Language Growth in Dutch School-Age Children with Specific Language Impairment

Published by

LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration:

Glass fusion object (detail) by Elly Rosseel (www.elly-rosseel.nl)

ISBN: 978-94-6093-138-3

NUR 616

Language Growth in Dutch School-Age Children with Specific Language Impairment

Taalgroei van Nederlandse Basisschoolkinderen met een Primaire Taalontwikkelingsstoornis

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
maandag 2 juni 2014 des middags te 14.30 uur

door

Robert Lodewijk Maria Zwitserlood

geboren op 29 mei 1959 te Schaesberg

Promotoren: Prof. Dr. F.N.K. Wijnen
Prof. Dr. L. Verhoeven
Copromotor: Dr. M. van Weerdenburg

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Koninklijke Auris Groep.

Contents

Acknowledgments

First and foremost, I would like to thank the Koninklijke Auris Groep for making this research project possible, especially the former board of directors, Han van Beers and Theo van Munnen, and the present board of directors, Dick de Boer. I am also very thankful to Marjan Bruins, who has been my manager for most of this period. Marjan, I know that you have played a very important role in making this project happen and you have greatly supported me all these years. I also want to thank my current manager Erik van Lingen, all my past and present colleagues at Auris Centraal Bureau in Gouda, and especially my Auris K&D colleagues for their warm support and interest in my work.

I also want to thank the UiL OTS: Martin Everaert, Henriette de Swart, and Maaike Schoorlemmer for being such wonderful hosts. Although I started as an external PhD student, you welcomed me to participate in the UiL OTS graduate program and made me feel more an internal than an external PhD student. Thank you for the opportunities to visit the LOT schools and international conferences, to organize workshops and attend PhD courses. I also want to thank the UiL OTS secretaries for all their help: Yvonne van Adrichem, Mariëtte Bonenkamp, Martien Camphuis, and Martine Paulissen.

I have been truly blessed with my three supervisors, my first promotor Frank Wijnen, my second promotor Ludo Verhoeven, and my co-promotor Marjolijn van Weerdenburg. Frank, you inspired me to follow my intuitions and you were always interested in my own ideas on SLI. You sharpened my ideas by asking me the right questions and encouraged me not to be too cautious in my interpretations and conclusions. I benefited greatly from your amazingly vast knowledge on language acquisition, SLI, and speech fluency. When times got rough and I got lost in the huge amount of data, you were able bring the focus back again. Ludo, you maintained a clear view of the whole project and always came up with creative titles for a chapter or a section. I have learned a lot from your strategic writing skills. Your suggestions for shifting a paragraph up or down or adding a well-chosen sentence really added to the quality of the chapters. Marjolijn, you are definitely the best co-promotor anyone could wish for. You generously provided me with the SLI data from your own research project. I thoroughly enjoyed our meetings in Nijmegen and Utrecht and your great help and support have been very important for me. You were my guide in statistics and you are also a walking APA manual! Your unremitting enthusiasm and positive feedback often made my day.

During my stay at the UiL OTS I got to know many wonderful, bright and friendly people. I started at Achter de Dom 22/24, where Jacqueline van Kampen welcomed me on my first day. Jacqueline, you were a great help with the Childes programs and all my questions on language acquisition. Together we supervised my first master student and I really learned quite a lot from you. It was a special atmosphere at Achter de Dom, with Paola Monachesi, Sharon Unsworth, Kiki Kushartanti, Jacub Dotlacil, various foreign guests, and the occasional PhD students living in the basement to finish their thesis undisturbed. At the beginning of my project, I also participated in the Lab meetings, where I met Maartje de Klerk, Annemarie Kerkhoff, Desiree Capel, and the amazing Elise de Bree. Elise, thank you for reading and commenting on some of my chapters and your humour and enthusiasm. Our meetings and coffees were always so delightful and always much too short.

The past five years, I have also been very lucky to have such wonderful roommates, in order of appearance: Gaetano Fiorin, Eline Westerhout, Thomas Marcus, Lizet van Ewijk, Brigitta Keij, Eva Poortman, and Yipu Wei. I want to thank you all for your companionship, work- and non-work related discussions, fun and laughs, coffees and cookies. Thanks also to the members of my intervision group (André Krom, Cynthia Han, Daria Bahtina, Liv Persson, Marko Simonovic, Mulusew Wondem, Rogier Kraf, and Tom Wang). In our intervision meetings, we could share our PhD problems and discuss our research projects and life in general (pets, weddings, movies, food, politics, cycling, and music).

Moving from Achter de Dom 22/24 to Trans 2, brought me to the crossroads of the Korte Nieuwstraat (Nederlands department) and the Trans (UiL OTS). I really loved working at the corner room 0.51 with so many great neighbours. There even was a connection with the upper floors through the scary spiral staircase. Sometimes people would just fall down the stairs and land on our doorstep. Thanks to all (PhD) colleagues for their interest in my work and for making my stay at the UiL OTS fun: Elena Tribushinina, Pim Mak, Nivja de Jong, Huub van den Bergh, Assaf Toledo, Desiree Capel, Marjolein van Egmond, Anna Sara Romøren, Bert le Bruyn, Xiaoli Dong, Marijn Struiksma, Hannah de Mulder, Jacqueline Evers-Vermeul, Jacomine Nortier, Anneloes Canestrelli, Loes Koring, Marko Hladnik, Carolien van den Hazelkamp, Dagmar Schadler, Ileana Grama, Bettina Gruber, Alexia Guerra Rivera, Mirjam Hachem, Sander van der Harst, Anne-France Pinget, Anna Volkova, Hanna de Vries, Sophia Manika, Rianne Schippers, Nadya Goldberg, Anna Chernilovskaya, Liquan Liu, Arnout Koorneef, Hans Rutger Bosker (who helped me quit smoking), Sandrien van

Ommen, Choonkyu Lee, Anne van Leeuwen, Tom Lentz, Arnout Koorneef, Sergey Avrutin, Hugo Quené, Sergio Baauw, and everyone else.

So many people outside the institute helped and advised me throughout the years. Mary Hirschman's paper on metalinguistic therapy for older children with SLI sparked my interest in metalinguistic intervention approaches. Mary, our email contact has become very dear and special to me, and I really hope we'll meet in person one day. Kirsten Thyme, you kindly provided me with all the materials you still had from the Lego therapy program 'Grammar in Form and Color'. Your program was the basis for my 'Metataal' intervention study. I also want to thank Corinna Sipkema for sending me her thesis on the Dutch version of the Lego therapy. Pauline Frizelle kindly provided me with information on her relative clause tests for children with SLI. I also want to thank Gerard Bol, Jan de Jong, Judith Rispens, and Evelien Krikhaar for their warm support and interest in my work.

This research project would have been impossible without the help of all the children, parents, staff and schools that participated. I want to thank the Howiblo in Montfoort, the Wilde Wingerd in Maarssen, the Paulusschool in Utrecht, and the Regenboogschool in Utrecht for their cooperation and hospitality. I also want to thank my UU students for their help with the transcriptions of the narratives: Irene Sormani, Marjolein van der Horst, Henaly Leijenhorst, Marij van Ewijk, and Merel van Witteloostuijn. Marian Strik from the Radboud University did a great job on the content scores of the narratives. The staff at the Opleiding Logopedie, Hogeschool Utrecht (Mieke Stumpel, Esther van Niel, Leonoor Oonk) and students (Liset Bergevoet, Hanneke Creemers, Jonna Genuit, and Nadia Lemouesset) were a great help in developing the tests and therapy program of the Metataal intervention study. I want to thank the parents, children and staff of Auris De Taalkring in Utrecht for making this intervention study possible. A special thanks goes to Marietje van Berkel for coordination, Evelien Burema and Lieke Kuipers for testing and providing the Metataal intervention, and Brigitta Keij and Margo Zwitserlood for testing the children.

I also want to express my gratitude to the members of my dissertation assessment committee: Richard Schwartz, Anne Baker, Hans van Balkom, Peter Coopmans, and Ellen Gerrits for their time and their valuable comments on the manuscript. Thank you Brigitta Keij and Wim Tersteeg for being my amazing paranymphs.

And finally, of course, thank you my sweet and lovely Margo for everything.

1 General introduction

Primary language acquisition is an amazingly rapid and robust process. By four years of age, a typically developing child can speak clearly and intelligibly, produce long and complex sentences and has a vocabulary already containing some thousands of words. However, many children do not learn their native language effortlessly and are delayed in their speech and/or language development. Some children overcome these problems, but others have persistent language difficulties. Follow-up studies of children with severe early language delays show that these children are at risk of later literacy problems, lack of friendships, unemployment, and higher rates of psychiatric disorder (Clegg, Hollis, Mawhood, & Rutter, 2005). However, the amount of research focusing on these speech and language disorders is relatively small, in spite of the fact that they are very common and have an enormous impact on individuals and society (Bishop, 2010). One might say that people with persistent language impairment never become native speakers in their mother tongue. Speech and communication will always remain burdensome, as illustrated with a brief case study, in which a 19-year-old young man tells us about his language problems:

Ik … ik heb gerecepteerd[a] …wie ik ben. Vroeger niet, maar nu wel. Ehm…het kan z-zijn dat ik problemen kan krijgen qua eh taalprobleem natuurlijk. Want bepaalde post als ik lees, snap ik het niet. Ga ik naar mijn moeder toe, of naar mijn vriendin toe, of naar een vriend toe van… hoe … hoe gaat het eigenlijk?[….]Dan kan ik in een…eh…kan die zin zo begrijpen en ik kan het ook anders begrijpen en dat weet ik soms niet. Dus daarom vraag ik ook …wie er omheen is. Vraag ik het na. Ja. Maar een taalprobleem blijf je altijd houden.

I … I have recepted[a] …who I am. Not in the past, but now I do. Uhm…it can b-be that I can get problems qua uh language problem of course. Because certain mail when I read, I don't understand it. I go to my mother, or to my girlfriend, or to a friend…how … how goes this actually? [….] Then I can in a…uh…can understand that sentence this way and I can also understand it in another way and sometimes I don't know which. So therefore I ask … who is around. I check it. Yes. But a language problem is something you always keep.
Note: [a]*gerecepteerd*/recepted is probably a blend of respected/accepted (*respecteren/accepteren*).

This dissertation is about children with specific language impairment (SLI). These children show significant limitations in language abilities, but do not have evident conditions that could explain their language problems, such as mental retardation, hearing loss, or neurological damage. Researchers have been trying to answer the question what causes SLI and how it can be prevented or remedied. If a better understanding of the causal mechanisms of SLI can be reached, we might also be able to design more effective methods of assessment, treatment and prevention. However, there are also theoretical reasons for studying SLI, as it can offer a window onto the processes and mechanisms underlying normal language acquisition. Research on SLI has nurtured debates and theories about modularity and innateness on the one hand, and usage-based approaches and statistical learning on the other hand.

The present thesis aims to theoretically validate processes of language development in Dutch children with SLI in perspective of clinical applications. The grammatical development of Dutch school-age children with SLI is examined longitudinally between the ages 6 and 10 years. To date, little is known about the grammatical development of Dutch children with SLI in the primary school years, especially in the middle and higher grades. In addition, the development of speech disfluencies is studied in the older (i.e., 8- to 10-year old) school-age children with SLI. Although these older children have overcome some of their language difficulties, sentence formulation may still be challenging, leading to higher disfluency rates, even when grammatical accuracy is high (Finneran, Leonard, & Miller, 2009). Furthermore, prospects for language intervention in older school-age children with SLI are investigated. In Dutch special schools for children with SLI, the bulk of speech therapy resources are allocated to the young children (up to Grade 3). The general opinion is that therapy for older school-age children is not very beneficial anymore. These opinions are based on (unsubstantiated) notions of critical or sensitive periods for language development in children with SLI. However, a growing body of research suggests that older school-age children with SLI can still benefit from intervention, but with approaches different from those commonly used with younger children with SLI. In general, older school-age children with SLI have overcome their phonological and lexical difficulties to a certain extent. Metalinguistic intervention approaches are therefore usually aimed at remediating (remaining) difficulties with complex syntax. In order to select appropriate targets for grammatical intervention, it is necessary to chart grammatical development (i.e., grammatical complexity and grammatical correctness) of school-age children with SLI in the middle and higher grades.

Specific language impairment: Definition and characteristics

SLI is defined as a developmental language disorder in which children fail to acquire their native language properly/completely, despite having normal non-verbal intelligence, no hearing problems, and no known neurological dysfunctions or behavioral, emotional or social problems (Leonard, 1998). The prevalence of SLI has been estimated to affect approximately 5-7% of a kindergarten population (Tomblin et al., 1997). SLI is basically defined by way of exclusion, meaning that the language problems cannot be explained by a known cause. However, there are also a number of (more clinically pragmatic) inclusion criteria. For instance, children have to show language test scores of -1.25 standard deviations below the age-normed average (or worse), and the obtained non-verbal IQ score must be higher than 85 (Leonard, 1998).

It has been shown that SLI runs in families. This genetic component has been demonstrated in various heritability studies, such as family evaluations (Tallal, Ross, & Curtiss, 1989; Tomblin, 1989; Whitehurst, Arnold, Smith, & Fischel, 1991), twin studies (Bishop, Adams, & Norbury, 2006; Bishop, North, & Donlan, 1995; Tomblin & Buckwalter, 1998), and in genetic etiology studies (Fisher, Vargha-Khadem, Watkins, & Monaco et al., 1998; Lai et al., 2001; Vernes, et al., 2008). Furthermore, SLI affects males more often than females (Tomblin et al., 1997) as is also the case with other developmental disorders, such as dyslexia and autism.

The word 'specific' in SLI is somewhat confusing or misleading. The specificity lies in the fact that only language development is regarded as impaired when compared to other developmental domains. However, recent evidence suggests that children with SLI also experience problems in particular non-linguistic cognitive skills, e.g., executive functions, mental rotation, or motor ability (Leonard, 1998; Ullman & Pierpont, 2006; Windsor & Kohnert, 2009). Some children with SLI also exhibit limitations in other cognitive domains such as working memory (Archibald & Gathercole, 2006; Ellis Weismer & Evans, 2002; Gathercole & Baddeley, 1990; Montgomery, Magimairaj, & Finney, 2010). Furthermore, a relatively high incidence of reading impairment has been identified in children with SLI (Bishop & Snowling, 2004; Tager-Flusberg & Cooper, 1999; van Weerdenburg, Verhoeven, Bosman, & van Balkom, 2011).

Another characteristic of SLI is the large heterogeneity. Language profiles tend to vary greatly among children with SLI. Although difficulty with morphosyntax is generally seen as a hallmark, these children can, but need not, also suffer from impairments in other linguistic domains, such as phonology, semantics, or pragmatics. This heterogeneity has inspired researchers to search for subgroups of children with SLI (Conti-Ramsden, Crutchley, & Botting, 1997; Rapin & Allen, 1983). For example, Rapin and Allen identified six subtypes (i.e., verbal auditory agnosia, phonologic/syntactic deficit disorder, verbal dyspraxia, speech programming deficit disorder, lexical deficit disorder, and semantic pragmatic deficit disorder). Conti-Ramsden et al. (1997) found basically the same six subgroups. Although these subgroups were stable over time, a large part of the children with SLI moved to a different subgroup within one year (Conti-Ramsden & Botting, 1999). SLI should therefore be regarded as a dynamic condition (Van Weerdenburg, Verhoeven, & van Balkom, 2006). Yet another subgroup was identified by Van der Lely (1998; 2004) and labeled grammatical SLI. However, only very few children seemed to fit a subtype that exclusively exhibits grammatical problems. The fact that children with SLI generally receive treatment for their language difficulties, may be regarded as a confounding factor in research for subgroups of SLI. Conceivably, language domains that are most resistant to treatment would then constitute the residue of the language difficulties. In older children with SLI, this remainder could well be their grammatical problems, which could subsequently be interpreted as grammatical SLI (Bishop, Bright, James, Bishop, & van der Lely, 2000).

