0**	**0**	**0**	**0**	**0**
Total number of 2-component U		**7**	**6**	**8**	**12**	**67**

TABLE II Three-component utterances in group C

ID			**11**	**2**	**27**	**7**	**23**
Name			**Jaakko**	**Riikka**	**Janna**	**Taru**	**Anniina**
IPSyn			**62**	**68**	**69**	**70**	**72**
Three elaborated components			**4**	**10**	**3**	**0**	**1**
[M]	[M]	[M]	1	1			1
[S2]	[M]	[M]		2	1		
[S2]	[S2]	[M]	1				
[M2]	[M]	[M]		2			
[S2+(M)]	[M]	[M]	1	1	1		
[S2+(M)]	[M2]	[M]		2			
[S2+(M)]	[S2+(M)]	[M]	1				
[S2+(2M)]	[M]	[M]		2			
[S2+(S2)]	[S2+(2M)]	[M]					
[S2+(M2+M)]	[S2+(M)]	[M]					
[S2+((S2+(M))+M)]	[M]	[M]					
[S2+((S2+(M))+M)]	[S2+(2M)]	[M]					
[S2+((S2+(M))+M)]	[S2+(S2+M)]	[M]			1		
Two elaborated components			**13**	**22**	**10**	**22**	**10**
[M]	[M]		6	7	5	6	6
[S2]	[M]		2	1	1	4	2
[S2]	[S2]		1				
[M2]	[M]					1	
[M2]	[M2]						
[M2]	[S2]			2	1		
[S2+(M)]	[M]		2	8	2	5	
[S2+(M)]	[M2]			1			
[S2+(M)]	[S2]		1	1			
[S2+(2M)]	[M]		1	1	1	4	1
[S2+(2M)]	[M2]						
[S2+(2M)]	[S2+(M)]					1	
[S2+(2M)]	[S2+(2M)]					1	
[S2+(2M)]	[S2+(S2+M)]			1			
[M3]	[M]						1
[S2+(M2)]	[M]						
[S2+(S2+(2M))]	[M]						
[S2+((S2+(M))+M)]	[M]						
[S2+((S2+(2M))+M)]	[M]						
[S2+((S2+(2M))+M)]	[S2+(M)]						

TABLE II continues							
ID			**30**	**19**	**3**	**1**	
Name			**Elina**	**Tuomo**	**Laura**	**Tuija**	**Total**
IPSyn			**74**	**75**	**76**	**84**	
Three elaborated components			**4**	**7**	**5**	**2**	**36**
[M]	[M]	[M]	2	2	1		8
[S2]	[M]	[M]	1	2		1	7
[S2]	[S2]	[M]					1
[M2]	[M]	[M]	1				3
[S2+(M)]	[M]	[M]		1	2	1	7
[S2+(M)]	[M2]	[M]					2
[S2+(M)]	[S2+(M)]	[M]					1
[S2+(2M)]	[M]	[M]					2
[S2+(S2)]	[S2+(2M)]	[M]		1			1
[S2+(M2+M)]	[S2+(M)]	[M]		1			1
[S2+((S2+(M))+M)]	[M]	[M]			1		1
[S2+((S2+(M))+M)]	[S2+(2M)]	[M]			1		1
[S2+((S2+(M))+M)]	[S2+(S2+M)]	[M]					1
Two elaborated components			**18**	**22**	**31**	**24**	**172**
[M]	[M]		9	10	8	5	62
[S2]	[M]			1	2	3	16
[S2]	[S2]				1		2
[M2]	[M]		2			2	5
[M2]	[M2]		1				1
[M2]	[S2]		1			1	5
[S2+(M)]	[M]		1	5	8	5	36
[S2+(M)]	[M2]		1		2		4
[S2+(M)]	[S2]					1	3
[S2+(2M)]	[M]		2	3	5	5	23
[S2+(2M)]	[M2]				1		1
[S2+(2M)]	[S2+(M)]				2		3
[S2+(2M)]	[S2+(2M)]				1	1	3
[S2+(2M)]	[S2+(S2+M)]						1
[M3]	[M]				1		2
[S2+(M2)]	[M]			1			1
[S2+(S2+(2M))]	[M]			1			1
[S2+((S2+(M))+M)]	[M]		1				1
[S2+((S2+(2M))+M)]	[M]					1	1
[S2+((S2+(2M))+M)]	[S2+(M)]			1			1

TABLE II continues

ID	**11**	**2**	**27**	**7**	**23**
Name	**Jaakko**	**Riikka**	**Janna**	**Taru**	**Anniina**
IPSyn	**62**	**68**	**69**	**70**	**72**
One elaborated component	**17**	**13**	**23**	**19**	**24**
[M]	4	1	12	5	9
[S2]	8		5	1	4
[S9]					1
[M2]		1			1
[S2+(M)]	1	6	2	8	3
[S2+(2M)]	3	2	2	5	4
[S2+(S2)]					
[S3+(M)]			1		
[S3+(S2+2M)]					1
[S2+(M2)]	1				
[S2+(M2+M)]		1			
[S2+(S2+(M))]					
[S2+((S2+(2M))+M)]		2			
[S3+((S2+(2M))+M)]					1
[S2+((S2+(M2+M)))]			1		
Unelaborated 3-component U	**0**	**0**	**1**	**0**	**5**
Total number of 3-component U	**34**	**45**	**37**	**41**	**40**

ID	**30**	**19**	**3**	**1**	
Name	**Elina**	**Tuomo**	**Laura**	**Tuija**	**Total**
IPSyn	**74**	**75**	**76**	**84**	
One elaborated component	**14**	**18**	**10**	**10**	**148**
[M]	5	7	1	1	45
[S2]	1				19
[S9]					1
[M2]	4	1			7
[S2+(M)]		6	6	7	39
[S2+(2M)]	1	3	2	2	24
[S2+(S2)]			1		1
[S3+(M)]					1
[S3+(S2+2M)]					1
[S2+(M2)]					1
[S2+(M2+M)]	1				2
[S2+(S2+(M))]	1	1			2
[S2+((S2+(2M))+M)]	1				3
[S3+((S2+(2M))+M)]					1
[S2+((S2+(M2+M)))]					1
Unelaborated 3-component U	**0**	**0**	**0**	**0**	**6**
Total number of 3-component U	**36**	**47**	**46**	**36**	**362**