To date, the attempts to identify subgroups in SLI populations have not been very successful and the clinical value of such subclassifications can therefore be questioned. So far, research on subgroups has either been based on clinical judgments or psychometric tests. However, in the near future, the rapid developments in neuroimaging techniques and genetics will probably advance the research on subgroups.

Explanatory frameworks of SLI

Theories of SLI are generally divided into two types: those that explain SLI as a deficit in (acquiring) linguistic knowledge, also known as representational accounts, and those that explain SLI as a deficit in domain-general or domain-specific cognitive processes (which underlie or subserve language acquisition processes), also known as processing accounts.

The representational accounts have often started from the phonological and morphological characteristics of the verb and its syntactic relations. These linguistic theories were clustered by De Jong (2004) into four groups that will be discussed here briefly.

The first group of explanations centers on the saliency and richness of the inflectional verb paradigm. The *sparse morphology hypothesis* (Leonard, 1987) claimed that children can benefit from the richness of the inflectional system, like for instance in Spanish, where verb inflectional suffixes are syllabic, stressed, and have full vowels. Children learning a language with a sparse morphology do not have this benefit. This would explain why Spanish children with SLI experience fewer problems with verb morphology than English children with SLI. In the same line, but more phono-morphologically orientated is the *surface hypothesis* (Leonard, 1989). Here, the claim is that non-salient morphemes, such as determiners and verb affixes, which are generally unstressed, are more difficult to learn/acquire.

The second group of explanations deals with agreement relationships. Clahsen (1992) proposed the *missing agreement hypothesis*. According to this hypothesis, problems with determiners, adjectival inflection and subject-verb agreement are caused by a lack of knowledge of asymmetrical relations between grammatical categories, where one category controls the other, like gender and number markings on determiners and articles, which are dependent upon the grammatical properties of the noun they accompany. Originally proposed for German SLI, this proposal was later suggested to also hold for other languages, such as English.

The third group of explanations focusses on optionality. Rice and Wexler (1996) proposed the *extended optional infinitive stage hypothesis*. This hypothesis states that all children are supposed to pass through a developmental stage in which they do not know that verbs in main clauses should be marked for tense features, and consequently, non-finite forms can be used where the adult grammar would require finite forms. In children with SLI, this stage is assumed to be prolonged. Later, this hypothesis was revised into the *agreement/tense omission model*. Both accounts presume a developmental delay in children with SLI. Their grammar is not qualitatively different but develops at a slower pace than in typically developing children (Wexler, Schütze & Rice, 1998).

Finally, the fourth group of representational accounts discussed here, concerns deficits of linguistic knowledge in children with SLI. These theories, such as the *missing feature hypothesis* (Gopnik, 1990) and the *implicit rule deficit hypothesis* (Gopnik & Crago, 1991) assume that linguistic knowledge is innate, and that this knowledge is partly absent or incomplete in children with SLI. Because of this lack of competence, children with SLI have to rely on rote learning and memory and according to this hypothesis, correctly produced forms would only appear by chance.

The second type of theories are processing accounts that in general explain SLI as the result of a limitation in information-processing capacity. For instance, children with SLI also have (subtle) deficits or weaknesses in non-linguistic domains, like motor skills, visuo-spatial abilities, memory skills and attention span. Therefore, the specificity of SLI is questioned by these theories (Joanisse & Seidenberg, 1998; Ullman & Pierpont, 2005). It has been argued that the selection of children with SLI based on exclusionary criteria could severely bias the subject sample and thus create the false impression that SLI is a distinct disorder restricted to language. Some of these non-linguistic deficits could explain the language problems of children with SLI. For instance, deficits in phonological working memory could lead to imprecise phonological representations. Phonological memory traces could also decay more rapidly, and a limited storage capacity could lead to fewer items stored. These limitations would hamper lexical learning, as well as sentence processing (Gathercole & Baddeley, 1990). Another proposal concerns deficits in the speed of information processing (Tallal, 1980). Because of an auditory temporal processing deficit language input is degraded which in turn would cause impairments in phonology, lexical learning and grammatical deficits.

More recently, Ullman and Pierpont (2005) proposed the *procedural deficit hypothesis*. According to this hypothesis, the linguistic and non-linguistic deficits of children with SLI are to be explained by an abnormal development of the brain structures that make up the procedural memory system. This brain network, rooted in the frontal lobe and basal ganglia, is dedicated to learning and execution of motor and cognitive skills, including important aspects of grammar and lexical retrieval. The declarative memory system, which depends on other brain structures is supposed to be largely intact. The function of the procedural system would be the learning and use of rule-governed computations (mental grammar), whereas the declarative memory is the storage for idiosyncratic mappings (mental lexicon). Ullman and Pierpont (2005) argued that a portion of the heterogeneity in SLI could be explained by the variation between individuals as to which structures of the procedural memory system are affected and to what degree. Furthermore, the declarative memory system is believed to compensate for deficits in the procedural memory system. Children with SLI would recruit declarative memory to learn particular grammatical forms as unanalyzed wholes (e.g., past tense verb forms) and even explicit grammatical rules (e.g., past tense rule). It was suggested that improvement in language skills over time observed in children with SLI would partly depend on the strength of their lexical/declarative abilities. It was also argued that the variation in symptoms of SLI across languages could be explained by the degree of saliency or learnability of certain types of grammatical rules: the rules that are more salient or simpler should be easier to learn in declarative memory (Ullman & Pierpont, 2005).

Theories explaining SLI are in continuous development. New theories are formulated and older ones are revised due to new insights. Some linguistic theories of SLI have been proposed for specific languages, such as the *missing agreement hypothesis* (Clahsen, 1992) for German and some theories only apply to early developmental stages, like the *extended optional infinitive stage* (Wexler & Rice, 1996). Most linguistic theories only deal with the grammatical symptoms of SLI and do not include the difficulties that children with SLI face in other language domains. A major problem for all linguistic accounts is that they are not able to adequately explain the inconsistency in morphological marking by children with SLI on the one hand and the variation in grammatical symptoms of SLI between languages on the other hand.

The processing accounts of SLI can explain some of the linguistic and non-linguistic deficits seen in children with SLI, such as difficulties in processing brief or rapidly presented (non)verbal stimuli, difficulties in word retrieval and phonological discrimination, and the simultaneous execution of multiple tasks (Leonard, 1998). However, constructing testable hypotheses constitutes a problem for limited processing accounts. Any kind of cognitive functional impairment could potentially be explained by a reduction in processing resources. Although the processing accounts addressing processing speed or insufficient working memory can explain some of the language problems seen in children with SLI, not all children with SLI seem to suffer from problems in these domains (Leonard, 1998; Bishop, Adams, & Rosen, 2006). To date, behavioral evidence supporting the *procedural deficit hypothesis* largely comes from reports of errors in past tense forms observed in children with SLI (Schwartz, 2009).

Research on Dutch children with SLI

In the Netherlands, a series of studies focusing on grammatical development in Dutch children with SLI have been conducted. The first large-scale study was conducted by Bol and Kuiken (1990). In order to construct a Dutch version of LARSP (Crystal, Fletcher, & Garman, 1976) they examined grammatical development in typically developing (TD) children and three different clinical groups. Their research concentrated on the presence or absence of linguistic phenomena in the spontaneous speech of the children and not so much on grammatical errors. The SLI group (n = 18) had a mean age of 5;11 years (age range 4;8-8;2 years). In comparison to the TD children, the SLI group overused structures with two constituents and underused structures with four or more constituents as well as sentences with coordination. They also used fewer phrase structures containing personal and possessive pronouns. In addition, instances of determiner + noun combinations and repletive and partitive '*er*' (there) were also underrepresented. With regard to morphology, fewer diminutives and agreement of first person singular was found. The SLI group did not differ from TD children in their use of modals and copulas.

The inclusion of the Bol and Kuiken (1990) corpus in the CHILDES database certainly paved the way for further research on Dutch children with SLI. Over the years, this corpus was used in a number of studies. For instance, Bol and de Jong (1992) examined the use of modal and aspect auxiliaries and found no group differences between children with SLI and MLU-matched TD controls. However, 6 out of 16 children with SLI omitted aspect auxiliaries *'hebben'* (to have) and *'zijn'* (to be) in sentences with participles. It was suggested that the Dutch auxiliary system was less complex and forms were more salient compared to the English system, which could account for the relatively good performance of Dutch speaking children with SLI, as compared to their English-speaking peers.

The second comprehensive study was performed by de Jong (1999) who examined verb inflectional morphology and verb argument structure in 35 children with SLI, with a mean age of 7;8 years. It was found that the children with SLI produced fewer past tense forms when compared to younger TD children. The SLI group performed poorly on past tense inflection. Past tense inflections and sometimes the entire verb were omitted. However, substitutions also occurred, which was not observed in English speaking children with SLI. Furthermore, the children with SLI frequently used dummy auxiliaries combined with infinitives. De Jong regarded this as a 'strategy' to avoid movement and inflection of the main verb. The use of a past tense dummy auxiliary was suggested to function as an early past tense carrier.

The examination of subject-verb agreement showed that the children with SLI either produced a verb stem form, a singular verb inflection with a plural subject, or an infinitival form in sentence final position. Agreement errors were disproportionally high in the SLI group compared to TD children. However, verb morphology errors were not produced consistently and the children with SLI did not perform at chance level. Therefore, de Jong concluded that the children with SLI did succeed in acquiring the verb inflectional paradigms. However, these representations should be seen as vulnerable. In addition, verb argument structure was examined in the Bol and Kuiken (1990) corpus. Results showed that the children with SLI showed a preference for a simple argument structure and a propensity to select intransitive frames. Problems with verb argument structure were found to coexist with problems involving functional categories.

The Bol and Kuiken (1990) corpus was also examined by Wexler, Schaeffer and Bol (2004), who found that children with SLI hardly ever violated the verb-second rule. Only 0.2% of the verb second positions were filled with an infinitival verb, and the SLI group even outperformed the younger TD group (age range 1;7 to 3;7). In both groups, proportions of root infinitives declined as MLU increased. However, even the older children with SLI (age range 6;0 - 8;2) still produced 15% root infinitives, whereas in TD children (age range 3;1 - 3;7) this rate had dropped to 7%. Wexler et al. (2004) doubted whether children with SLI would ever leave the root infinitive stage.

In a more recent study, Bol and Kasparian (2009) re-examined the use of pronouns in the Bol and Kuiken (1990) corpus. The children with SLI were found to make case errors in personal pronouns, by substituting subject forms with object forms. Furthermore, personal pronouns instead of possessive pronouns were produced.

Another line of research concerned the comparison between Dutch SLI and children learning Dutch as a second language (L2). Researchers noticed that the morphosyntactic problems of Dutch children with SLI closely resembled those of L2 children and adults (Orgassa, 2009; Steenge, 2006; Verhoeven, Steenge, & van Balkom, 2011). Notably, the problems with determiner-noun agreement, verb placement, the extended use of dummy auxiliaries, and the deletion of '*er*' (there) stood out. Evidently, some characteristics of Dutch are more difficult to acquire than others. Therefore, L2 children could be misdiagnosed as having SLI because of these similarities between L2 and SLI symptoms, which has led researchers to design a number of studies aimed at disentangling SLI and bilingualism. For instance, verb morphology production in narratives of monolingual TD and SLI children was compared to bilingual TD and SLI children, aged 7 and 9 years (Steenge, 2006; Verhoeven, Steenge, & van Balkom, 2011). Omission of an agreement marker in 3rd person singular verb forms was regarded as a clinical marker of SLI in L1 and L2 learners. Furthermore, the children with SLI told longer stories than the TD controls, but MLU was lower and proportions of ungrammatical sentences were higher. In a similar group design, but with an elicitation task, verb placement and verb inflection was studied (Orgassa, 2009). The 7-year-old children with SLI had almost no verb placement errors in main clauses. However, these errors were found in main clauses with inversion and were especially prominent in relative clauses. The SLI group also used more dummy auxiliaries than TD L1 and L2 learners. Contrastively, no differences in the use of dummy auxiliaries was found between monolingual and bilingual SLI groups. Clearly, the overuse of dummy verbs was related to having SLI.

Productive and receptive knowledge (i.e., grammatical judgment) of definite determiners in Dutch children with SLI was investigated by Keij, Cornips, van Hout, Hulk, and van Emmerik (2012). Three groups were compared, namely monolingual TD children (age range 6;7–9;11), monolingual children with SLI (age range 8;4–12;0), and TD L2 children (age range 6;7–10;0). The three groups demonstrated different stages in mastery of the Dutch gender paradigm. The children with SLI performed poorly on the production and the grammatical judgment tasks. Performance on common gender determiner *'de'* was poor and on neuter determiner *'het'* at chance level, although considerable variation was found. Keij et al. (2012) concluded that the children with SLI were only at the first stage of the discovery of the gender paradigm, showing awareness of gender distinction, but not able to produce neuter determiner-noun combinations correctly.

A recent study compared performance differences of 7-year-old children with SLI (range 6;1-9;9 years) and age-matched TD controls on a story-retelling task versus a story-generation task (Duinmeijer, de Jong, & Scheper, 2012). On both tasks, the SLI group was outperformed by the TD group on plot score, MLU and grammaticality. On the story-retelling task, the SLI group also used fewer embedded clauses than TD controls. When both tasks were compared within the SLI group, it was found that scores on measures of complexity, embedding, and disfluency were higher on the retelling task. It was concluded that the children with SLI appeared to benefit from the language input in the retelling condition, but having to recall and reproduce the story led to more disfluencies.

To conclude, this overview of the research demonstrates that, although Dutch SLI has had its share of research interest, longitudinal studies have so far been scarce. Furthermore, a large part of the studies made use of the same Bol and Kuiken (1990) corpus of spontaneous speech. The size of this corpus is somewhat limited and older school-age children with SLI can be regarded as underrepresented. Moreover, some studies examined relatively small groups with considerable age ranges in the SLI groups. In some studies experimental tasks were used, but a large portion of the research findings was based on the analysis of spontaneous speech. So far, only a few studies used narratives to investigate the language skills of Dutch children with SLI. Narratives have been argued to be more revealing than spontaneous speech in examining older children with SLI (Blankenstijn & Scheper, 2003; Norbury & Bishop, 2003; Reilly, Losh, Bellugi, & Wulfeck, 2004; Wetherell, Botting, & Conti-Ramsden, 2007).

The present thesis

In this thesis, grammatical development of school-age children with SLI is related to a number of issues: the delay versus deviance debate in SLI, the role of the lexicon in grammatical development, speech disfluency as a marker of expressive language impairment, and the effectiveness of a metalinguistic intervention approach for older school-age children with SLI. To begin with, a longitudinal examination of grammatical development was conducted in two groups of children with SLI while comparing them to language-matched (LA) and age-matched (CA) TD control groups. The children with SLI participating in this study were all diagnosed as having severe language impairments by a team of specialists at all three time points in two years with a 12-month interval. At the three time points, the children in the first SLI group were aged 6, 7, and 8 years and the children in the second SLI group were aged 8, 9, and 10 years, respectively. The age range in the groups never exceeded 3.5 months. Each SLI group and each LA and CA control group consisted of 30 children. The CA group in the first study also functioned as the LA group in the second study, so in total there were five groups of children (150 participants). Grammatical development was examined with a narrative task that was administered three times in two years, amounting to 450 narratives. In the third study, speech disfluency was investigated in the older SLI group at the ages of 8, 9, and 10 years. To date, only Duinmeijer et al. (2012) studied speech disfluency in Dutch children with SLI. It has been suggested that speech disfluencies can be a marker of expressive language impairment, especially in older children with SLI (Finneran, Leonard, & Miller, 2009). In the fourth study, a metalinguistic intervention approach was examined in twelve children with SLI enrolled in a school for children with severe speech and language disorders. To this day, intervention studies aimed at the remediation of grammatical targets are virtually non-existent in the Netherlands. The only study that could be found is Braam-Voeten (1997) who compared two therapy approaches in seven children with SLI. The first therapy approach targeted the expansion of sentence structures, the second one was aimed at improving children's marking for agreement. Results indicated that the children omitted fewer words from sentences, but no improvement was found for subject-verb agreement. Because this was a small-scale study results have to be interpreted with caution.

The outline of the present thesis is as follows. In Chapter 2, the delay versus deviance dichotomy in children with SLI is investigated. To this aim, grammatical development in monolingual Dutch-speaking children with SLI and TD control groups is studied using a narrative task. All children were tested at three points in time with an interval of 12 months during a period of two years. The SLI and CA groups were measured at ages 6, 7, and 8 years and the LA children were two years younger. The three groups were compared on a wide range of measures of grammatical complexity and grammatical correctness. Developmental patterns in the SLI group were compared to those in TD children.

In Chapter 3, the role of vocabulary in grammatical development of older school-age children with SLI is investigated. In normal language development, lexicon and grammar are strongly related. The question is whether this relation also exists in children with SLI, and whether vocabulary can be related to grammatical complexity as well as to grammatical correctness. Again, all children were tested at three points in time with an interval of 12 months during a period of two years, but now, the SLI and CA groups were measured at ages 8, 9, and 10 years and the LA children were two years younger. Scores on vocabulary measures were correlated with grammatical complexity and correctness.

In Chapter 4, the development of speech disfluency is examined in the same sample of children that was examined in Chapter 3. Speech disfluencies are examined for type, frequency, syntactic distribution and the duration of silent pauses. The location, types and frequencies of speech disfluencies in children with SLI may inform research on the nature of (subtle) language difficulties that cannot always be deduced from standardized tests.

In Chapter 5, the effectiveness of a metalinguistic intervention program to remediate an aspect of grammar in older school-age children with SLI is evaluated. The intervention was targeted at the production of relative clauses. A specific therapy program and tests were especially designed for this study, based on ideas and programs developed in the 1970s. The intervention program 'Metataal' is a multimodal approach that deploys visual, auditory and motor skills, and literally teaches children to build sentences with Lego blocks.

In Chapter 6, the general discussion and conclusions of the findings from the four different studies are presented. Furthermore, some theoretical and clinical implications are discussed as well as suggestions for future research and directions.

References

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in SLI. *International Journal of Language and Communication Disorders, 41,* 675–742.

Bishop, D. V. M. (2010). Which neurodevelopmental disorders get researched and why? *PloS One*, *5*(11), e15112. doi:10.1371/journal.pone.0015112.

Bishop, D. V. M., Adams, C. V., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: Evidence from 6-year-old twins. *Genes, Brain and Behavior, 5,* 158–169.

Bishop, D. V. M., Adams, C. V., & Rosen, S. (2006). Resistance of grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language and Communication Disorders*, *41,* 19-40.

Bishop, D. V. M., Bright, P., James, C., Bishop, S. J., & Van der Lely, H. K. J. (2000). Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics*, *21*, 159–181.

Bishop, D. V. M., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental Medicine and Child Neurology, 37,* 56–71.

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 30,* 858–886.

Blankenstijn, C., & Scheper, A. (2003). *Language development in children with psychiatric impairment*. Utrecht: LOT Dissertation Series 82. (Doctoral dissertation, University of Amsterdam). Retrieved from http://www.lotpublications.nl/publish/articles/000542/bookpart.pdf.

Bol, G. W., & de Jong, J. (1992). Auxiliary verbs in Dutch SLI children. *Scandinavian Journal of Logopedics and Phoniatrics*, *17*, 17–21.

Bol, G. W., & Kasparian, K. (2009). The production of pronouns in Dutch children with developmental language disorders: A comparison between children with SLI, hearing impairment, and Down's syndrome. *Clinical Linguistics and Phonetics*, *23*, 631–646.

Bol, G. W., & Kuiken, F. (1990). Grammatical analysis of developmental language disorders: A study of the morphosyntax of children with specific language disorders, with hearing impairment and with Down's syndrome. *Clinical Linguistics and Phonetics*, *4*, 77-86.

Braam-Voeten, M. W. J. A. (1997). *Van Taaltheorie naar taaltherapie. Evaluatie en analyse van een congruentie- en capaciteitentherapie bij kinderen met taalstoornissen.* (Unpublished doctoral dissertation). Twente University, Enschede.

Clahsen, H. (1992). *Linguistic perspectives on specific language impairment.* Düsseldorf: University of Düsseldorf.

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42,* 1195–1204.

Conti-Ramsden, G., Crutchley, G., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research, 40,* 765-777.

Crystal, D., Fletcher, P., & Garman, M. (1976). *The grammatical analysis of language disability: A procedure for assessment and remediation*. London: Edward Arnold.

De Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure.* (Doctoral dissertation). Groningen: Grodil. Retrieved from: http://irs.ub.rug.nl/ppn/297976826.

De Jong, J. (2004). Grammatical impairment: An overview and a sketch of Dutch. In L. Verhoeven & H. Van Balkom, (Eds.), *Classification of Developmental Language Disorders*, 261-281, Mahwah, NJ: Erlbaum.

Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language and Communication Disorders*, *47*, 542–55.

Ellis Weismer, S., & Evans, J. L. (2002). The role of processing limitations in early identification of specific language impairment. *Topics in Language Disorders*, *22*, 15–29.

Finneran, D. A., Leonard, L. B., & Miller, C. A. (2009). Speech disruptions in the sentence formulation of school-age children with specific language impairment. *International Journal of Language and Communication Disorders, 44,* 271-286.

Fisher, S. E., Vargha-Kadhem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localization of a gene implicated in a severe speech and language disorder. *Nature Genetics, 18,* 168–170.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language 29,* 336–360.

Gopnik, M. (1990). Feature-blindness. *Language Acquisition, 1,* 139–164.

Gopnik, M., & Crago, M. B. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39,* 1–50.

Joanisse, M. F. & Seidenberg, M. S. (1998). Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Neuroscience, 2,* 240-247.

Keij, B., Cornips, L., van Hout, R., Hulk, A., & van Emmerik, J. (2012). Knowing versus producing: The acquisition of grammatical gender and the definite determiner in Dutch by L1-TD, L1-SLI, and eL2 children. *Linguistic Approaches to Bilingualism*, *2*, 379–403.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., Levy, E. R., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature, 413,* 519–523.

Leonard, L. B. (1987). Is specific language impairment a useful construct? In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Disorders of first language development* (pp. 1–39). Cambridge, NY: Cambridge University Press.

Leonard, L.B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10,* 179-202.

Leonard, L. B. (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*, *19*, 78–94.

Norbury, C. F., & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders*, *38*, 287–313.

Orgassa, A. (2009). *Specific language impairment in a bilingual context. The acquisition of Dutch by Turkish-Dutch learners* (Doctoral dissertation University of Amsterdam). Utrecht: LOT Dissertation Series 220. Retrieved from: http://www.lotpublications.nl/ publish/articles/003604/bookpart.pdf

Rapin, I., & Allen, D. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading, and spelling* (pp. 155–184). New York, NY: Academic Press.

Rice, L., K. Wexler, P. L. Cleave (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research*, *38,* 850-863.

Reilly, J., Losh, M., Bellugi, U. & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language, 88*, 229-247.

Schwartz, R.G. (2009). Specific language impairment. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 3-43). New York, NY: Psychological Press.

Steenge, J. (2006). *Bilingual children with specific language impairment: Additionally disadvantaged?* (Doctoral dissertation). Nijmegen: EAC, Research Centre on Atypical Communication, Radboud University Nijmegen. Retrieved from: http://repository.ubn.kun.nl/handle/2066/29807.

Tager-Flusberg, H. & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1275–1278.

Tallal, P. (1980). Auditory temporal perception, phonics and reading disabilities in children. *Brain and Language, 9,* 182-198.

Tallal, P., Ross, R., & Curtiss, S. (1989). Familial aggregation in specific language impairment. *Journal of Speech and Hearing Disorders*, *54,*167–173.

Tomblin, J. B. (1989). Familial concentration of developmental language impairment. *Journal of Speech and Hearing Disorders*, *54,* 287–295.

Tomblin, J. B., & Buckwalter, P. (1998). The heritability of poor language achievement among twins. *Journal of Speech and Hearing Research, 41,* 188–199.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & Brien, M. O. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research*, *40*, 1245–1260.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*, 399–433.

Van der Lely, H. K. J. (1998). SLI in children: Movement, economy and deficits in the computational-syntactic system. *Language Acquisition, 7*, 161–192.

Van der Lely, H. K. J. (2004). Evidence for and implications of a domain-specific grammatical deficit. In L. Jenkins (Ed.), *The genetics of language* (pp. 117–145). Oxford: Elsevier.

Van Weerdenburg, M., Verhoeven, L., Bosman, A., & van Balkom, H. (2011). Predicting word decoding and word spelling development in children with specific language impairment. *Journal of Communication Disorders*, *44*, 392-411.

Van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry, 47*, 176-189.

Verhoeven, L., Steenge, J., & Van Balkom, H. (2011). Verb morphology as clinical marker of specific language impairment: Evidence from first and second language learners. *Research in Developmental Disabilities*, *32*, 1186–1193.

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M. Alarcón, M., Oliver, P. L., Davies, K. E., Geschwind, D. H., Monaco, A. P., & Fisher, S.E. (2008). A functional genetic link between distinct developmental language disorders. *The New England Journal of Medicine, 359,* 2337–2345.

Wetherell, D., Botting, N., & Conti-Ramsden, G. (2007). Narrative in adolescent specific language impairment (SLI): A comparison with peers across two different narrative genres. *International Journal of Language and Communication Disorders, 42*, 583-605.

Wexler, K., Schaeffer, J., & Bol, G. W. (2004). Verbal syntax and morphology in typically developing Dutch children and children with SLI: How developmental data can play an important role in morphological. *Syntax*, *7*, 148–198.

Wexler, K., Schütze, C. T. & Rice, M. (1998). Subject case in children with SLI and unaffected normal controls: Evidence for the Agr/Tns omission model. *Language Acquisition, 7,* 317–344.

Whitehurst, G. J., Arnold, D. S., Smith, M., & Fischel, J. E. (1991). Family history in developmental expressive language delay. *Journal of Speech and Hearing Research, 34*, 1150–1157.

Windsor, J. & Kohnert, K. (2009). Processing speed, attention, and perception: Implications for child language disorders. In R.G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 445-461). New York, NY: Psychology Press.

2 Grammatical development in Dutch school-age children with SLI: Delay versus deviance [1]

Abstract

Purpose: The purpose of this study was to identify delay or deviance patterns in the grammatical development of Dutch school-age children with specific language impairment (SLI).

Method: Grammatical complexity, grammatical correctness (verb-related and non-verb related errors), and the use of dummy auxiliaries were assessed using a narrative task that was administered at three points in time (T1, T2, and T3) with an interval of 12 months during a two-year period. Participants were 30 monolingual Dutch children with SLI, 30 age-matched (CA), and 30 language-matched (LA) children. The SLI and CA group were aged 6 years at T1 and the LA group was 2 years younger.

Results: On complexity, the SLI group was only outperformed by the CA group on relative clause use at T3. On correctness measures, the SLI group performed more poorly than both LA and CA groups and their development would best fit a delay-within-delay pattern. The use of dummy auxiliaries in the SLI group did not change over time, and this pattern did not fit that of the control groups.

Conclusions: The narrative analysis demonstrates different developmental trajectories for the grammatical complexity and correctness measures in the SLI and control groups. In the SLI group, grammatical skills continue to develop. The pattern of dummy auxiliary use may be regarded as deviant compared to typically developing children.

Keywords: narrative, specific language impairment (SLI), grammatical complexity, grammatical correctness.

---

[1] Submitted (Authors: Rob Zwitserlood[a,c], Marjolijn van Weerdenburg[b], Ludo Verhoeven[b], and Frank Wijnen[a]), [a]Utrecht Institute of Linguistics (OTS), University Utrecht, The Netherlands, [b]Behavioural Science Institute, Radboud University Nijmegen, The Netherlands, [c]Koninklijke Auris Groep, The Netherlands).

An on-going debate is whether language development in children with SLI should be regarded as delayed or deviant relative to normal development. Language development in children with SLI can be described as delayed if performance falls below that of typically developing children matched on chronological age (CA), but resembles performance of a younger control group matched on language age (LA). Language development can be termed as deviant when performance does not resemble the development in LA matched control children.

The delay versus deviance issue in children with SLI is clinically relevant, because information on relative strengths and weaknesses in their language development can inform diagnostics (Rice, Warren, & Betz, 2005) and intervention (Paradis, 2010). The question whether language development of children with SLI should be characterized as delayed or deviant also has theoretical implications. Deviant language profiles provide evidence of unique characteristics in impaired language development, whereas delay-only profiles do not. In a deviant profile, children may have specific and more severe delays in sub-domains that are inconsistent with a general delay pattern. Children may also make errors that are not observed in typical language development.

According to Leonard (1996), a strict dichotomy in terms of delay and deviance has been both misleading and helpful in our understanding of language disorders in children. Delay should not be taken to mean that children with SLI simply lag behind their peers. Deviance, on the other hand, should not imply that the language development of children with SLI "borders on the bizarre" (Leonard, 1996, p. 217). Most often these children do produce the same grammatical errors as seen in younger typically developing (TD) children. Leonard (1998) identified five possible patterns of language development in children with SLI. First, children with SLI can exhibit a straightforward language *delay*, consisting of a late language emergence and a slower language development. Eventually, these children with SLI may catch up. If they do, it is questionable whether these children should have been considered as having SLI in the first place. A second pattern would be that their language development reaches a *plateau* before mastery levels are achieved. A third pattern can be seen when language development in children with SLI shows an *uneven profile*. For example, in certain elements of morphosyntax the children with SLI resemble TD children that are one year younger, but in other elements they can lag behind three years or more. A fourth pattern is found when children with SLI exhibit an *abnormal frequency of error*. A particular type of grammatical error, which is also seen in younger TD children, persists in SLI and occurs in high frequencies that are never seen in TD children. Finally, a fifth pattern is observed when children with SLI show *qualitative differences*. According to Leonard (1998), such highly unusual and unique patterns have been mentioned in the literature, but especially with regard to phonological development. However, the evidence for truly qualitatively different patterns remains anecdotal. If delay is interpreted as just lagging behind in development, only the first pattern would be considered a delay, and the other four would classify as deviant. On the other hand, if deviance is defined as truly qualitatively different, then only the fifth pattern would fit this definition.

Thomas et al. (2009) took a different approach than Leonard (1998) and based their taxonomy of developmental disorders on the analysis of growth trajectories. They posed that typical development could be regarded as following a linear trajectory. Simply speaking, children progress in their cognitive skills as they grow older. Growth curves depicting developmental delay could then take three forms: a delayed onset; a slowed developmental rate; or a combination of the two. Deviance in development could take the form of a non-linear trajectory; a premature asymptote; a zero trajectory (no growth); or a pattern that shows no systematic relation with age. Naturally, development and growth are the key issue, and therefore carefully planned longitudinal studies are preferable to cross-sectional designs to investigate delay versus deviance patterns.

Longitudinal Studies of Delay-Deviance in SLI

In order to identify deviance and delay patterns in children with SLI, a three-group design with typically developing LA and CA control groups is the most appropriate choice. However, to disclose different or even unique developmental patterns in children with SLI compared to control groups, a longitudinal approach examining a wide range of morphosyntactic measures is necessary (Paradis, 2010). To date, only a few studies have taken such an extensive longitudinal approach.