TABLE III Four-component utterances in group C								
ID				**11**	**2**	**27**	**7**	**23**
Name				**Jaakko**	**Riikka**	**Janna**	**Taru**	**Anniina**
IPSyn				**62**	**68**	**69**	**70**	**72**
Four elaborated components				**0**	**0**	**0**	**2**	**0**
[M]	[M]	[M]	[M]					
[M2]	[M]	[M]	[M]				1	
[S2+(2M)]	[S2]	[M]	[M]				1	
Three elaborated components				**2**	**5**	**2**	**1**	**1**
[M]	[M]	[M]		1	3			1
[S2]	[M]	[M]		1				
[S2+(M)]	[M]	[M]				1		
[S2+(M)]	[S2+(M)]	[M]			1			
[S2+(2M)]	[M]	[M]						
[S2+(2M)]	[M2]	[M]			1			
[S2+(2M)]	[S2+(M)]	[M]					1	
[S2+(2M)]	[S2+(M)]	[S2]						
[S2+(2M)]	[S2+(2M)]	[M]				1		
[S2+(M2+M)]	[M]	[M]						
[S2+(S2+(M)]	[S2+(M)]	[M]						
Two elaborated components				**16**	**9**	**5**	**11**	**8**
[M]	[M]			7	6	3	2	4
[M2]	[M]			1	2		1	
[M2]	[M2]							
[M2]	[S3]							1
[S2]	[M]			6			3	
[S2+(M)]	[M]			1		2	1	1
[S2+(M)]	[S2]						1	1
[S2+(M)]	[S2+(M)]				1			
[S2+(2M)]	[M]						2	1
[S2+(2M)]	[S2]							
[S2+(M2+M)]	[M]							
[S2+((S2+(M))+M)]	[S3]			1				
[S3+((S2+(M))+M)]	[S2+(2M)]						1	

TABLE III continues

ID				**30**	**19**	**3**	**1**	
Name				**Elina**	**Tuomo**	**Laura**	**Tuija**	**Total**
IPSyn				**74**	**75**	**76**	**84**	
Four elaborated components				**0**	**0**	**1**	**0**	**3**
[M]	[M]	[M]	[M]			1		1
[M2]	[M]	[M]	[M]					1
[S2+(2M)]	[S2]	[M]	[M]					1
Three elaborated components				**1**	**5**	**3**	**7**	**27**
[M]	[M]	[M]		1	1	2	1	10
[S2]	[M]	[M]			1		1	3
[S2+(M)]	[M]	[M]			1	1	1	4
[S2+(M)]	[S2+(M)]	[M]						1
[S2+(2M)]	[M]	[M]					1	1
[S2+(2M)]	[M2]	[M]			1			2
[S2+(2M)]	[S2+(M)]	[M]					1	2
[S2+(2M)]	[S2+(M)]	[S2]					1	1
[S2+(2M)]	[S2+(2M)]	[M]						1
[S2+(M2+M)]	[M]	[M]			1			1
[S2+(S2+(M)]	[S2+(M)]	[M]					1	1
Two elaborated components				**12**	**10**	**8**	**7**	**86**
[M]	[M]			6	5	4	4	41
[M2]	[M]			1				5
[M2]	[M2]			1				1
[M2]	[S3]							1
[S2]	[M]			2				11
[S2+(M)]	[M]				3		3	11
[S2+(M)]	[S2]			1		1		4
[S2+(M)]	[S2+(M)]							1
[S2+(2M)]	[M]				1	2		6
[S2+(2M)]	[S2]				1			1
[S2+(M2+M)]	[M]			1		1		2
[S2+((S2+(M))+M)]	[S3]							1
[S3+((S2+(M))+M)]	[S2+(2M)]							1

TABLE III continues					
ID	**11**	**2**	**27**	**7**	**23**
Name	**Jaakko**	**Riikka**	**Janna**	**Taru**	**Anniina**
IPSyn	**62**	**68**	**69**	**70**	**72**
One elaborated component	**13**	**4**	**14**	**6**	**8**
[M]	6	4	10	3	6
[M2]			1		
[S2]	5		2	1	1
[S2+(M)]	1		1	2	
[S2+(2M)]					
[S2+(S2)]					
[S2+(M2)]					1
[S2+(S2+(M))]					
[S3]	1				
Unelaborated 4-component U	**2**	**1**	**3**	**5**	**5**
Total number of 4-component U	**33**	**19**	**24**	**25**	**22**
ID	**30**	**19**	**3**	**1**	
Name	**Elina**	**Tuomo**	**Laura**	**Tuija**	**Total**
IPSyn	**74**	**75**	**76**	**84**	
One elaborated component	**14**	**6**	**7**	**4**	**76**
[M]	9	5	2	2	47
[M2]	2			1	4
[S2]			1		10
[S2+(M)]	1		1	1	7
[S2+(2M)]	1				1
[S2+(S2)]	1		3		4
[S2+(M2)]					1
[S2+(S2+(M))]		1			1
[S3]					1
Unelaborated 4-component U	**0**	**1**	**3**	**0**	**20**
Total number of 4-component U	**27**	**22**	**22**	**18**	**212**

APPENDIX 7 Inventory of elaboration types and combinations produced by children in all three groups (ordered by MLU values)

TABLE I Elaboration types when only one component is elaborated in an utterance

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

ID	**35**	**22**	**18**	**36**	**17**	**6**
Name	**Seppo**	**Anna**	**Saku**	**Sampo**	**Tatu**	**Ronja**
MLU	**3,688**	**3,888**	**3,962**	**3,987**	**4,063**	**4,1**
2-layered components						
[M]	234	234	23	234	234	234
[S2]		4	23		23	2
3-layered components						
[M2]	2			23		
[S2+(M)]	2	234	2	23	23	23
[S2+(2M)]	2	2	23	3		2
[S2+(S2)]						2
[S3+(M)]						2
[S3+(2M)]		3				
4-layered components						
[S2+(M2)]	2					
[S2+(2M2)]				3		
[S2+(S2+(M))]			2			2
[S2+(S2+(2M))]		2				
[S2+((S2+(M))+M)]					2	