A relatively early longitudinal study investigating the delay versus deviance issue in children with SLI was Curtiss, Katz, and Tallal (1992). Children with SLI aged 4;4 years (range 4;0-4;9) were compared to LA children aged 2;9 years (range 2;1-3;8), but no CA control group participated. An omnibus language test was administered over a 5-year period assessing both comprehension (picture pointing tasks) and production (sentence completion tasks) of grammatical features. Results indicated that the developmental trajectories of children with SLI were similar to those of LA children. The conclusion from this study was that grammatical development in SLI should be interpreted as delayed. However, Curtiss et al. (1992) defined deviance in terms of abnormality in reaching successive stages of grammatical development, implying that children would show acquisition of later-learned "difficult" structures before the mastering of earlier-learned "easier" ones, which was not the case. Furthermore, there are a few points that affect the validity of this study. First, a data loss of 20% was reported during test rounds, partly in the higher level subtests of the children with SLI. Second, the large difference in age range between the LA group (19 months) and the SLI group (9 months) may have influenced the results. A third point is that a sentence completion task might not be the most suitable way to study expressive language, because of the highly structured nature of such a task. Moreover, the exclusive use of a standardized language test, instead of a combination with the analysis of conversational or narrative samples may have limited the opportunity to find deviance patterns.

Rice, Wexler, and Hershberger (1998) used a longitudinal three-group design to study the development of verb and noun morphology in 5-year-old children with SLI over a three-year period). The SLI group was compared with LA and CA control groups. The children with SLI performed poorly compared to LA controls on the correct production of tense and agreement, but had age appropriate scores on noun morphology (regular *-s* plurals). With respect to verb morphology, the children with SLI showed no signs of catching up over time. Rice et al. (1998) argued that the problems in the representation of grammatical tense in the SLI group did not correspond to a simple delay account. Tense and agreement lagged further behind than other elements of morphosyntactic development. Rice et al. (1998) termed this an extended delay.

Another selective delay pattern was described by Rice (2003). Vocabulary development of children with SLI was similar to that of LA children and could be regarded as delayed. Contrastively, the SLI group performed much poorer than the LA control group on tense and agreement tasks. This unexpected delay in grammatical morphology was regarded as deviant compared to the development of other language skills. Rice (2003) described this unevenness in language development as a delay-within-delay. This delay-within-delay pattern best fits the uneven profile from Leonard's (1998) taxonomy.

The studies by Rice and colleagues (1998; 2003) focused mainly on the acquisition of verb morphology and compared verb morphology to development in vocabulary and noun morphology. It is unknown as yet if the conclusions drawn by Rice and colleagues would extend to other domains, such as grammatical complexity, or functional categories such as determiners, pronouns and prepositions. So far, these elements have not been studied extensively in a longitudinal design.

The Present Study

This study aims to contribute further to the delay versus deviance discussion in grammatical development in children with SLI in several ways. To begin with, the focus of the present study is on Dutch, whereas to this date this issue has only been studied in English speaking children with SLI. Dutch differs considerably from English, as will be elaborated on below, and therefore we do not know whether Dutch SLI is identical in phenotype to English SLI. Second, this study opens up a wider perspective on grammatical development of children with SLI than Curtiss et al. (1992) and Rice and colleagues (1998; 2003) by comparing the development of grammatical complexity (complex syntax) to grammatical correctness (verb-related and non-verb related grammatical errors). Different developmental patterns of delay or deviance might show up in either or both dimensions. Third, previous studies used standardized tests (Curtiss et al., 1992), or spontaneous speech samples and elicited probes (Rice et al., 1998). In the present study a narrative task was used, because narratives provide valuable information on the development of vocabulary, morphosyntax and discourse, particularly in older children with SLI (Reilly, Losh, Bellugi, & Wulfeck, 2004). In contrast with spontaneous speech, assessment using a narrative task may lead to more and different errors. The children are prompted to express semantic relations that may be difficult for them, and which they may avoid in spontaneous speech. Therefore, a narrative task may be better at revealing the relative strengths and weaknesses in the language performance of children with SLI (Wetherell, Botting, & Conti-Ramsden, 2007). According to Blankenstijn and Scheper (2003) narratives are a complex genre that makes a great demand on the morphological/syntactic skills of children. In their study, children with psychiatric impairment showed clear morphological and syntactic difficulties in a narrative task and more difficulties than in conversation. Using narratives to study grammatical development can therefore be considered as an ecologically valid way to study grammatical development in children with SLI.

The main research question for the present study was which delay or deviance patterns can be found in the development of grammatical complexity and grammatical correctness in Dutch school-age children with SLI. We also wanted to investigate to what extent the children show progress in their grammatical skills. Do we see continuous development or does the development of grammatical skills level off at a certain age?

SLI in Dutch

To this day, most studies on SLI concern English speaking children. However, English can be regarded as a language with a sparse morphology, whereas Dutch has some enriched morphosyntactic features that merit studying in SLI. First, Dutch is an SOV + verb-second language, which entails that the inflected verb takes second position in main clauses and final position in subordinate clauses. In main clauses, only one constituent can precede the inflected verb. Infinitival verbs and participles always appear in clause final position. In the course of acquiring the verb inflection and verb placement rules, Dutch children go through a developmental stage in which they use dummy auxiliaries as placeholders for inflected lexical verbs in second position (Van Kampen, 1997; Wijnen & Verrips, 1998). Once the rules for verb inflection are mastered, this periphrastic use of dummy auxiliaries fades out. However, Dutch children with SLI tend to overuse these dummy auxiliaries. The proposed explanations state that Dutch children with SLI use these dummy verbs as a 'strategy' to avoid movement and inflection of the lexical verb (de Jong, Blom, & Orgassa, 2013; Orgassa, 2009). De Jong (1999) also suggested that dummy verbs appearing in past tense form might function as an early tense carrier. Second, the Dutch verb inflectional paradigm is somewhat richer than the English system. English children with SLI predominantly omit inflectional morphemes of lexical verbs, whereas Dutch children with SLI not only omit these morphemes, but also often make substitution errors (de Jong, 1999; 2003). Third, Dutch has a more elaborate determiner system with two definite determiners: common gender *'de'* and neuter gender *'het'*. For singular nouns there has to be gender agreement between determiner and noun. English speaking children with SLI are known to omit determiners, but analogous to the errors in verb morphology, Dutch children with SLI also produce substitution errors (Keij, Cornips, van Hout, Hulk, & van Emmerik, 2012). A fourth characteristic is related to this gender agreement between determiner and noun. In relative clauses, the relative pronoun ('*die'* or *'dat'*) has to agree with the gender of the relativized noun in the main clause. Finally, a feature of Dutch known to be very problematic for children with SLI (Bol & Kuiken, 1990), is the correct use of *'er'* (there). This adverb can have many different applications in Dutch sentences and is often omitted in obligatory contexts by Dutch children with SLI.

Based on this short overview of the morphosyntactic characteristics of Dutch and language specific difficulties of Dutch children with SLI, it can be expected that Dutch children with SLI demonstrate difficulties that are at least partly language specific. An investigation of grammatical complexity and correctness can reveal which types of grammatical errors and which difficulties in complex syntax are most prominent for Dutch speaking children with SLI.

Method

Participants

Participants were 30 monolingual Dutch children with SLI, aged 6 years at first measurement, 30 LA control children (matched on language test scores), aged 4 years at first measurement, and 30 CA control children, aged 6 years at first measurement. Informed consent was obtained from all parents. The data from the SLI group originated from a previous study by Van Weerdenburg, Verhoeven, and van Balkom (2006). The children with SLI (23 boys, 7 girls) were all enrolled in special education for children with severe speech and language impairments. The children in the LA group (18 boys, 12 girls) and the CA group (16 boys, 14 girls) were recruited from four different primary schools in the central part of the Netherlands. The 4-year-old LA children can be regarded as cognitively mature enough to participate in various language tasks (Southwood, 2007). Furthermore, around this age, their morphosyntactic development has nearly reached completion. Numerous studies on SLI have used LA control groups that were two years younger. This two year age difference was also found when children were matched on MLU (Paradis, 2010). As can be seen in Table 1, all children obtained non-verbal IQ scores within the normal range, measured with the Raven Colored Progressive Matrices (Raven CPM; van Bon, 1986). However, on the sequential memory tasks from the Dutch version of the Kaufman Assessment Battery for Children (Kaufman ABC; Kaufman & Kaufman, 1983) the children with SLI scored -1.1 standard deviation below the mean. Parisse and Maillart (2009) claimed that typical children with SLI have a profile where, apart from poor scores on language tests, a clear gap exists between scores on sequential memory tasks and other subtests of a non-verbal intelligence test. The children with SLI in the present study clearly fit this description and can thus be regarded as typical SLI.

During first test round (T1), the language abilities of all children were assessed with the Dutch standardized Language Proficiency Test (LPT) for all Children (TAK, *Taaltoets Alle Kinderen*; Verhoeven & Vermeer, 2001). The CA group performed significantly higher on all LPT subtests compared to both the SLI and the LA groups. There were no significant differences between the SLI group and the LA group on any of the LPT subtests. However, the SLI group did perform poorly on the Kaufman ABC sequential memory tasks compared to both control groups ($p < .001$).

Table 1. *Age in months, Raven Coloured Progressive Matrices Non-verbal IQ, Kaufman ABC Sequential Memory, and LPT Language Tests in the three Groups at First Test Round*

	Max. Score[d]	SLI ($n = 30$) *M (SD)*	LA ($n = 30$) *M (SD)*	CA ($n = 30$) *M (SD)*
Age in months		77 (1.5)	55 (2.6)	78 (2.4)
Raven CPM[a]		5.92 (1.92)	5.89 (1.50)	6.32 (1.79)
Kaufman Seq. memory[b]		83.23 (9.44)	98.43 (9.84)	96.03 (11.24)
LPT Receptive vocabulary[c]	96	56.93 (13.40)	53.63 (11.70)	77.33 (6.85)
LPT Sentence comprehension I[c]	42	32.67 (4.48)	31.07 (5.73)	36.90 (2.17)
LPT Sentence comprehension II[c]	42	30.80 (4.15)	28.27 (3.71)	35.40 (2.51)
LPT Morphology[c]	24	11.03 (4.49)	12.77 (3.00)	16.57 (5.56)

Note. [a]standard score (-1 *SD* to +1 *SD* ranges from 3.0 - 7.0), [b]quotient score, [c]raw score, [d]maximum raw score. LPT Sentence comprehension I measures function words, LPT Sentence comprehension II measures syntactic patterns, and LPT Morphology measures production of noun plurals and past participles.

Materials

The Storytelling tasks of the LPT (Figure 1 and 2) were used to elicit narratives. Two monochrome picture series show a sequence of events that form a coherent story each. The children are instructed to look at the pictures carefully and then tell the story in such a way that someone who cannot see the pictures will be able to understand the story in full. The investigator does not ask questions, but may encourage the children to continue if they stop midway. In contrast with other narrative tasks such as the Renfrew Bus Story (Renfrew, 1969) and the Frog Story (Mayer, 1969), both male, female and plural referents appear in the LPT story generation tasks. This variety in characters enlarges the chance to observe a wider range of morphosyntactic errors in pronouns (e.g., case, gender, and number), determiners (e.g., gender) and subject-verb agreement. In the first story all characters are introduced in the first picture. The second story might be more taxing for the children, because some characters are introduced later on or are reintroduced. In this story a shop attendant and a clown are acting in the background and mentioning them is not necessary for a complete and comprehensible narration. The narratives from the SLI group were recorded on audio cassette and later digitized. The stories told by the control groups were digitally recorded on a laptop.

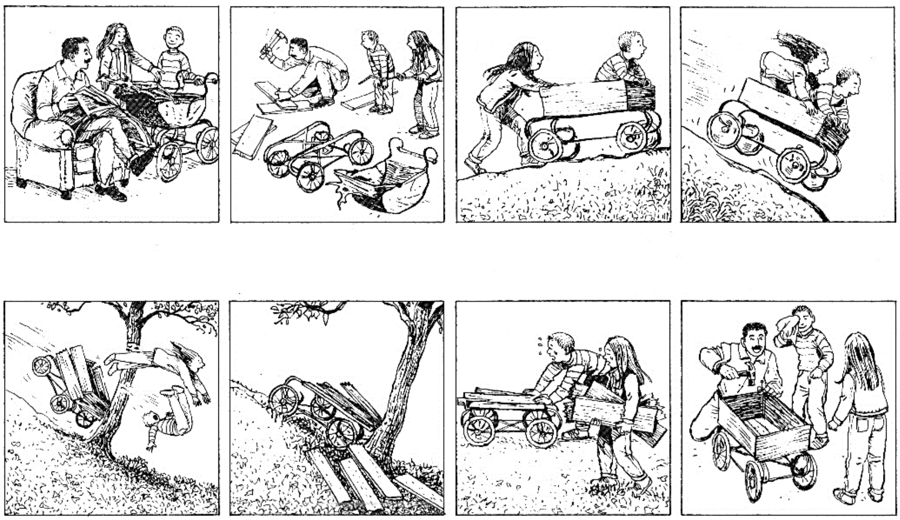

Figure 1. LPT Storytelling task 1. Copyright 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

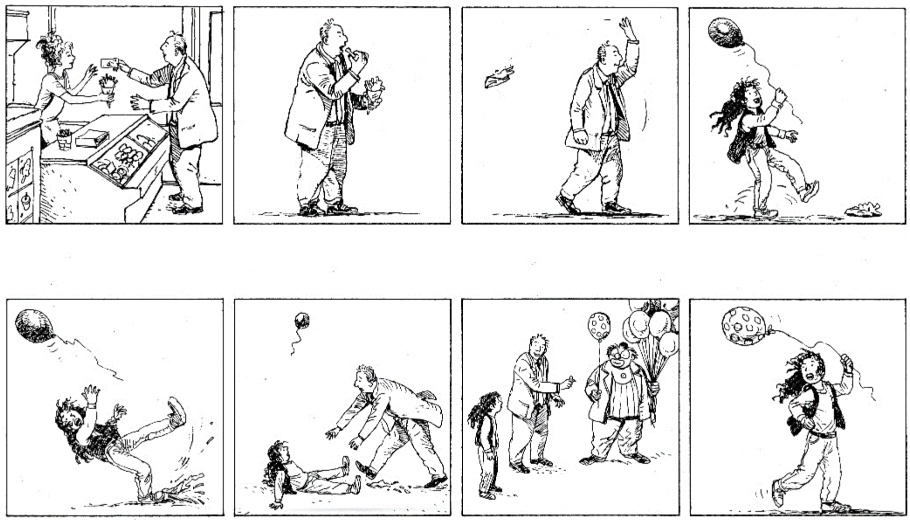

Figure 2. LPT Storytelling task 2. Copyright 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

Procedure

All children were tested at three points in time with an interval of 12 months during a period of two years. The narratives were orthographically transcribed in accordance with CHAT conventions, the coding system of the Child Language Data Exchange System (CHILDES, MacWhinney, 2000). Each transcript contained Storytelling tasks 1 and 2 told in succession in this fixed order. The basic unit of analysis was the T-unit, defined as a single main clause plus any subordinate clause or non-clausal structure that is attached to it or embedded in it (Hunt, 1970). Coordinate clauses were transcribed and counted as separate utterances, unless there was conjunction reduction. Sentences with quoted speech, where the quote forms a full clause containing a subject and a verb, were also transcribed as separate utterances. In addition to the transcriptions on the main tier, additional dependent tiers were created in the CHAT files to code grammatical complexity and grammatical errors.

The reliability of the transcriptions was checked by re-transcribing 10% of the files by either trained speech-language pathologists (for the transcripts of the children with SLI) or trained master students in linguistics (for the transcripts of the control groups). The point-to-point reliability at word level reached 90%.

There were no missing data in the three groups. All children told all stories at all three time points. Furthermore, there were also almost no unanalyzable utterances. When utterances were intelligible or otherwise unanalyzable (false starts or utterances that were broken off) these were excluded from the analysis.

Analysis

In this study, both grammatical complexity and grammatical correctness were investigated. The operationalizations of the dependent measures are presented in Table 2. Because the narratives varied in length, percentages had to be computed for most variables, in order to make the dependent measures comparable across children and time points.

The first step in the analysis concerned grammatical complexity, and four different measures were computed. Mean length of T-units (*MLTU*) in words was used as a general measure of grammatical complexity. Grammatical complexity was further operationalized by the composite measure *percentage of complex sentences*. All sentences with subordinate clauses, coordinated sentences with conjunction reduction, direct speech, and infinitival clauses were counted as complex sentences. Subordinate clauses included all forms of adverbial, nominal, and relative clauses. The complex sentences were subdivided further by computing the measures *percentage of subordinate clauses* and *percentage of relative clauses*. This subdivision was motivated by the fact that subordinate clauses, and relative clauses in particular, are known to pose difficulties for school-age children with SLI (Marinellie, 2004; Novogrodsky & Friedmann, 2006; Schuele & Nicholls, 2000; Schuele & Tolbert, 2001).

The second step in the analysis was the examination of the grammatical correctness in the narratives. The measures related to grammatical correctness can be found in Table 2. As a general measure of grammatical correctness the *percentage of T-units correct* (i.e., error-free T-units) was used. However, a T-unit can contain several grammatical errors, therefore all grammatical errors in the narratives were tallied. According to the SLI literature, children with SLI especially perform poorly on verb morphology and verb argument structure. However, the narratives also provided extensive information on other grammatical errors such as errors in noun morphology, pronouns, missing words and word order. In order to distinguish the different types of errors and to examine possible changes in error patterns over time, the grammatical errors were further specified into a number of different error types which were subsequently arranged in two broad categories: *percentage of verb-related errors* and *percentage of non-verb related errors*. The composite measure *percentage of verb-related errors* contains all observed errors in verb morphology (e.g., subject-verb agreement, tense, auxiliaries, and participles), verb placement and verb argument structure. The means and standard deviations of these different verb-related error measures are presented in Appendix 1. The second composite measure *percentage of non-verb related errors* contains all remaining grammatical errors. This composite measure was not further subdivided, because many different error types were counted and percentages of individual error types could therefore be too low for a meaningful quantitative analysis.

Finally, the measure *percentages of dummy auxiliaries* was included in the analysis. Overuse of these dummy verbs has been reported frequently in studies on Dutch SLI (de Jong, 1999; de Jong et al., 2013; Orgassa, 2009). Using a dummy auxiliary does not render a sentence ungrammatical. Therefore counts of dummy auxiliaries were not included in the grammatical correctness measures. However, a prolonged and frequent use of dummy verbs can be argued to reflect an immature stage of verb morphology mastery, which makes dummy auxiliaries a developmental feature worth investigating.

As we were looking for delay or deviance patterns, it was decided to regard developmental patterns that were not found in either the LA or the CA control groups as deviant. Developmental patterns that would resemble those of the typically developing LA group would be considered as delayed.

Table 2. *Operationalizations of Grammatical Complexity and Correctness Measures*

Grammatical complexity:	
MLTU	Mean length of T-units in words (speech disfluencies such as filled pauses, interjections, and repetitions are excluded from this count).
Complex sentences	Total number of complex sentences (= sum of all sentences containing subordinate, infinitival and reduced clauses; conjunction reduction, and direct speech) divided by the total number of T-units.
Subordinate clauses	Total number of subordinate clauses divided by the total number of T-units.
Relative clauses	Total number of relative clauses divided by the total number of T-units.
Grammatical correctness:	
T-units correct	Number of error-free T-units divided by the total number of T-units.

Verb-related errors	Sum of all errors related to verbs (defined below) divided by clauses containing a(t least one) verb: - errors in subject-verb agreement (*hij loop*: he walk). - errors in verb-second placement (*dan de man komt:* then the man comes). - tense errors (present or past tense adverb with an incorrectly tensed verb: *toen*~PAST~ *valt*~PRESENT~ *het meisje*: then falls the girl). - root infinitives (*hij ballon geven:* he balloon give). - past tense verb overregularisations (*hij brengde*: he bringed). - omissions (*Ø)* and substitutions of aspect auxiliaries *zijn/hebben* (be/have) with a past participle (*toen Ø ze naar de clown gelopen*: then Ø they to the clown walked). - past participles errors: deletion of pre- and/or suffix or use of wrong suffix (*hij heeft het meisje (ge-) pak(-t) /gepakken*: he has the girl take). - verb argument structure errors: subject and object omissions were divided by the number of clauses where a subject or object was expected and obligatory (instances of allowed subject or object drop were not counted as errors).
Non-verb related errors	Sum of all non-verb related errors divided by clauses containing a(t least one) verb: all errors in word order, deletion of nouns, substitution and omission errors in determiners, prepositions, pronouns (case, gender and number), conjunctions, omission of adverb '*er*' (there), and errors in adjectival inflection.
Dummy auxiliaries	Number of dummy auxiliaries divided by clauses containing a(t least one) verb.

Results

Grammatical Complexity

This first section of the results concerns grammatical complexity in the narratives of the three groups at the three time points. The descriptives of the grammatical complexity measures and chronological age at the different measurement points are presented in Table 3.

Table 3. *Age, MLTU and Grammatical Complexity Measures (Percentages of Total Number of T-units in the Narratives) for the three Groups at the three Time Points*

	SLI (*n* = 30) *M (SD)*			LA (*n* = 30) *M (SD)*			CA (*n* = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Age (years; *SD* in months)									
	6;5 (1.5)	7;4 (2.1)	8;5 (2.0)	4;7 (2.6)	5;7 (2.6)	6;7 (2.6)	6;6 (2.4)	7;6 (2.4)	8;6 (2.4)
Mean length of T-units in words									
	5.75 (0.77)	5.94 (0.80)	6.24 (0.60)	5.55 (0.85)	6.21 (0.89)	6.53 (0.53)	6.35 (0.78)	6.59 (0.86)	6.84 (0.74)
Complex sentences									
	5.0 (6.9)	7.7 (6.7)	9.6 (7.8)	2.2 (4.7)	5.3 (6.7)	6.6 (5.9)	6.0 (6.8)	9.3 (9.3)	11.8 (9.7)
Subordinate clauses									
	2.6 (4.6)	2.9 (4.0)	4.8 (6.8)	1.2 (3.4)	2.7 (3.8)	2.7 (4.0)	3.2 (4.2)	4.3 (5.5)	4.3 (5.5)
Relative clauses									
	0.0 (0.0)	0.8 (2.0)	0.1 (0.5)	0.4 (1.4)	0.5 (1.4)	0.5 (1.6)	1.0 (2.6)	0.8 (1.7)	1.1 (2.0)

MLTU

For mean length of T-units, no significant interaction was found between Time and Group ($F_{(4,174)}$ = 1.87, p = .118, partial η^2 = .041). However, there was a significant main effect of Group ($F_{(2,87)}$ = 9.66, p < .001, partial η^2 = .182), as well as of Time ($F_{(2,174)} = 24.96$, $p < .001$, partial $\eta^2 = .223$). The changes in *MLTU* over time and differences between groups are presented in Figure 3(a). This schematic outline, with an arbitrary scale on the y-axis, sketches the developmental trajectories of the three groups and differences between the groups. One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)} = 7.99$, $p = .001$), T2 ($F_{(2,87)} = 4.26$, $p = .017$), and T3 ($F_{(2,87)} = 6.97$, $p = .002$). Post-hoc tests revealed that the SLI group had a significantly lower *MLTU* than the CA group at T1 ($p = .015$), at T2 ($p = .014$), and at T3 ($p = < .001$). Between the SLI and the LA group, no significant differences were found at any of the time points. The LA group had a lower *MLTU* than the CA group at T1 ($p = .001$). However, this difference was not significant at T2 and T3. Repeated measures (RM) ANOVA for the separate groups revealed that in the SLI group *MLTU* only increased significantly between T1 and T3 ($F_{(1,29)} = 9.55$, $p = .004$ partial $\eta^2 = .248$). The LA control group showed a significant increase in *MLTU* between T1 and T2 ($F_{(1,29)} = 19.91$, $p < .001$, partial $\eta^2 = .407$). This difference was not significant between T2 and T3 ($F_{(1,29)} = 3.97$, $p = .056$, partial $\eta^2 = .120$). Between T1 and T3, *MLTU* increased significantly ($F_{(1,29)}$) = 31.74, $p < .001$, partial $\eta^2 = .523$). In the CA group, *MLTU* increased significantly between T1 and T3 ($F_{(1,29)} = 9.13$, $p < 005$, partial $\eta^2 = .239$).

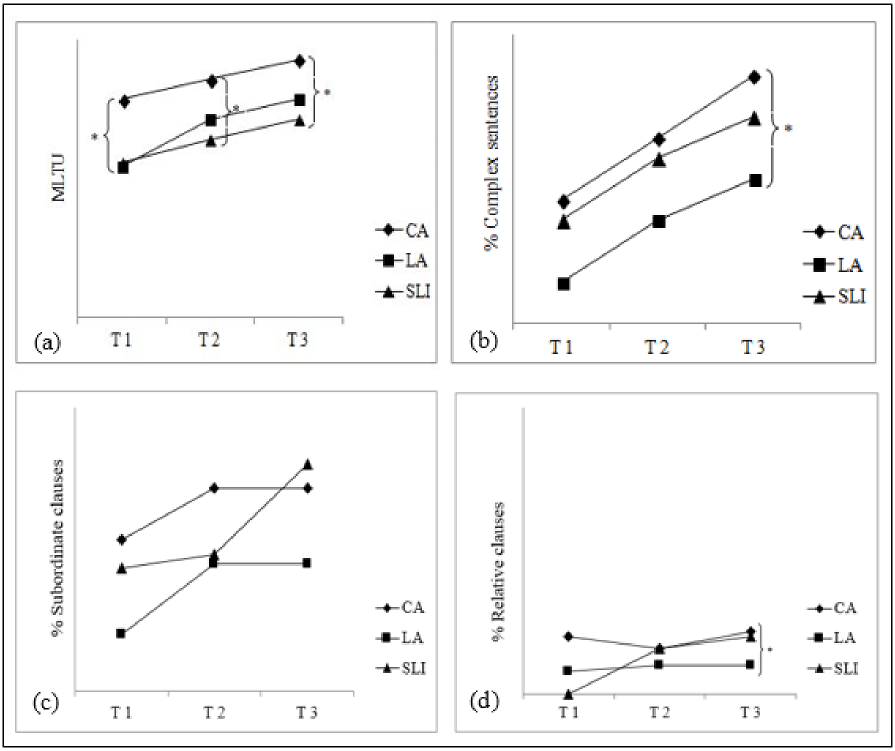

Figure 3. Development of markers of grammatical complexity at T1, T2, and T3 in the three groups: (a) MLTU, (b) Complex sentences, (c) Subordinate clauses, and (d) Relative clauses (schematic outline; * = significant difference between groups).

Complex Sentences

This composite measure represents the number of complex sentences divided by the number of T-units in the narratives. There was no significant interaction between Time and Group ($F_{(4,174)} = 0.13, p = .971$, partial $\eta^2 = .003$). However, there was a significant main effect of Time ($F_{(2,174)} = 12.91, p < .001$, partial $\eta^2 = .129$) and of Group ($F_{(2,87)} = 5.70, p = .005$, partial $\eta^2 = .116$). The use of complex sentences increases steadily but gradually over time in all three groups, as can be seen in Figure 3(b). One-way ANOVAs revealed that the only significant difference between groups was found at T3 ($F_{(2,87)} = 3.27, p = .043$). Post-hoc testing showed that at T3, the CA group used more complex sentences than the LA group ($p = .038$). The children with SLI did not differ from both control groups. RM ANOVAs for the separate groups revealed that in all three groups, differences between T1-T2 and T2-T3 were not significant. However, differences were significant in all groups between T1 and T3, for SLI ($F_{(1,29)}$ = 7.34, p = .011, partial η^2 = .202), for LA ($F_{(1,29)}$ = 8.66, p = .006, partial η^2 = .230) and for CA ($F_{(1,29)}$ = 8.49, p = .007, partial η^2 = .227).

Subordinate clauses

The total number of subordinate clauses was divided by the number of T-units to compute the *percentages of subordinate clauses*. For this measure, no interaction between Time and Group was found ($F_{(4,174)}$ = 0.48, p = .751, partial η^2 = .011). The main effect of Time did not reach significance ($F_{(2,174)}$ = 3.00, p = .053, partial η^2 = .033). The main effect of Group also was not significant ($F_{(2,87)}$ = 2.44, p = .093, partial η^2 = .053). Figure 3(c) illustrates that no significant differences between groups (one-way ANOVAs) or changes over time within groups (RM ANOVAs for the separate groups) were found.

Relative clauses

The measure *percentage of relative clauses* was calculated by dividing counts of relative clauses by the number of T-units. Results for the three groups across time are illustrated in Figure 3(d). Percentages of relative clauses produced in the narratives were low, and only once exceeded 1%. We found no significant interaction between Time and Group ($F_{(4,174)}$ = 1.19, p = .318, partial η^2 = .027). No significant main effect of Time was found ($F_{(2,174)}$ = 0.619, p = .619, partial η^2 = .005) but there was a significant main effect of Group ($F_{(2,87)}$ = 3.48, p = .035, partial η^2 = .065). One-way ANOVAs revealed that the difference between groups was only significant at T3 ($F_{(2,87)}$ = 3.25, p = .043). Post-hoc testing indicated that the SLI group used fewer relative clauses than the CA group at T3 (p = .039). RM ANOVAs for the separate groups revealed no significant differences between time points.

Summary of grammatical complexity

The SLI group had a lower *MLTU* than the CA control group at all three time points, but did not differ from the LA group. The LA children produced fewer complex sentences at age 6 than the CA children at age 8, but the children with SLI did not differ from both control groups at any time point. The use of subordinate clauses did not differ significantly between the three groups at all three time points. However, at age 8 the children with SLI produced fewer relative clauses than the CA group. With respect to the development of grammatical complexity, we found that all three groups showed a significant increase in *MLTU* and percentages of complex sentences between the three time points. Contrastively, percentages of subordinate clauses and percentages of relative clauses did not change over time in the three groups.

Grammatical Correctness and the Use of Dummy Auxiliaries

This second section of the results concerns the grammatical correctness calculated from the narratives. The descriptives of the measures are presented in Table 4.

Table 4. *Grammatical Correctness Measures and Dummy Auxiliaries*

	SLI (n = 30) *M (SD)*			LA (n = 30) *M (SD)*			CA (n = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
T-units correct[a]									
	40.1	49.6	56.3	68.3	77.9	86.3	81.4	91.1	92.8
	(16.1)	(18.0)	(14.7)	(19.8)	(13.2)	(10.2)	(13.1)	(7.1)	(8.6)
Verb-related errors[b]									
	49.2	32.3	25.8	21.1	12.6	5.5	9.3	3.3	3.9
	(29.8)	(20.0)	(14.6)	(21.3)	(10.4)	(6.2)	(8.8)	(3.9)	(4.9)
Non-verb related errors[b]									
	35.7	34.0	27.3	19.7	11.8	6.3	10.0	5.2	3.3
	(14.1)	(14.8)	(11.0)	(10.4)	(11.2)	(8.0)	(8.6)	(5.9)	(5.8)
Dummy auxiliaries[b]									
	20.8	17.8	15.2	11.3	21.9	14.9	16.6	11.7	8.3
	(14.4)	(11.4)	(12.5)	(11.4)	(16.2)	(8.9)	(13.4)	(10.9)	(8.2)

Note. [a]Total number of correct T-units divided by the total number of T-units produced in the narratives. [b]Total number of verb-related errors, non-verb related errors, or dummy auxiliaries divided by the total number of clauses containing a(t least one) verb produced in the narratives.