TABLE I continues						
ID	**31**	**34**	**29**	**13**	**8**	**15**
Name	**Aleksi**	**Leo**	**Elisa**	**Joel**	**Aino**	**Mika**
MLU	**4,2**	**4,263**	**4,388**	**4,713**	**4,775**	**4,838**
2-layered components						
[M]	234	234	234	234	234	234
[S2]	23	4	23	23	23	23
3-layered components						
[M2]	3	34	34	3	3	23
[S2+(M)]	23	3	23	23	23	2
[S2+(2M)]	23	3	2	2	3	23
[S3+(2M)]					2	
[S3+(3M)]				3		
4-layered components						
[M3]				2		
[S2+(M2)]	2			2	3	
[S2+(S2+(M))]			2			
[S2+(S2+(2M))]			2			
[S2+(S3+(3M))]	3					
ID	**24**	**38**	**10**	**23**	**11**	**19**
Name	**Sanni**	**Pekka**	**Santeri**	**Anniina**	**Jaakko**	**Tuomo**
MLU	**4,963**	**5,275**	**5,713**	**6,05**	**6,112**	**6,237**
2-layered components						
[M]	234	34	234	34	234	34
[S2]	3	3	34	34	34	
[S3]					4	
3-layered components						
[M2]	3	3	34	23		3
[S2+(M)]	23	34	234	3	34	23
[S2+(2M)]	3	3		23	3	23
[S3+(S2+2M)]				3		

TABLE I continues						
4-layered components						
[S2+(M2)]			3	4	3	
[S2+(S2+(M))]	2			2		34
[S3+(M2+2M)]	3					
[S3+((S2+(2M))+M)]				3		
5-layered components						
[S2+(M3+M)]			4			
ID	**1**	**3**	**2**	**27**	**30**	**7**
Name	**Tuija**	**Laura**	**Riikka**	**Janna**	**Elina**	**Taru**
MLU	**6,25**	**6,25**	**6,287**	**6,3**	**6,325**	**6,35**
2-layered components						
[M]	34	34	34	34	34	34
[S2]		4		34	3	34
3-layered components						
[M2]	4		23	4	234	
[S2+(M)]	34	34	3	234	4	34
[S2+(2M)]	3	23	23	23	34	3
[S2+(S2)]		34			4	
[S3+(M)]				3		
[S3+(2M)]	2					
4-layered components						
[S2+(M2+M)]			23		23	
[S2+(S2+(M))]	2		2		3	
5-layered components						
[S2+(S2+(M2+M))]				23		

TABLE II Elaboration combinations when two components are elaborated

2 = elaboration combination occurs in two-component utterances
3 = elaboration combination occurs in three-component utterances
4 = elaboration combination occurs in four-component utterances

ID		**35**	**22**	**18**	**36**	**17**	**6**
Name		**Seppo**	**Anna**	**Saku**	**Sampo**	**Tatu**	**Ronja**
MLU		**3,688**	**3,888**	**3,962**	**3,987**	**4,063**	**4,1**
2+2 layers							
[M]	**[M]**	23	234	234	234	234	234
[S2]	**[M]**	2	23	23		3	234
3+2 layers							
[M2]	**[M]**	3			23		
[S2+(M)]	**[M]**	23	23	23		34	2
[S2+(M)]	**[S2]**		3				
[S2+(2M)]	**[M]**				2	3	2
[S3+(3M)]	**[M]**			2			
3+3 layers							
[S2+(M)]	**[S2+(M)]**		3				
[S2+(2M)]	**[S2+(M)]**	2		2			
4+2 layers							
[S2+(M2)]	**[M]**	3					
[S2+(S2+(2M))]	**[M]**	3					
[S2+((S2+(M))+M)]	**[M]**	2					

TABLE II continues							
ID		**31**	**34**	**29**	**13**	**8**	**15**
Name		**Aleksi**	**Leo**	**Elisa**	**Joel**	**Aino**	**Mika**
MLU		**4,2**	**4,263**	**4,388**	**4,713**	**4,775**	**4,838**
2+2 layers							
[M]	**[M]**		234	2 4	3	23	34
[S2]	**[M]**	23	2	23	23	2	23
3+2 layers							
[M2]	**[M]**		3	23			4
[M2]	**[S2]**		3				
[S2+(M)]	**[M]**	3	23	23	2	23	234
[S2+(M)]	**[S2]**				2	2	4
[S2+(M)]	**[S3]**	2					
[S2+(2M)]	**[M]**		34		3	23	23
[S2+(2M)]	**[S2]**					2	
[S3+(2M)]	**[M]**						4
[S3+(3M)]	**[S2]**						4
3+3 layers							
[S2+(M)]	**[M2]**			4	3		
[S2+(M)]	**[S2+(M)]**	3			3	2	
[S2+(2M)]	**[S2+(M)]**					2	3
4+2 layers							
[M3]	**[M]**				3		
[S2+(M2)]	**[M]**	3					
[S2+(M2+M)]	**[M]**		4				
[S2+(S2+(M))]	**[M]**		2				
[S2+((S2+(M))+M)]	**[M]**						3
[S2+((S2+(M))+M)]	**[S3]**					4	

TABLE II continues							
ID		**24**	**38**	**10**	**23**	**11**	**19**
Name		**Sanni**	**Pekka**	**Santeri**	**Anniina**	**Jaakko**	**Tuomo**
MLU		**4,963**	**5,275**	**5,713**	**6,05**	**6,112**	**6,237**
2+2 layers							
[M]	**[M]**	23	234	234	34	34	34
[S2]	**[M]**	4	3	3	3	34	3
[S2]	**[S2]**				2	3	
3+2 layers							
[M2]	**[M]**	3	4	23		4	
[M2]	**[S2]**						2
[M2]	**[S3]**				4		
[S2+(M)]	**[M]**	234	234	23	4	34	234
[S2+(M)]	**[S2]**	3		3	3	3	
[S2+(2M)]	**[M]**		3	23	34	3	34
[S2+(2M)]	**[S2]**						4
[S2+(S2)]	**[M]**				2		
[S2+(S2+M)]	**[M]**	2					
3+3 layers							
[S2+(M)]	**[M2]**		3	3			
[S2+(M)]	**[S2+(M)]**	4					
[S2+(2M)]	**[S2+(2M)]**					2	
4+2 layers							
[M3]	**[M]**				3		
[S2+(M2)]	**[M]**						3
[S2+(M2+M)]	**[M]**						2
[S2+(S2+(M))]	**[M]**		2				
[S2+(S2+(2M))]	**[M]**						3
[S2+((S2+(M))+M)]	**[S3]**					4	
4+3 layers							
[S2+((S2+(2M))+M)]	**[S2+(M)]**						3
5+2 layers							
[S2+(M3+M)]	**[M]**			3			