T-units correct

The *percentages of T-units correct* (i.e., counts of error free T-units divided by the sum of T-units) were regarded as a general measure of grammatical correctness. There was no significant interaction between Group and Time ($F_{(3.39,147.23)}$ = .99, p = .405, partial η^2 = .022). However, there was a significant main effect of Time

($F_{(1.69,147.23)}$ = 43.97, p < .001, partial η^2 = .336) and of Group ($F_{(2,87)}$ = 109.92, p < .001, partial η^2 = .716).). Figure 4(a) illustrates differences between groups for correct T-units and changes over time. One-way ANOVAs showed that significant differences between groups were found at T1 ($F_{(2,87)}$ = 48.59, p < .001), T2 ($F_{(2,87)}$ = 73.76, p < .001) and T3 ($F_{(2,87)}$ = 86.95, p < .001). Post-hoc tests revealed that at all time points the SLI group had fewer correct T-units than the LA and CA control groups (p < .001 for all comparisons). The LA group had fewer correct T-units than the CA group at T1 (p = .009) and T2 (p = .001). At T3 this difference did not reach significance (p = .093). RM ANOVAs indicated that *percentages of T-units correct* increased significantly in the SLI group between T1 and T2 ($F_{(1,29)}$ = 8.68, p = .006, partial η^2 = .230) and T2 and T3 ($F_{(1,29)}$ = 7.16, p = .012, partial η^2 = .198). In the LA group, *percentages of T-units correct* increased significantly between T1 and T2 ($F_{(1,29)}$ = 7.65, p = .010, partial η^2 = .209) and T2 and T3 ($F_{(1,29)}$ = 15.30, p = .001, partial η^2 = .345). In the CA group, the difference between T1 and T2 was significant ($F_{(1,29)}$ = 14.38, p < .001, partial η^2 = .332), but did not reach significance between T2 and T3 ($F_{(1,29)}$ = 0.99, p =.328, partial η^2 = .033).

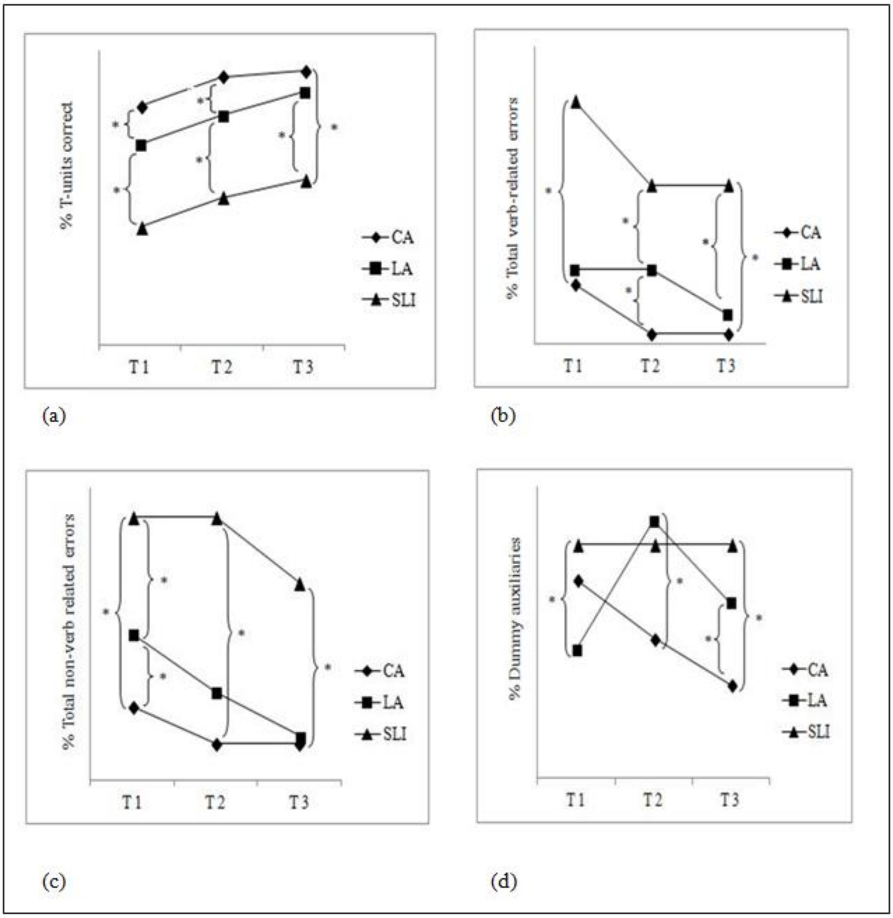

Figure 4. Development of markers of grammatical correctness at T1, T2, and T3 in the three groups: (a) Correct T-units, (b) Verb-related errors, (c) Non-verb related errors, and (d) Dummy auxiliaries (schematic outline, * = significant difference between groups).

Verb-related errors

The *percentage of verb-related errors* were all verb-related errors divided by the total number of clauses containing a(t least one) verb. There was a significant interaction between Group and Time ($F_{(3.10,134.76)} = 3.81$, $p = .011$, partial $\eta^2 = .081$). There also was a significant main effect of Group ($F_{(2,87)} = 57.49$, $p < .001$, partial $\eta^2 = .569$) as well as of Time ($F_{(1.55,134.76)} = 29.57$, $p < .001$, partial $\eta^2 = .081$). Development of *verb-related errors* and group differences are presented in Figure 4(b). One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)}$ = 26.63, p < .001), T2 ($F_{(2,87)}$ = 37.74, p < .001) and T3 ($F_{(2,87)}$ = 48.59, p < .001). Post-hoc tests revealed that at all time points, the SLI group made more verb-related errors than both control groups (p < .001 for all comparisons). At T1, the difference between LA and CA groups was not significant (p = .115). At T2, the difference between LA and CA groups did reach significance (p = .022), but at T3, this difference was not significant (p = 1.00). RM ANOVAs for the separate groups revealed that in the SLI group verb-related errors decreased over time between T1 and T2 ($F_{(1,29)}$ = 20.00, p < .001, partial η^2 = .408) but not between T2 and T3 ($F_{(1,29)}$ = 3.03, p = .092, partial η^2 = .095). In the LA group, verb-related errors did not decrease between T1 and T2 ($F_{(1,29)}$ = 4.02, p = .054, partial η^2 = .0122), but a significant decrease was found between T2 and T3 ($F_{(1,29)}$ = 13.82, p = .001, partial η^2 = .323). In the CA group, the decrease was significant between T1 and T2 ($F_{(1,29)}$ = 11.30, p = .002, partial η^2 = .280), but not between T2 and T3 ($F_{(1,29)}$ = 0.56, p =.461, partial η^2 = .019).

Non-verb related errors

For this measure, the count of all non-verb related errors was divided by the total number of clauses containing a(t least one) verb. There was no significant interaction between Group and Time ($F_{(4,174)}$ = 1.75, p = .142, partial η^2 = .039). However, there was a significant main effect of Group ($F_{(2,87)}$ = 101.84, p < .001, partial η^2 = .701) as well as of Time ($F_{(2,174)}$ = 25.11, p < .001, partial η^2 = .224). Developmental trajectories and group differences are shown in Figure 4(c). One-way ANOVAs revealed significant differences between groups at T1 ($F_{(2,87)}$ = 39.70, p < .001), at T2 ($F_{(2,87)}$ = 53.94, p < .001) and at T3 ($F_{(2,87)}$ = 70.92, p < .001). Post-hoc tests revealed that at all time points the SLI group produced more non-verb related errors than both control groups (p < .001 for all comparisons). At T1, the LA group had more errors than the CA group (p = .004). At T2 and T3, differences between LA and CA children were not significant (p = .077 and p = .541 respectively). RM ANOVAs for the separate groups revealed that *percentages of non-verb related errors* decreased significantly in the SLI group between T2 and T3 ($F_{(1,29)}$ = 6.32, p = .018, partial η^2 = .179), but not between T1 and T2 ($F_{(1,29)}$ = 0.23, p = .633, partial η^2 = .008). In the LA group, *percentages of non-verb related errors* decreased significantly between T1 and T2 ($F_{(1,29)}$ = 12.49, p = .001, partial η^2 = .301) and between T2 and T3 ($F(_{1,29})$ = 9.44, p = .005, partial η^2 = .246). In the CA control group, the decrease was significant between T1 and T2 ($F_{(1,29)}$ = 9.76, p = .004, partial η^2 = .252), but not between T2 and T3 ($F_{(1,29)}$ = 2.05, p = .163, partial η^2 = .066).

Dummy auxiliaries

To compute *percentages of dummy auxiliaries*, counts of dummy auxiliaries were divided by clauses containing a(t least one) verb. There was a significant interaction between Group and Time ($F_{(4,174)} = 5.82$, $p < .001$, partial $\eta^2 = .118$). There was a significant main effect of Group ($F_{(2,87)} = 3.17$, $p = .047$, partial $\eta^2 = .068$) as well as of Time ($F_{(2,174)} = 4.76$, $p = .010$, partial $\eta^2 = .052$). Development in the three groups is presented in Figure 4(d). One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)} = 4.00$, $p = .022$), at T2 ($F_{(2,87)} = 4.66$, $p = .012$) and at T3 ($F_{(2,87)} = 4.55$, $p = .013$). Post-hoc tests revealed that at T1, the SLI group used more dummy auxiliaries than the LA group ($p = .018$). The difference between the children with SLI and the CA group was not significant ($p = .664$). At T2, no significant differences were found between the SLI group and the control groups. However, the LA group used significantly more dummy auxiliaries than the CA group ($p = .009$). At T3, the SLI group used more dummy auxiliaries than the CA group ($p = .028$), and the LA group also used more dummies than the CA group ($p = .037$). No difference was found between SLI and LA groups. RM ANOVAs for the separate groups revealed that in the SLI group the use of dummy auxiliaries did not decrease significantly between any of the time points. In the LA group, the use of dummy verbs increased significantly between T1 and T2 ($F_{(1,29)} = 16.07$, $p < .001$, partial $\eta^2 = .356$), followed by a significant decrease between T2 and T3 ($F_{(1,29)} = 5.99$, $p = .021$, partial $\eta^2 = .171$). In the CA group, the use of dummy verbs only decreased significantly between T1 and T3 ($F_{(1,29)} = 10.33$, $p = .003$, partial $\eta^2 = .263$).

Summary of grammatical correctness and the use of dummy auxiliaries

At all three time points, the SLI group performed more poorly on the measure *percentages of correct T-units* compared to LA and CA group. Both the SLI and LA groups improved steadily over time, and the CA group seemed to reach a plateau at age 8 (93% correct). With respect to *percentages of verb-related errors*, the SLI group was again outperformed by LA and CA groups at all time points. The SLI group and the CA group improved significantly between the ages 6 and 7, whereas the LA group improved significantly between the ages 5 and 6. The children with SLI also performed poorer than the LA and CA groups on *percentages of non-verb related errors*. On this measure, the SLI group improved between the ages 7 and 8, the LA group improved steadily between the ages 4, 5, and 6. The CA group improved between the ages 6 and 8. With respect to the use of *dummy auxiliaries*, no development across time points was seen in the SLI group. The LA children showed an increase followed by a decrease, and in the CA group the use of dummy verbs decreased steadily but gradually.

Discussion

The purpose of this study was to identify delay or deviance patterns in the development of grammatical complexity and grammatical correctness in Dutch school-age children with SLI. The delay versus deviance issue was investigated in Dutch children with SLI for the first time. Narratives were used to analyze a wide range of grammatical complexity and grammatical correctness measures (Blankenstijn & Scheper, 2003). Our first observation is that comparing the SLI and TD groups with a standardized language test yields a different picture than results obtained from a narrative analysis. Scores on all subtests of the LPT at T1 indicate that the SLI group lags two years behind the CA group, and performs just like the LA control group. On the basis of the LTP-scores alone, one could be inclined to label the SLI group as simply two years delayed. However, our elaborate analysis of the narratives provides a more differentiated picture. Results on grammatical complexity and correctness measures will be discussed and related to the developmental patterns described in the literature.

Grammatical Complexity

No significant differences between the SLI group and the LA and CA groups were found for the grammatical complexity measures *complex sentences* and *subordinate clauses*. For these two measures, neither delay nor deviance was found. Nevertheless, the 8-year-old children with SLI did use fewer *relative clauses* than the CA group. Furthermore, mean length of T-unit (*MLTU*) in the SLI group did not differ from the LA group but remained lower compared to the CA group at all three time points. The results on the measures *MLTU* and *relative clauses* would seem to fit a developmental delay pattern. The children with SLI performed similarly to the LA group and showed no sign of catching up. However, whether the SLI group eventually will reach a plateau cannot be inferred from our study. Up to age 8, there is still growth in *MLTU* levels, and relative clause production may not have attained its full development at age 8.

The finding that relative clauses are problematic for older children with SLI has also been reported by other studies (Schuele & Nicholls, 2000; Schuele & Tolbert, 2001; Marinellie, 2004; Novogrodsky & Friedmann, 2006). The question is why relative clauses are so difficult for children with SLI. Hestvik, Schwartz, and Tornyova (2010) used a cross-modal picture naming experiment to examine automatic on-line gap-filling in relative clauses in children with SLI. The same sentences were also used in an off-line comprehension task. The children with SLI showed lack of immediate gap-filling after the relative clause verb. On the comprehension task the SLI group performed as well as the controls. Hestvik et al. (2010) argued that processing mechanisms in children with SLI are impaired, but their grammatical knowledge is not. This processing explanation is in accordance with the poor performance on the Kaufman ABC sequential memory tasks of the SLI group in the present study. This limitation in sequential memory abilities might affect the production of relative clauses. This argument is also supported by Coco, Garraffa, and Branigan (2012), who found that poor performance on subject relative clauses was related to working memory constraints.

An opposing view by Marinellie (2004) holds that problems with relative clauses are not caused by grammatical features or working memory demands, but by their semantic properties. The function of relative clauses is to further specify a noun, and lexical knowledge of descriptive, specific vocabulary is needed for this specification. Lexical deficits, such as a sparse vocabulary and poor lexical access and retrieval, are suggested to have a negative effect on the production of relative clauses.

Yet another explanation concerns the exposure of children with SLI to complex syntax. Relative clause production might remain poor in children with SLI because the input frequency of these sentence types could be too low (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Sarilar, Matthews, & Küntay, 2013). Arguably, many parents of children with SLI also experience language problems and possibly will not use these sophisticated grammatical constructions extensively. Furthermore, in special education settings, teachers tend to simplify their language output to adapt to the poor language skills of their language impaired pupils. As a consequence, children might not encounter enough exemplars to learn these complex sentence structures.

To conclude, the poor results of the children with SLI on relative clause production may be adequately explained by either processing limitations, semantic deficits, or a limited input frequency. These explanations may even be interrelated and demand further research.

Grammatical Correctness and the Use of Dummy Auxiliaries

The analysis of grammatical correctness yielded more differences between the children with SLI and the TD control groups. On all three grammatical correctness measures, namely *T-units correct*, *verb-related errors* and *non-verb related errors*, the SLI group performed much poorer than both LA and CA control groups. The results of the SLI group on *MLTU* and all standardized LPT subtests would imply a plain two year delay. When the very poor results on the grammatical correctness measures are also taken into account, development in the SLI group would best fit a delay-within-delay pattern (Rice, 2003) or in Leonard's (1998) terms, an uneven profile.

At age 7, the CA group already obtained a high level of grammatical correctness and appears to have reached ceiling. The SLI and LA groups both showed continuous improvement on *percentages T-units correct* between the three measurement points. The examination of *verb-related errors* and *non-verb related errors* revealed even more differences in developmental trajectories between the three groups. For instance, in the SLI group *verb-related errors* decreased between T1 and T2, and between T2 and T3 stagnation was found. In the LA group, this pattern was reversed. On the other hand, *non-verb related errors* decreased steadily in the LA group, but in the SLI group only after age 7. These different developmental trajectories for *verb-related errors* and *non-verb related errors* in the SLI and LA groups, indicate that children with SLI do not always follow the same developmental trajectories as TD children. This result differs from the "tracking hypothesis" proposed by Law, Tomblin, and Zhang (2008) which claims that language in children with SLI takes off slower, but once started, parallels that of TD children. They state that children with SLI do not get better or worse, rather they tend to stay on the same trajectory. It must be noted that Law et al. (2008) only used a standardized language test to examine developmental trajectories of receptive language in SLI and their study did not include TD control groups.

The different developmental trajectories for *verb-related errors* and *non-verb related errors* in the SLI group demand an explanation. As all children with SLI were enrolled in special schools for children with severe speech and language impairments, specific goals and intensity of intervention may have influenced the outcomes. In this educational setting, speech therapy is most intensive until the end of grade 3 (around age 7). Possibly, therapy goals were centered at verb morphology, and a reduction in therapy after age 7, may have contributed to the stagnation observed for *verb-related errors*. This speculation immediately leads to the question why then *non-verb related errors* only started to decrease in the 8-year-old SLI group. The fact that the 8-year-olds have attained a certain level of reading proficiency may offer an explanation. Determiners, pronouns, and adverb '*er*' (there) formed the most prominent error categories in *non-verb related errors*. In Dutch, these functional elements can all be regarded as having a low phonetic saliency. These monosyllabic items are short, unstressed and often contain schwas. According to Leonard's "surface hypothesis" (1989) grammatical items with weak surface characteristics, such as unstressed monosyllabic words, can easily be neglected in input and output. However, in written language these monosyllabic items are all surrounded by spaces and are thus easier to identify. This additional visual information may help children with SLI to build up more robust mental representations of these functional items.

The results from the present study show that even at age 8, the percentages of *non-verb related errors* remain substantial in the SLI group. Perhaps, in the end, verb morphology rules turn out to be easier to learn for children with SLI than other morphosyntactic rules. For instance, verb inflection rules can be regarded as rather transparent, but determiner-noun combinations in Dutch all have to be learned. Older Dutch school-age children with SLI are still at the start of learning the gender paradigm for determiners and continue to overgeneralize common gender for neuter gender determiners (Keij et al., 2012). The rules for the correct use of adverb '*er*' (there) are known to be very opaque. Furthermore, pronouns and prepositions can be viewed as less fixed or stable in language use. The correct use of pronouns and prepositions partly depends on discourse requirements. In discourse, speakers and listeners have to deal with interlocutor perspectives, which is in turn related to the development of theory of mind. According to Farrant, Fletcher, and Maybery (2006), the development of visual perspective-taking abilities and theory of mind are delayed in children with SLI. It is argued that the limited perspective taking skills of children with SLI contribute to delays and errors in the acquisition of pronouns (Bol & Kasparian, 2009).

From a different perspective, the relatively high non-verb related error rates may be related to the limited executive functioning skills of children with SLI. There is growing evidence that children with SLI show weaknesses in executive functioning (Gillam, Montgomery, & Gillam, 2009; Henry, Messer, & Nash, 2012). In a narrative, the speaker must retain in memory to a certain extent what has been said already, and in what form. Consequently, executive functions such as attention, inhibition and working memory can be supposed to influence narrative skills. For instance, Seiger-Gardner and Schwartz (2008) found that previously uttered words remained active for a longer period in older children with SLI than in TD control groups. In the present study, a large portion of the errors made by the SLI group were substitution errors in closed class items. In addition, perseverations within one story or between the first and the second story were also observed. Poor inhibition skills in the children with SLI may explain these substitution errors and perseverations.

Finally, the use of *dummy auxiliaries* in the SLI, LA, and CA groups is considered. An overuse of dummy auxiliaries can be regarded as an immature stage of grammatical development. The inverted U-shape of *percentages of dummy auxiliaries* in the LA group in this study forms a puzzling result for which we have no plausible explanation. In the CA control group, *percentages of dummy auxiliaries* decreased steadily, whereas in the SLI group this was not the case. The pattern observed in the SLI group fits a developmental plateau in Leonard's (1998) taxonomy. As this pattern does not resemble the trajectories of either the LA or CA control groups, the use of dummy verbs could be interpreted as deviant compared to normal development. The percentages of dummy verbs in the SLI group in our study correspond roughly with those reported in Orgassa (2009). She reported that Dutch children with SLI (mean age 7;3; range 6;3-8;5) produced 22% dummy auxiliaries in an elicitation task targeted at verb inflection. Orgassa (2009) found that the use of dummy verbs decreased with age in the SLI group. However, this was not confirmed in the present study, where percentages of dummy verbs in the SLI group did not change significantly between the ages 6 and 8. In our study, the unchanging use of dummy auxiliaries paralleled a steady decrease of error rates in subject-verb agreement and verb placement (see Appendix 1). Apparently, the children with SLI learn the verb inflection paradigms and verb-second rule. This would render insertion of dummy verbs as an "economy strategy" superfluous, and the use of dummy verbs would be expected to fade out. On the basis of our results, the explanation that dummy verbs are used to avoid verb placement and inflection (de Jong, 1999; de Jong et al., 2013; Orgassa, 2009), should be reconsidered. Possibly, the unchanged rates of dummy auxiliaries should be interpreted as a form of fossilization. The children with SLI have used this 'strategy' so extensively that dummy verbs are prone to activation and selection from the mental lexicon. This interpretation would also be in line with the poor inhibition skills observed in children with SLI (Seiger-Gardner & Schwartz, 2008).

The overuse of dummy auxiliaries might also be related to poor lexical retrieval in children with SLI. A good part of the children with SLI experience word-finding difficulties. Conceivably, the insertion of a dummy verb at verb-second position might function as a stalling device, offering the children extra time to retrieve the lexical verb. By using a dummy verb, production of the lexical verb is postponed to clause-final position. This stalling hypothesis could be investigated further with a cross-modal priming task or by investigating silent and filled pauses preceding inflected lexical verbs compared to stalling behavior before dummy verbs.

To conclude, in the present study the children with SLI did not show a simple delay compared to LA and CA control groups, as argued by Curtiss et al. (1992). In our analysis a number of different developmental trajectories for the different grammaticality measures emerged. On the grammatical complexity measures, the SLI group exhibited no delay at all, except for *MLTU* at all three time points and at age 8 for *relative clauses*. The results on the grammatical correctness measures *T-units correct*, *verb-related errors* and *non-verb related errors* best fit a delay-within-delay model (Rice, 2003), or an uneven profile (Leonard, 1998). On these measures, the SLI group performed very poorly compared to the LA controls. The continuous and unchanging use of *dummy auxiliaries* in the SLI group would best fit a plateau, and might even be regarded as a form of fossilization. As this profile was not observed in the CA and LA control groups, this pattern could be viewed as deviant compared to normal development.

This study has revealed different developmental trajectories for grammatical complexity measures and correctness measures in Dutch school-age children with SLI. These trajectories did not always match those of the LA group. When standardized tests are used exclusively, only delay patterns are found in children with SLI. The combination of standardized tests with the analysis of narrative tasks in a longitudinal three group design appears to be an appropriate method to investigate developmental trajectories in children with SLI.

Finally, a positive finding of this study is that the children with SLI showed continuous progress on some of the grammatical complexity measures and all correctness measures. Whether this improvement is the merit of the special education system, or just a reflection of a slowed and prolonged development, is a question that cannot be answered here. Although the children with SLI in this study continued to have severe language problems, no evidence was found for an overall stagnation in grammatical development. A follow-up study applying this longitudinal three-group design to an older school-age SLI group and TD control children would be very informative. Such a study could reveal whether the grammatical skills will continue to develop, and whether the observed plateaus or stagnation patterns in the SLI group found in this study are permanent or transitory.

Clinical Implications

Narrative tasks have considerable diagnostic value next to standardized tests. A narrative analysis of grammatical complexity and grammatical correctness can offer a more detailed evaluation than can be obtained with standardized tests. MLTU and the grammatical correctness measures differentiated school-age children with SLI from their age peers at all three measurement points. Grammatical profiles obtained from narratives can inform clinicians in choosing adequate therapy goals and can also be useful in evaluating the effects of intervention. The present study also showed that SLI is a dynamic condition and that grammatical skills of school-age children with SLI continue to develop. This was especially the case for the grammatical correctness. Therefore, the provision of language intervention beyond grade 3 still seems beneficial for children with SLI.

Acknowledgments

This study was supported by the Dutch Royal Auris Group, Gouda, The Netherlands. We thank Cito for their permission to use the images of the LPT Storytelling tasks. We also thank the children, parents and staff of all the schools that participated; and Marij van Ewijk, Henaly Leijenhorst, Marjolein van der Horst, and Merel van Witteloostuijn for their help with the transcriptions.

References

Blankenstijn, C., & Scheper, A. (2003). *Language development in children with psychiatric impairment*. Utrecht: LOT Dissertation Series 82. (Doctoral dissertation, University of Amsterdam). Retrieved from http://www.lotpublications.nl/publish/articles/000542/bookpart.pdf.

Bol, G. W., & Kasparian, K. (2009). The production of pronouns in Dutch children with developmental language disorders: A comparison between children with SLI, hearing impairment, and Down's syndrome. *Clinical Linguistics and Phonetics*, *23*, 631–646.

Bol, G. W., & Kuiken, F. (1990). Grammatical analysis of developmental language disorders: A study of the morphosyntax of children with specific language disorders, with hearing impairment and with Down's syndrome. *Clinical Linguistics and Phonetics*, 4, 77-86.

Coco, M. I., Garraffa, M., & Branigan, H. P. (2012). Subject relative production in SLI children during syntactic priming and sentence repetition. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012)*, Sapporo, Japan, August, 2012.

Curtiss, S., Katz, W., & Tallal, P. (1992). Delay versus deviance in the language acquisition of language-impaired children. *Journal of Speech and Hearing Research, 35*, 373–383.

De Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure*. Groningen: Grodil. (Doctoral dissertation, University of Groningen). Retrieved from: http://irs.ub.rug.nl/ppn/297976826.

De Jong, J. (2003). Specific language impairment and linguistic explanation. In Y. Levy and J. Schaeffer (Eds.) *Language competence across populations. Toward a definition of specific language impairment* (pp. 151-170). Mahwah: Lawrence Erlbaum Associates.

De Jong, J. Blom, E. Orgassa, A. (2013). Dummy auxiliaries in children with SLI - a study on Dutch, in monolinguals and bilinguals. In E. Blom, I. van de Craats, & J. Verhagen (Eds), *Dummy auxiliaries in first and second language acquisition* (pp. 251-278). Berlin: De Gruyter Mouton.

Farrant, B. M., Fletcher, J., & Maybery, M. T. (2006). Specific language impairment, theory of mind, and visual perspective taking: evidence for simulation theory and the developmental role of language. *Journal of Child Development*, 77, 1842–1853.

Gillam, R. B., Montgomery, J. W., & Gillam, S.L. (2009). Attention and memory in child language disorders. In R. G. Schwartz (Ed.). *The handbook of child language disorders* (pp. 201-215). New York, NY: Psychological Press.

Hestvik, A., Schwartz, R. G., & Tornyova, L. (2010). Relative clause gap-filling in children with specific language impairment. *Journal of Psycholinguistic Research*, *39*, 443–56.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry*, *53*(1), 37–45.

Hunt, K. W. (1970). Syntactic maturity in school children and adults. *Monograph of the Society for Research in Child Development, 35*(1), 1-67.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S.C. (2002). Language input at home and at school: Relation to syntax. *Cognitive Psychology*, 45, 337-374.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children: Interpretative Manual.* Circle Pine, MN: American Guidance Service.

Keij, B., Cornips, L., Van Hout, R., Hulk, A., & Emmerik, J. van (2012). Knowing versus producing: The acquisition of grammatical gender and the definite determiner in Dutch by L1-TD, L1-SLI, and eL2 children. *Linguistic Approaches to Bilingualism*, *2*, 379–403.

Law, J., Tomblin, J. B., & Zhang, X. (2008). Characterizing the growth trajectories of language-impaired children between 7 and 11 years of age. *Journal of Speech, Language, and Hearing Research*, *51*, 739-749.

Leonard, L.B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10*, 179-202.

Leonard, L. B. (1996). Assessing morphosyntax in clinical settings. In D. McDaniel, C. McKee, and H. Smith Cairns (Eds.) *Methods for Assessing Children's Syntax (pp.* 288-302). Cambridge, MA: MIT Press.

Leonard, L. B. (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Marinellie, S. A. (2004). Complex syntax used by school-age children with specific language impairment (SLI) in child-adult conversation. *Journal of Communication Disorders*, *37*, 517–533.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.

Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *International Journal of Speech-Language Pathology 8*, 364–375.

Orgassa, A. (2009). *Specific language impairment in a bilingual context. The acquisition of Dutch by Turkish-Dutch learners*. Utrecht: LOT Dissertation Series 220.

(Doctoral dissertation, University of Amsterdam). Retrieved from: http://www.lotpublications.nl/publish/articles/003604/bookpart.pdf.

Paradis, J. (2010). Comparing typically developing children and children with specific language impairment. In E. Blom and S. Unsworth (Eds.) *Experimental Methods in Language Acquisition* (pp. 223-243). Amsterdam: John Benjamins.

Parisse, C., & Maillart, C. (2009). Specific language impairment as systemic developmental disorders. *Journal of Neurolinguistics, 22*, 109-122.

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy and J. Schaeffer (Eds.) *Language competence across populations. Toward a definition of specific language impairment (*pp. 63-95). Mahwah: Lawrence Erlbaum Associates.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense Over Time: The longitudinal course of tense acquisition in children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1412–1431.

Reilly, J., Losh, M., Bellugi, U. & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language, 88*, 229-247.

Renfrew, C. (1969). *The Bus Story: A Test of Continuous Speech*. North Place, Old Headington: Oxford.

Sarilar, A., Matthews, D., & Küntay, A. C. (2013). Hearing relative clauses boosts relative clause usage (and referential clarity) in young Turkish language learners. *Applied Psycholinguistics*, available on CJO2013. doi: 10.1017/S0142716413000192.

Seiger-Gardner, L., & Schwartz, R. G. (2008). Lexical access in children with and without specific language impairment: A cross-modal picture-word interference study. *International Journal of Language and Communication Disorders*, *43*, 528–551.

Schuele, C. M., & Nicholls, L. (2000). Relative clauses: evidence of continued linguistic vulnerability in children with specific language impairment. *Clinical Linguistics and Phonetics*, *14*, 563–585.

Schuele, C. M., & Tolbert, L. (2001). Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics and Phonetics*, *15*, 257–274.

Southwood, F. (2007). *Specific language impairment in Afrikaans. Providing a Minimalist account for problems with grammatical features and word order* (Doctoral dissertation). Utrecht: LOT Dissertation Series 166. Retrieved from: http://www.lotpublications.nl/publish/articles/002526/bookpart.pdf.

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*, *52*, 336–358.

Van Bon, W. H. J. (1986). *Coloured Progressive Matrices: Manual of Dutch Norms.* London: Lewis.

Van Kampen, N. J. (1997). *First Steps in Wh-movement* (Doctoral dissertation, Utrecht University). Delft: Eburon.

Van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry, 47*, 176-189.

Verhoeven, L. & Vermeer, A. (2001). *Taaltoets Alle Kinderen* [LPT, Language Proficiency Test for All Children]. Arnhem: Cito.

Wetherell, D., Botting, N., & Conti-Ramsden, G. (2007). Narrative in adolescent specific language impairment (SLI): A comparison with peers across two different narrative genres. *International Journal of Language and Communication Disorders,* 42, 583-605.

Wijnen, F., & Verrips, M. (1998). The acquisition of Dutch syntax. In S. Gillis, and A. De Houwer (Eds) *The Acquisition of Dutch* (pp. 223-299). Amsterdam: John Benjamins.

Appendix 1

Percentages of verb-related errors and the subtypes of verb errors (total counts of errors divided by the total counts of clauses with a(t least one verb)) in the three groups at the three time points.