TABLE II continues							
ID		**1**	**3**	**2**	**27**	**30**	**7**
Name		**Tuija**	**Laura**	**Riikka**	**Janna**	**Elina**	**Taru**
MLU		**6,25**	**6,25**	**6,287**	**6,3**	**6,325**	**6,35**
2+2 layers							
[M]	**[M]**	34	234	234	34	234	34
[S2]	**[M]**	3	3	3	3	4	34
[S2]	**[S2]**			3			
3+2 layers							
[M2]	**[M]**		23			234	34
[M2]	**[S2]**	2	3		3	3	
[S2+(M)]	**[M]**	234	234	23	234	23	234
[S2+(M)]	**[S2]**		3	3		3	24
[S2+(2M)]	**[M]**	34	3	34	23	3	234
[S2+(2M)]	**[S2]**	4					
[S3+(2M)]	**[M]**		2				
3+3 layers							
[M2]	**[M2]**					34	
[S2+(M)]	**[M2]**			3		3	
[S2+(M)]	**[S2+(M)]**					2	
(S2+(2M)]	**[M2]**			2			
[S2+(2M)]	**[S2+(M)]**			23			3
[S2+(2M)]	**[S2+(2M)]**		3	3			3
[S2+(S2)]	**[S2+(M)]**		2				
4+2 layers							
[M3]	**[M]**			3			
[S2+(M2)]	**[M]**	3					
[S2+(M2+M)]	**[M]**	2		4		4	
[S2+(S2+(2M))]	**[M]**	3		2			2
[S2+((S2+(M))+M)]	**[M]**					3	2
[S2+((S2+(2M))+M)]	**[M]**		3				
4+3 layers							
[S2+((S2+(2M))+M)]	**[S2+(M)]**	3					4
5+3 layers							
[S2+(M3)]	**[S2+(M)]**			2			

TABLE III Elaboration combinations when three components are elaborated in an utterance

2 = elaboration combination occurs in two-component utterances
3 = elaboration combination occurs in three-component utterances
4 = elaboration combination occurs in four-component utterances

ID			**35**	**22**	**28**	**36**	**17**	**6**
Name			**Seppo**	**Anna**	**Saku**	**Sampo**	**Tatu**	**Ronja**
MLU			**3,688**	**3,888**	**3,962**	**3,987**	**4,063**	**4,1**
2+2+2 layers								
[M]	**[M]**	**[M]**	34	3	3	34		
[S2]	**[M]**	**[M]**	3					
3+2+2 layers								
[M2]	**[M]**	**[M]**				3		
[S2+(M)]	**[M]**	**[M]**			4		3	

ID			31	34	29	13	8	15
Name			**Aleksi**	**Leo**	**Elisa**	**Joel**	**Aino**	**Mika**
MLU			**4,2**	**4,263**	**4,388**	**4,713**	**4,775**	**4,838**
2+2+2 layers								
[M]	**[M]**	**[M]**	4	4		3		4
[S2]	**[M]**	**[M]**	4			34		
3+2+2 layers								
[S2+(M)]	**[M]**	**[M]**		4			4	4
3+3+2 layers								
[S2+(M)]	**[M2]**	**[M]**	4					
[S2+(2M)]	**[M2]**	**[M]**						3
4+4+2 layers								
[S3+(M2+M)]	**[S2+(S2+(2M))]**	**[M]**	4					

TABLE III continues								
ID			**24**	**38**	**10**	**23**	**11**	**19**
Name			**Sanni**	**Pekka**	**Santeri**	**Anniina**	**Jaakko**	**Tuomo**
MLU			**4,963**	**5,275**	**5,713**	**6,05**	**6,112**	**6,237**
2+2+2 layers								
[M]	**[M]**	**[M]**	3			34	34	34
[S2]	**[M]**	**[M]**		3	34		4	34
[S2]	**[S2]**	**[M]**					3	
3+2+2 layers								
[S2+(M)]	**[M]**	**[M]**			3		3	34
[S2+(M)]	**[S2]**	**[M]**			4			
[S2+(2M)]	**[S2]**	**[M]**	4					
3+3+2 layers								
[S2+(M)]	**[S2+(M)]**	**[M]**					3	
[S2+(2M)]	**[M2]**	**[M]**						4
[S2+(S2)]	**[S2+(2M)]**	**[M]**						3
4+2+2 layers								
[S2+(M2+M)]	**[M]**	**[M]**						4
4+3+2 layers								
[S2+(M2+M)]	**[S2+(M)]**	**[M]**						3

TABLE III continues

ID			1	3	2	27	30	7
Name			**Tuija**	**Laura**	**Riikka**	**Janna**	**Elina**	**Taru**
MLU			**6,25**	**6,25**	**6,287**	**6,3**	**6,325**	**6,35**
2+2+2 layers								
[M]	**[M]**	**[M]**	4	34	34		34	
[S2]	**[M]**	**[M]**	34		3	3	3	
3+2+2 layers								
[M2]	**[M]**	**[M]**			3		3	
[S2+(M)]	**[M]**	**[M]**	34	34	3	34		
[S2+(2M)]	**[M]**	**[M]**	4		3			
3+3+2 layers								
[S2+(M)]	**[M2]**	**[M]**			3			
[S2+(M)]	**[S2+(M)]**	**[M]**			4			
[S2+(2M)]	**[M2]**	**[M]**			4			
[S2+(2M)]	**[S2+(M)]**	**[M]**	4					4
[S2+(2M)]	**[S2+(M)]**	**[S2]**	4					
[S2+(2M)]	**[S2+(2M)]**	**[M]**				4		
4+2+2 layers								
[S2+((S2+(M))+M)]	**[M]**	**[M]**		3				
4+3+2 layers								
[S2+(S2+(M))]	**[S2+(M)]**	**[M]**	4					
[S2+((S2+(M))+M)]	**[S2+(2M)]**	**[M]**		3				
[S2+((S2+(M))+M)]	**[S2+(S2+M)]**	**[M]**				3		

TABLE IV Elaboration combinations when four components are elaborated in an utterance

2 = elaboration combination occurs in two-component utterances
3 = elaboration combination occurs in three-component utterances
4 = elaboration combination occurs in four-component utterances

ID				**35**	**22**	**18**	**36**	**17**	**6**
Name				**Seppo**	**Anna**	**Saku**	**Sampo**	**Tatu**	**Ronja**
MLU				**3,688**	**3,888**	**3,962**	**3,987**	**4,063**	**4,1**
2+2+2+2 layers									
[M]	**[M]**	**[M]**	**[M]**						4
ID				**31**	**34**	**29**	**13**	**8**	**15**
Name				**Aleksi**	**Leo**	**Elisa**	**Joel**	**Aino**	**Mika**
MLU				**4,2**	**4,263**	**4,388**	**4,713**	**4,775**	**4,838**
				No utterances with four components elaborated					
ID				**24**	**38**	**10**	**23**	**11**	**19**
Name				**Sanni**	**Pekka**	**Santeri**	**Anniina**	**Jaakko**	**Tuomo**
MLU				**4,963**	**5,275**	**5,713**	**6,05**	**6,112**	**6,237**
				No utterances with four components elaborated					
ID				**1**	**3**	**2**	**27**	**30**	**7**
Name				**Tuija**	**Laura**	**Riikka**	**Janna**	**Elina**	**Taru**
MLU				**6,25**	**6,25**	**6,287**	**6,3**	**6,325**	**6,35**
2+2+2+2 layers									
[M]	**[M]**	**[M]**	**[M]**		4				
3+2+2+2 layers									
[M2]	**[M]**	**[M]**	**[M]**						4
3+3+2+2 layers									
[S2+(2M)]	**[S2]**	**[M]**	**[M]**						4

APPENDIX 8 Inventory of elaboration types in the risk and control group (ordered by MLU)

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

TABLE I The component inventory of the risk children

ID	**35**	**22**	**36**	**31**	**34**	**29**
Name	**Seppo**	**Anna**	**Sampo**	**Aleksi**	**Leo**	**Elisa**
MLU	**3,688**	**3,888**	**3,987**	**4,2**	**4,263**	**4,388**
Two-layered components						
[M]	234	234	234	234	234	234
[S2]	23	234		234	234	23
[S3]				2		
Three-layered components						
[M2]	23		23	34	34	234
[S2+(M)]	23	234	23	234	23	234
[S2+(2M)]	2	2	23	23	34	2
[S2+(S2)]						
[S2+(S2+M)]						
[S3+(M)]						
[S3+(2M)]		3				
[S3+(S2+2M)]						
Four-layered components						
[M3]						
[S2+(M2)]	23			23		
[S2+(M2+M)]					4	
[S2+(2M2)]			3			
[S2+(S2+(M))]					2	2
[S2+(S2+(2M))]	3	2		4		2
[S2+((S2+(M))+M)]	2					
[S2+((S2+(2M))+M)]						
[S2+(S3+(3M))]				3		
[S3+(M2+M)]				4		
[S3+(M2+2M)]						
[S3+((S2+(M))+M)]						
[S3+((S2+(2M))+M)]						
Five-layered components						
[S2+(S2+(M2+M))]						
Total	**8**	**6**	**5**	**9**	**7**	**7**

TABLE I continues					
ID	**24**	**38**	**23**	**27**	**30**
Name	**Sanni**	**Pekka**	**Anniina**	**Janna**	**Elina**
MLU	**4,963**	**5,275**	**6,05**	**6,3**	**6,325**
Two-layered components					
[M]	234	234	234	234	234
[S2]	34	3	234	34	34
[S3]			4		
Three-layered components					
[M2]	3	34	234	4	234
[S2+(M)]	234	234	34	234	234
[S2+(2M)]	34	3	234	23	34
[S2+(S2)]			2		4
[S2+(S2+M)]	2			3	
[S3+(M)]				3	
[S3+(2M)]					
[S3+(S2+2M)]			3		
Four-layered components					
[M3]			3		
[S2+(M2)]			4		
[S2+(M2+M)]					234
[S2+(2M2)]					
[S2+(S2+(M))]	2	2	2		3
[S2+(S2+(2M))]					
[S2+((S2+(M))+M)]				3	
[S2+((S2+(2M))+M)]					3
[S2+(S3+(3M))]					
[S3+(M2+M)]					
[S3+(M2+2M)]	3				
[S3+((S2+(M))+M)]					3
[S3+((S2+(2M))+M)]			3		
Five-layered components					
[S2+(S2+(M2+M))]				23	
Total	**8**	**6**	**12**	**9**	**10**

TABLE II The component inventory of the control children

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

ID	**18**	**17**	**6**	**13**	**8**	**15**	**10**
Name	**Saku**	**Tatu**	**Ronja**	**Joel**	**Aino**	**Mika**	**Santeri**
MLU	**3,962**	**4,063**	**4,1**	**4,713**	**4,775**	**4,838**	**5,713**
Two-layered components							
[M]	234	234	234	234	234	234	234
[S2]	23	23	234	234	23	234	34
[S3]					4		
Three-layered components							
[M2]				3	3	234	234
[S2+(M)]	234	234	23	23	23	234	234
[S2+(2M)]	23	3	2	23	23	23	23
[S2+(S2)]			2				
[S2+(S2+M)]							
[S3+(M)]	2		2				
[S3+(2M)]					2	4	
[S3+(3M)]				3			
Four-layered components							
[M3]				23			
[S2+(M2)]				2	3		3
[S2+(M2+M)]							
[S2+(S2+(M))]	2		2	3		3	
[S2+(S2+(2M))]							
[S2+((S2+(M))+M)]	2	2			4	3	
[S2+((S2+(2M))+M)]							
[S3+((S2+(M))+M)]							
[S3+((S2+(2M))+M)]							
Five-layered components							
[S2+(M3)]							
[S2+(M3+M)]							34
[S2+(S2+(M2))]							
Total	**7**	**5**	**7**	**9**	**9**	**8**	**7**

TABLE II continues						
ID	**11**	**19**	**1**	**3**	**2**	**7**
Name	**Jaakko**	**Tuomo**	**Tuija**	**Laura**	**Riikka**	**Taru**
MLU	**6,112**	**6,237**	**6,25**	**6,25**	**6,287**	**6,35**
Two-layered components						
[M]	234	234	234	234	234	234
[S2]	34	234	34	34	3	234
[S3]	4					
Three-layered components						
[M2]	4	234	234	3	234	34
[S2+(M)]	34	234	234	234	234	234
[S2+(2M)]	23	234	34	234	234	234
[S2+(S2)]		3	2	34		
[S2+(S2+M)]					3	
[S3+(M)]						
[S3+(2M)]			2			
[S3+(3M)]	4					
Four-layered components						
[M3]				3		
[S2+(M2)]	3	3				
[S2+(M2+M)]		234		4	23	
[S2+(S2+(M))]		34	24		2	
[S2+(S2+(2M))]		3		2		2
[S2+((S2+(M))+M)]				3		2
[S2+((S2+(2M))+M)]					3	
[S3+((S2+(M))+M)]						4
[S3+((S2+(2M))+M)]		3	3			
Five-layered components						
[S2+(M3)]				2		
[S2+(M3+M)]						
[S2+(S2+(M2))]					1	
Total	**8**	**11**	**9**	**11**	**10**	**8**

APPENDIX 9 Inventories of elaboration combinations in utterances produced by the risk and control children (ordered by MLU values)

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

TABLE I Inventory of the elaboration types when only one component is elaborated						
Risk children						
ID	**35**	**22**	**36**	**31**	**34**	**29**
Name	**Seppo**	**Anna**	**Sampo**	**Aleksi**	**Leo**	**Elisa**
MLU	**3,688**	**3,888**	**3,987**	**4,2**	**4,263**	**4,388**
Two-layered components						
[M]	234	234	234	234	234	234
[S2]		4		23	4	23
Three-layered components						
[M2]	2		23	3	34	34
[S2+(M)]	2	234	23	23	3	23
[S2+(2M)]	2	2	3	23	3	2
[S2+(S2)]						
[S3+(M)]						
[S3+(2M)]		3				
[S3+(S2+2M)]						
Four-layered components						
[S2+(M2)]	2			2		
[S2+(M2+M)]						
[S2+(2M2)]			3			
[S2+(S2+(M))]						2
[S2+(S2+(2M))]		2				2
[S2+(S3+(3M))]				3		
[S3+(M2+2M)]						
[S3+((S2+(2M))+M)]						
Five-layered components						
[S2+(S2+(M2+M))]						
Total number of types	**5**	**6**	**5**	**7**	**5**	**7**

TABLE I continues					
Risk children					
ID	**24**	**38**	**23**	**27**	**30**
Name	**Sanni**	**Pekka**	**Anniina**	**Janna**	**Elina**
MLU	**4,963**	**5,275**	**6,05**	**6,3**	**6,325**
Two-layered components					
[M]	234	34	34	34	34
[S2]	3	3	34	34	3
Three-layered components					
[M2]	3	3	23	4	234
[S2+(M)]	23	34	3	234	4
[S2+(2M)]	3	3	23	23	34
[S2+(S2)]					4
[S3+(M)]				3	
[S3+(2M)]					
[S3+(S2+2M)]			3		
Four-layered components					
[S2+(M2)]			4		
[S2+(M2+M)]					23
[S2+(2M2)]					
[S2+(S2+(M))]	2		2		3
[S2+(S2+(2M))]					
[S2+(S3+(3M))]					
[S3+(M2+2M)]	3				
[S3+((S2+(2M))+M)]			3		
Five-layered components					
[S2+(S2+(M2+M))]				23	
Total number of types	**7**	**5**	**9**	**7**	**8**

TABLE I continues							
<u>**Control group**</u>							
ID	**18**	**17**	**6**	**13**	**8**	**15**	**10**
Name	**Saku**	**Tatu**	**Ronja**	**Joel**	**Aino**	**Mika**	**Santeri**
MLU	**3,962**	**4,063**	**4,1**	**4,713**	**4,775**	**4,838**	**5,713**
Two-layered components							
[M]	23	234	234	234	234	234	234
[S2]	23	23	2	23	23	23	34
[S3]							
Three-layered components							
[M2]				3	3	23	34
[S2+(M)]	2	23	23	23	23	2	234
[S2+(2M)]	23		2	2	3	23	
[S2+(S2)]			2				
[S3+(M)]			2				
[S3+(2M)]					2		
[S3+(3M)]				3			
Four-layered components							
[M3]				2			
[S2+(M2)]				2	3		3
[S2+(M2+M)]							
[S2+(S2+(M))]	2		2				
[S2+((S2+(M))+M)]		2					
Five-layered components							
[S2+(M3+M)]							4
Total number of types	**5**	**4**	**7**	**8**	**7**	**5**	**6**

TABLE I continues						
Control group						
ID	**11**	**19**	**1**	**3**	**2**	**7**
Name	**Jaakko**	**Tuomo**	**Tuija**	**Laura**	**Riikka**	**Taru**
MLU	**6,112**	**6,237**	**6,25**	**6,25**	**6,287**	**6,35**
Two-layered components						
[M]	234	34	34	34	34	34
[S2]	34			4		34
[S3]	4					
Three-layered components						
[M2]		3	4		23	
[S2+(M)]	34	23	34	34	3	34
[S2+(2M)]	3	23	3	23	23	3
[S2+(S2)]				34		
[S3+(M)]						
[S3+(2M)]			2			
[S3+(3M)]						
Four-layered components						
[M3]						
[S2+(M2)]	3					
[S2+(M2+M)]					23	
[S2+(S2+(M))]		34	2		2	
[S2+((S2+(M))+M)]						
Five-layered components						
[S2+(M3+M)]						
Total number of types	**6**	**5**	**6**	**5**	**6**	**4**

TABLE II Inventory of elaboration combinations when two components are elaborated

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

Risk children							
ID		**35**	**22**	**36**	**31**	**34**	**29**
Name		**Seppo**	**Anna**	**Sampo**	**Aleksi**	**Leo**	**Elisa**
MLU		**3,688**	**3,888**	**3,987**	**4,2**	**4,263**	**4,388**
2+2 layers							
[M]	[M]	23	234	234		234	2 4
[S2]	[M]	2	23		23	2	23
[S2]	[S2]						
3+2 layers							
[M2]	[M]	3		23		3	23
[M2]	[S2]					3	
[M2]	[S3]						
[S2+(M)]	[M]	23	23		3	23	23
[S2+(M)]	[S2]		3				
[S2+(M)]	[S3]				2		
[S2+(2M)]	[M]			2		34	
[S2+(S2)]	[M]						
[S2+(S2+M)]	[M]						
3+3 layers							
[M2]	[M2]						
[S2+(M)]	[M2]						4
[S2+(M)]	[S2+(M)]		3		3		
[S2+(2M)]	[S2+(M)]	2					
4+2 layers							
[M3]	[M]						
[S2+(M2)]	[M]	3			3		
[S2+(M2+M)]	[M]					4	
[S2+(S2+(M))]	[M]					2	
[S2+(S2+(2M))]	[M]	3					
[S2+((S2+(M))+M)]	[M]	2					
Total number of combinations		**8**	**5**	**3**	**5**	**8**	**5**

TABLE II continues

Risk children						
ID		**24**	**38**	**23**	**27**	**30**
Name		**Sanni**	**Pekka**	**Anniina**	**Janna**	**Elina**
MLU		**4,963**	**5,275**	**6,05**	**6,3**	**6,325**
2+2 layers						
[M]	[M]	23	234	34	34	234
[S2]	[M]	4	3	3	3	4
[S2]	[S2]			2		
3+2 layers						
[M2]	[M]	3	4			234
[M2]	[S2]				3	3
[M2]	[S3]			4		
[S2+(M)]	[M]	234	234	4	234	23
[S2+(M)]	[S2]	3		3		3
[S2+(M)]	[S3]					
[S2+(2M)]	[M]		3	34	23	3
[S2+(S2)]	[M]			2		
[S2+(S2+M)]	[M]	2				
3+3 layers						
[M2]	[M2]					34
[S2+(M)]	[M2]		3			3
[S2+(M)]	[S2+(M)]	4				2
[S2+(2M)]	[S2+(M)]					
4+2 layers						
[M3]	[M]			3		
[S2+(M2)]	[M]					
[S2+(M2+M)]	[M]					4
[S2+(S2+(M))]	[M]		2			
[S2+(S2+(2M))]	[M]					
[S2+((S2+(M))+M)]	[M]					3
Total number of combinations		**7**	**7**	**9**	**5**	**12**

TABLE II continues

Control children								
ID		**18**	**17**	**6**	**13**	**8**	**15**	**10**
Name		**Saku**	**Tatu**	**Ronja**	**Joel**	**Aino**	**Mika**	**Santeri**
MLU		**3,962**	**4,063**	**4,1**	**4,713**	**4,775**	**4,838**	**5,713**
2+2 layers								
[M]	[M]	234	234	234	3	23	34	234
[S2]	[M]	23	3	234	23	2	23	3
[S2]	[S2]							
3+2 layers								
[M2]	[M]						4	23
[M2]	[S2]							
[S2+(M)]	[M]	23	34	2	2	23	234	23
[S2+(M)]	[S2]				2	2	4	3
[S2+(2M)]	[M]		3	2	3	23	23	23
[S2+(2M)]	[S2]					2		
[S3+(2M)]	[M]						4	
[S3+(3M)]	[M]	2						
[S3+(3M)]	[S2]						4	
3+3 layers								
[S2+(M)]	[M2]				3			3
[S2+(M)]	[S2+(M)]				3	2		
(S2+(2M)]	[M2]							
[S2+(2M)]	[S2+(M)]	2				2	3	
[S2+(2M)]	[S2+(2M)]							
[S2+(S2)]	[S2+(M)]							
[S2+(S2+M)]	[S2+(2M)]							
4+2 layers								
[M3]	[M]				3			
[S2+(M2)]	[M]							
[S2+(M2+M)]	[M]							
[S2+(S2+(M))]	[M]							
[S2+(S2+(2M))]	[M]							
[S2+((S2+(M))+M)]	[M]						3	
[S2+((S2+(M))+M)]	[S3]					4		
[S2+((S2+(2M))+M)]	[M]							
4+3 layers								
[S2+((S2+(2M))+M)]	[S2+(M)]							
5+2 layers								
[S2+(M3+M)]	[M]							3
[S2+(S2+(M2)]	[M]							
5+3 layers								
[S2+(M3)]	[S2+(M)]							
Total number of combinations		**5**	**4**	**4**	**8**	**9**	**10**	**8**

TABLE II continues

Control children							
ID		**11**	**19**	**1**	**3**	**2**	**7**
Name		**Jaakko**	**Tuomo**	**Tuija**	**Laura**	**Riikka**	**Taru**
MLU		**6,112**	**6,237**	**6,25**	**6,25**	**6,287**	**6,35**
2+2 layers							
[M]	[M]	34	34	234	234	34	34
[S2]	[M]	34	3	3	3	3	34
[S2]	[S2]	3			3		
3+2 layers							
[M2]	[M]	4		23		2 4	34
[M2]	[S2]		2	3		3	
[S2+(M)]	[M]	34	234	234	23	23	234
[S2+(M)]	[S2]	3		3	3	3	24
[S2+(2M)]	[M]	3	34	3	34	23	234
[S2+(2M)]	[S2]		4				
[S3+(2M)]	[M]			2			
[S3+(3M)]	[M]						
[S3+(3M)]	[S2]						
3+3 layers							
[S2+(M)]	[M2]				3	3	
[S2+(M)]	[S2+(M)]					4	
(S2+(2M)]	[M2]				2		
[S2+(2M)]	[S2+(M)]				23		3
[S2+(2M)]	[S2+(2M)]	2		3	3		3
[S2+(S2)]	[S2+(M)]			2			
[S2+(S2+M)]	[S2+(2M)]					3	
4+2 layers							
[M3]	[M]				3		
[S2+(M2)]	[M]		3				
[S2+(M2+M)]	[M]		2		4		
[S2+(S2+(M))]	[M]					2	
[S2+(S2+(2M))]	[M]		3		2		2
[S2+((S2+(M))+M)]	[M]						2
[S2+((S2+(M))+M)]	[S3]	4					
[S2+((S2+(2M))+M)]	[M]			3			
4+3 layers							
[S2+((S2+(2M))+M)]	[S2+(M)]		3				4
5+2 layers							
[S2+(M3+M)]	[M]						
[S2+(S2+(M2)]	[M]					2	
5+3 layers							
[S2+(M3)]	[S2+(M)]				2		
Total number of combinations		**9**	**10**	**11**	**14**	**12**	**11**

TABLE III Inventory of elaboration combinations when three components are elaborated in an utterance

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

Risk children								
ID			**35**	**22**	**36**	**31**	**34**	**29**
Name			**Seppo**	**Anna**	**Sampo**	**Aleksi**	**Leo**	**Elisa**
MLU			**3,688**	**3,888**	**3,987**	**4,2**	**4,263**	**4,388**
2+2+2 layers								
[M]	[M]	[M]	34	3	34	4	4	
[S2]	[M]	[M]	3			4		
3+2+2 layers								
[M2]	[M]	[M]			3			
[S2+(M)]	[M]	[M]					4	
[S2+(2M)]	[S2]	[M]						
3+3+2 layers								
[S2+(M)]	[M2]	[M]				4		
[S2+(2M)]	[S2+(2M)]	[M]						
[S2+(S2)]	[S2+(2M)]	[M]						
4+3+2 layers								
[S2+((S2+(M))+M)]	[S2+(S2+M)]	[M]						
4+4+2 layers								
[S3+(M2+M)]	[S2+(S2+(2M))]	[M]				4		
Total number of combinations			**2**	**1**	**2**	**4**	**2**	**0**

TABLE III continues

Risk children							
ID			**24**	**38**	**23**	**27**	**30**
Name			**Sanni**	**Pekka**	**Anniina**	**Janna**	**Elina**
MLU			**4,963**	**5,275**	**6,05**	**6,3**	**6,325**
2+2+2 layers							
[M]	[M]	[M]	3		34		34
[S2]	[M]	[M]		3		3	3
3+2+2 layers							
[M2]	[M]	[M]					3
[S2+(M)]	[M]	[M]				34	
[S2+(2M)]	[S2]	[M]	4				
3+3+2 layers							
[S2+(M)]	[M2]	[M]					
[S2+(2M)]	[S2+(2M)]	[M]				4	
[S2+(S2)]	[S2+(2M)]	[M]					
4+3+2 layers							
[S2+((S2+(M))+M)]	[S2+(S2+M)]	[M]				3	
4+4+2 layers							
[S3+(M2+M)]	[S2+(S2+(2M))]	[M]					
Total number of combinations			**2**	**1**	**1**	**4**	**3**

TABLE III continues

Control group									
ID			**18**	**17**	**6**	**13**	**8**	**15**	**10**
Name			**Saku**	**Tatu**	**Ronja**	**Joel**	**Aino**	**Mika**	**Santeri**
MLU			**3,962**	**4,063**	**4,1**	**4,713**	**4,775**	**4,838**	**5,713**
2+2+2 layers									
[M]	[M]	[M]	3			3		4	
[S2]	[M]	[M]				34			34
[S2]	[S2]	[M]							
3+2+2 layers									
[M2]	[M]	[M]							
[S2+(M)]	[M]	[M]	4	3			4	4	3
[S2+(M)]	[S2]	[M]							4
[S2+(2M)]	[M]	[M]							
3+3+2 layers									
[S2+(M)]	[M2]	[M]							
[S2+(M)]	[S2+(M)]	[M]							
[S2+(2M)]	[M2]	[M]						3	
[S2+(2M)]	[S2+(M)]	[M]							
[S2+(2M)]	[S2+(M)]	[S2]							
[S2+(S2)]	[S2+(2M)]	[M]							
4+2+2 layers									
[S2+(M2+M)]	[M]	[M]							
[S2+((S2+(M))+M)]	[M]	[M]							
4+3+2 layers									
[S2+(M2+M)]	[S2+(M)]	[M]							
[S2+(S2+(M))]	[S2+(M)]	[M]							
[S2+((S2+(M))+M)]	[S2+(2M)]	[M]							
Total number of components			**2**	**1**	**0**	**2**	**1**	**3**	**3**

TABLE III continues

Control group								
ID			**11**	**19**	**1**	**3**	**2**	**7**
Name			**Jaakko**	**Tuomo**	**Tuija**	**Laura**	**Riikka**	**Taru**
MLU			**6,112**	**6,237**	**6,25**	**6,25**	**6,287**	**6,35**
2+2+2 layers								
[M]	[M]	[M]	34	34	4	34	34	
[S2]	[M]	[M]	4	34	34		3	
[S2]	[S2]	[M]	3					
3+2+2 layers								
[M2]	[M]	[M]					3	
[S2+(M)]	[M]	[M]	3	34	34	34	3	
[S2+(M)]	[S2]	[M]						
[S2+(2M)]	[M]	[M]			4		3	
3+3+2 layers								
[S2+(M)]	[M2]	[M]					3	
[S2+(M)]	[S2+(M)]	[M]	3				4	
[S2+(2M)]	[M2]	[M]		4			4	
[S2+(2M)]	[S2+(M)]	[M]			4			4
[S2+(2M)]	[S2+(M)]	[S2]			4			
[S2+(S2)]	[S2+(2M)]	[M]		3				
4+2+2 layers								
[S2+(M2+M)]	[M]	[M]		4				
[S2+((S2+(M))+M)]	[M]	[M]				3		
4+3+2 layers								
[S2+(M2+M)]	[S2+(M)]	[M]		3				
[S2+(S2+(M))]	[S2+(M)]	[M]			4			
[S2+((S2+(M))+M)]	[S2+(2M)]	[M]				3		
Total number of components			**5**	**7**	**7**	**4**	**8**	**1**

TABLE IV Inventory of elaboration combinations when four components are elaborated in utterances

2 = the component type occurs in 2-component utterances
3 = the component type occurs in 3-component utterances
4 = the component type occurs in 4-component utterances

Risk children						
ID	**35**	**22**	**36**	**31**	**34**	**29**
Name	**Seppo**	**Anna**	**Sampo**	**Aleksi**	**Leo**	**Elisa**
MLU	**3,69**	**3,9**	**3,99**	**4,2**	**4,263**	**4,388**
Total	**0**	**0**	**0**	**0**	**0**	**0**

ID	**24**	**38**	**23**	**27**	**30**
Name	**Sanni**	**Pekka**	**Anniina**	**Janna**	**Elina**
MLU	**4,963**	**5,275**	**6,05**	**6,3**	**6,325**
Total	**0**	**0**	**0**	**0**	**0**

Control group										
ID				**18**	**17**	**6**	**13**	**8**	**15**	**10**
Name				**Saku**	**Tatu**	**Ronja**	**Joel**	**Aino**	**Mika**	**Santeri**
MLU				**3,962**	**4,063**	**4,1**	**4,713**	**4,775**	**4,838**	**5,713**
2+2+2+2										
[M]	[M]	[M]	[M]			4				
3+2+2+2										
[M2]	[M]	[M]	[M]							
3+3+2+2										
[S2+(2M)]	[S2]	[M]	[M]							
Total				**0**	**0**	**1**	**0**	**0**	**0**	**0**

TABLE IV continues

Control group									
ID				**11**	**19**	**1**	**3**	**2**	**7**
Name				**Jaakko**	**Tuomo**	**Tuija**	**Laura**	**Riikka**	**Taru**
MLU				**6,112**	**6,237**	**6,25**	**6,25**	**6,287**	**6,35**
2+2+2+2									
[M]	[M]	[M]	[M]				4		
3+2+2+2									
[M2]	[M]	[M]	[M]						4
3+3+2+2									
[S2+(2M)]	[S2]	[M]	[M]						4
Total				**0**	**0**	**0**	**1**	**0**	**2**