	SLI (*n* = 30) *Mean (SD)*			LA(*n* = 30) *Mean (SD)*			CA(*n* = 30) *Mean (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Verb-related errors	49.2 (29.8)	32.3 (20.0)	25.8 (14.6)	21.1 (21.3)	12.6 (10.4)	5.5 (6.2)	9.3 (8.8)	3.3 (3.9)	3.9 (4.9)
Agreement errors	13.7 (11.6)	8.1 (8.0)	4.6 (4.4)	7.6 (8.6)	4.7 (7.8)	1.7 (2.9)	2.9 (5.7)	1.2 (2.0)	0.8 (1.6)
Tense errors	5.0 (6.5)	6.5 (6.9)	5.0 (7.5)	1.9 (4.7)	2.3 (6.5)	1.3 (2.9)	1.1 (2.4)	0.8 (1.7)	1.0 (2.1)
Over-regularizations	1.8 (4.7)	1.7 (2.8)	2.6 (3.4)	1.4 (3.7)	0.5 (1.5)	0.8 (1.9)	1.9 (4.3)	0.1 (0.6)	1.1 (2.9)
Verb-second errors	3.5 (6.1)	1.9 (4.7)	2.5 (3.0)	2.6 (10.2)	1.0 (1.9)	0.0 (0.0)	0.4 (1.7)	0.1 (0.6)	0.4 (1.1)
Root infinitives	3.0 (6.1)	1.1 (2.7)	0.7 (1.6)	1.7 (4.6)	0.2 (1.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Subject omissions	8.3 (8.2)	4.9 (4.3)	4.5 (4.5)	3.7 (8.7)	1.0 (2.0)	0.4 (1.5)	0.8 (2.0)	0.2 (0.9)	0.3 (1.3)
Object omissions	13.8 (13.0)	7.7 (9.9)	8.8 (9.9)	4.5 (8.0)	4.8 (7.9)	1.2 (4.3)	1.3 (4.3)	0.2 (1.1)	0.3 (1.5)

3 The role of vocabulary in grammatical development of school-age children with SLI [2]

Abstract

The grammatical development of 30 Dutch-speaking children with SLI, aged 8 years at first round of testing, 30 language age-matched (LA) and 30 chronological age-matched (CA) control children was followed for three years. Narrative tasks were used to investigate development in grammatical complexity and grammatical correctness. Special attention was given to the overuse of dummy auxiliaries as a sign of delayed grammatical development. Results on grammatical complexity and correctness in the SLI group were related to their vocabulary development, assessed by means of standardized receptive and expressive vocabulary tests.

The grammatical skills of the children with SLI improved over the years with regard to complex sentences and verb-related errors, although stagnation was observed for non-verb related errors. On the complexity measures, performance of the SLI group was intermediate between the LA and CA control groups. The 10-year-old children with SLI produced fewer complex sentences than the CA controls and lagged behind in their use of complex-compound sentences. On the correctness measures, the SLI group was outperformed by both LA and CA control groups at all three time points. Individual grammatical correctness scores of the children with SLI were strongly associated with their vocabulary knowledge, most markedly so for non-verb related errors and the use of dummy auxiliaries.

Findings indicate a double deficit in children with SLI. Processing difficulties limit the use of complex syntax and stagnation of non-verb related errors hints at incomplete linguistic knowledge. Results of the correlation analyses between vocabulary and measures of grammatical errors and the use of

[2] To be submitted (Authors: Rob Zwitserlood[a,c], Marjolijn van Weerdenburg[b], Frank Wijnen[a], and Ludo Verhoeven[b]), [a]Utrecht Institute of Linguistics (OTS), University Utrecht, The Netherlands, [b]Behavioural Science Institute, Radboud University Nijmegen, The Netherlands, [c]Koninklijke Auris Groep, The Netherlands).

dummy auxiliaries lend some support to the procedural deficit hypothesis for SLI.
Keywords: specific language impairment (SLI), Dutch, grammatical complexity, grammatical correctness, vocabulary.

The definition of specific language impairment (SLI) is based on exclusionary criteria. SLI is characterized by difficulties with language acquisition that are not caused by identified neurological, sensory, intellectual or emotional deficits (Leonard, 1998). Children with SLI form a heterogeneous group (Stark & Tallal, 1981) and SLI can be regarded as a dynamic condition (van Weerdenburg, Verhoeven, & van Balkom, 2006) with changing (phenotypical) characteristics as the child matures (Conti-Ramsden & Botting, 1999). In spite of this phenotypical volatility, difficulties with grammatical morphology and syntax appear to be a hallmark symptom of SLI (Leonard, 1998; Oetting & Hadley, 2009). Difficulties in other language domains (e.g., phonology, vocabulary, and pragmatics) have been reported as well, but appear less dominant (Schwartz, 2009).

To date, longitudinal research on grammatical development, especially in older school-age children with SLI (i.e., from age 8 upwards), has remained scarce. It is often reported that grammatical development in children with SLI lags behind two years or more relative to their unaffected peers (Curtiss, Katz, & Tallal, 1992; Paradis, 2010; Rice, 2003). However, this lag is typically seen in young children, and little is known about its development at a later age. It remains unclear whether the development of grammatical skills in children with SLI levels off at a certain age, or that continued progress can be observed. For this reason, we investigated the development of grammatical complexity and grammatical correctness in school-age children with SLI. In addition, we related grammatical development to vocabulary development. In typical language acquisition, lexical development is predictive of grammatical development (Bates & Goodman, 1997; Marchman & Bates, 1994; Tomasello, 2003). Although there is some evidence for similar associations between vocabulary and grammar in clinical groups, this relationship has so far not been studied (longitudinally) in school-age children with SLI by taking both grammatical complexity and grammatical correctness into account.

Grammatical Complexity in SLI

Research on the acquisition of complex sentence structures in children with SLI has remained limited. Schuele and Dykes (2005) reported that, on average, children with SLI begin to produce complex sentences two years *later* than typically developing (TD) children. School-age children with SLI also use *fewer* complex sentences than TD children (Norbury & Bishop, 2003; Reilly, Losh, Bellugi, & Wulfeck, 2004). Specifically, comprehending and producing relative clauses is particularly difficult for children with SLI across different languages, e.g., English (Schuele & Dykes, 2005; Schuele & Nicholls, 2000; Schuele & Tolbert, 2001), Swedish (Hakansson & Hansson, 2000), Hebrew (Novogrodsky & Friedmann, 2006), Italian (Contemori & Garraffa, 2010), German (Koch et al., 2013), and Greek (Stavrakaki, 2001). Furthermore, the production of complex-compound clauses (i.e., sentences containing coordination and one or more dependent clauses) also remains difficult for children with SLI (Marinellie, 2004). Thus, the overall picture is that children with SLI *can* use complex sentential structures, but do so markedly less often than their TD peers. Furthermore, their complex sentences contain more errors than those of TD children (Gillam & Johnston, 1992; Schuele and colleagues, 2000; 2001; 2005). This suggests that producing complex structures exerts too heavy a pressure on their processing capacities.

At present, little is known about the growth trajectories for complex syntax in children with SLI. It is still far from clear whether children make progress in their use of different types of complex sentences and to what extent the grammatical correctness of these sentences improves over time.

Grammatical Correctness in SLI

The evidence for morphosyntactic difficulties in children with SLI is overwhelming (for an overview, see Leonard, 1998; Oetting & Hadley, 2009). Up to age 8, difficulties with verb morphology are very prominent in English speaking children with SLI, whereas deficits in functional categories such as determiners, pronouns and conjunctions are also often reported. However, the exact nature of the morphosyntactic difficulties is strongly related to characteristics of the language that is learned. For instance, in languages that only have one definite determiner, such as English, determiner omission is the predominant error made by children with SLI. In languages with more determiners, such as Dutch, substitutions also tend to occur and become the prominent error as the children get older (Keij, Cornips, Van Hout, Hulk, & Emmerik, 2012). Other factors such as the phonetic salience of functional morphemes also determine the error profiles of children with SLI in a given language, as grammatical items with a high salience tend to be easier for children with SLI (Leonard, 1989; 1998).

Older school-age children with SLI appear to have mastered verb morphology rules to a large extent. In children with a chronological age of 7;9 years, measures of verb morphology could not distinguish between SLI and TD groups (Moyle, Karasinski, Ellis Weismer, & Gorman, 2011). However, it is still unclear to what extent older school-age children with SLI reach mastery levels in functional elements such as determiners, pronouns, prepositions, and conjunctions. In fact, grammatical correctness appears to remain vulnerable in school-age children with SLI, and may easily degrade when communicative or conceptual demands are high (Colozzo et al., 2011).

Relations between Vocabulary and Grammar

Vocabulary acquisition is a complex and extended process involving the integration of phonological, semantic and grammatical knowledge with cognitive and social processes (Dockrell & Messer, 2004, p. 35). For typical language acquisition it has been argued that development of the lexicon and grammar are strongly related (Bates & Goodman, 1997). Lexical development, specifically attaining a critical mass of the lexicon, has been hypothesized to trigger the onset of multi-word utterances and, thus, grammar. Furthermore, children make use of lexical knowledge to bootstrap morphosyntax (Marchman & Bates, 1994; Tomasello, 2003). Later, they use morphosyntactic knowledge to bootstrap new word meanings (Naigles, 1990). Presumably, the relative importance of lexical versus syntactic bootstrapping changes with development (Dionne, Dale, Boivin, & Plomin, 2003). According to Tomblin and Zhang (2006), grammatical abilities and vocabulary abilities are intertwined in young children and only become two-dimensional during middle childhood, but McGregor et al. (2012) propose a continuous reciprocal support between lexical and syntactic bootstrapping in older language learners.

Although children with SLI show on average a one year delay in the onset of their first words (Trauner, Wulfeck, Tallal, & Hesselink, 1995) and generally perform below age-level on a variety of lexical measures, the literature usually presents vocabulary skills as a relative strength. Indeed, as a group, children with SLI do have greater deficits in syntax than in lexical semantics (Rice, 2003; Tomblin & Zhang, 1999). Nonetheless, children with SLI have difficulties acquiring new words, especially verbs (Dollaghan, 1987; Leonard, 1998). Children with SLI seem to be poor at constructing full phonological specifications of words and adequate semantic representations (Dockrell & Messer, 2004). On the one hand, phonological errors in naming tasks, especially in longer words, are seen as evidence for imprecise or inadequate phonological representations. On the other hand, word-finding difficulties, which affect many children with SLI, have been associated with poor semantic representations (Kail & Leonard, 1986). Most children with SLI never reach age-appropriate vocabulary levels (McGregor, Oleson, Bahnsen, & Duff, 2013) and their vocabulary knowledge remains sparse with regard to breadth (i.e., the number of words in the lexicon), as well as depth (i.e., the richness of the representations of words).

Leonard and Deevy (2004) sketch the difficulties children with SLI may encounter because of their lexical deficits. Most obviously, communication is seriously hindered when the words needed to frame one's thoughts are not available or when the words in the received message are not understood. In addition, a poor vocabulary can have an adverse effect on reading proficiency. Finally, poor representations of words (i.e., their phonological, semantic and morphosyntactic features) can be detrimental to the learning of sentence structure and grammatical morphology.

When measured with standardized tests, the variation in vocabulary size tends to be substantially larger in children with SLI than in TD children (Sheng & McGregor, 2010). Apparently, some of the children with SLI are more successful at word learning than others. According to Gathercole and Baddeley (1993), a phonological memory deficit underpins poor word-learning in children with SLI. Syntactic bootstrapping limitations are seen as another possible cause for poor vocabulary learning, especially for verbs (for an overview, see McGregor, 2009).

Some recent studies have provided evidence for a strong relationship between lexicon and grammar in school-age children with SLI, other clinical groups, and TD children. For instance, Marchman, Saccuman, and Wulfeck (2004) found that receptive vocabulary predicted performance on a past tense elicitation task in 8-year-old children with SLI, children with focal cerebral lesions, and TD control groups. A study by McGregor et al. (2012) compared syntactic abilities and vocabulary abilities in 10-year-old children with SLI, language-matched (LA) and age-matched (CA) control groups, and 11-year-old children with autism (ASD). The children without syntactic deficits also demonstrated age-appropriate lexical knowledge, whereas the children with SLI and the ASD group with concomitant syntactic deficits both had poor lexical knowledge. Positive correlations were found between expressive syntax and vocabulary depth, as measured with word definition and word association tasks. It was concluded that the relationship between lexicon and syntax that characterizes early typical language development is still present in later and atypical language development (McGregor et al., 2012).

A theoretical account for SLI that links lexical knowledge to grammatical competence is the neurobiologically framed procedural deficit hypothesis (PDH), proposed by Ullman and Pierpont (2005). The PDH is based on the declarative/procedural (DP) model of lexicon and grammar (Ullman, 2001, 2004). The DP model posits that the mental lexicon and mental grammar are both subserved by two at least partially distinct memory systems in the brain. The declarative memory system underlies explicit episodic and semantic (lexical) knowledge, including word forms and meanings. The procedural memory system supports the implicit acquisition, storage and use of knowledge. It is supposed to subserve the learning and use of various perceptual, motor and cognitive skills, such as sequencing and probabilistic categorization, and rule-governed aspects of grammar (Lum, Conti-Ramsden, Page, & Ullman, 2012).

According to the PDH (Ullman & Pierpont, 2005), abnormalities of brain structures underlying procedural memory can largely explain the grammar problems in children with SLI. Children with SLI are claimed to have an impaired procedural memory system, leading to deficits in implicit sequence learning and grammar. Lexical retrieval and working memory, which partly rely on the same affected brain structures, are also impaired. On the other hand, the declarative memory system, which supports the acquisition of vocabulary, is supposed to remain largely intact. The deficit in the procedural memory system will be compensated for by other brain structures, and according to the PDH, certain grammatical functions will be taken over by the declarative memory system. Grammatical forms (e.g., verb inflection paradigms) can be learned as chunks, and the declarative memory system can even be used to learn grammatical rules explicitly (e.g., for past tense add *–ed* to the verb stem, for ongoing actions add *–ing* to the verb stem). Ullman and Pierpont (2005) also suggest that the heterogeneity in SLI populations can be explained in part by the severity of the procedural deficit on the one hand and by individual differences in strength of the declarative memory system on the other hand. Given that the lexicon is supported by declarative memory, a well-developed vocabulary would suggest a high capacity for compensation of a procedural deficit by the declarative system.

The Present Study

The present study aims to investigate the relation between vocabulary knowledge and grammatical development in Dutch older school-age children with SLI. Based on the literature, we assume that even in older school-age children with

SLI, vocabulary skills may be related to their grammatical development. Two research questions will be addressed in this study: (1) Which patterns can be identified in the development of grammatical complexity and correctness in older school-age children with SLI? (2) To what extent is vocabulary knowledge related to the development of grammatical complexity and correctness in older school-age children with SLI?

Research on SLI has been lopsided in two respects: most of it has focused on the early stages of development, and languages other than English are underrepresented in the research literature. The present paper contributes to redrawing this (dis)balance by focusing on later stages of development in children with SLI who are native speakers of Dutch. Research on SLI in Dutch children adds a new perspective to the existing research. For instance, the Dutch determiner system and inflectional paradigms for verbs and adjectives are richer than in English. Another difference is that Dutch is an SOV + verb second language. The finite verb takes second position in main clauses and sentence final position in dependent clauses. In the acquisition of finiteness and verb placement in Dutch, a developmental stage exists wherein TD children use dummy auxiliaries as placeholders for inflected main verbs in second position (Van Kampen, 1997; Wijnen & Verrips, 1998). In TD children, this behaviour gradually fades out, but Dutch children with SLI tend to overuse these dummy auxiliaries. An "avoidance strategy" to sidestep inflection and movement of the lexical verb was proposed as an explanation (de Jong, 1999; de Jong, Blom, & Orgassa, 2013; Orgassa, 2009). This overuse of dummy auxiliaries may be a hallmark of Dutch SLI. This feature is special because linguistic material is not omitted or substituted, as is often observed in language production of children with SLI. Rather, linguistic material is added to the utterance. An abundant use of dummy auxiliaries does not render a sentence ungrammatical, but can be regarded as an immature stage of grammatical development in older children. To date, it remains unknown up to what age Dutch children with SLI continue to overuse these dummy verbs.

In this study, grammatical development of school-age children with SLI and LA and CA control groups is investigated in a longitudinal design. First, the development of grammatical complexity and grammatical correctness is examined by using a narrative task. Subsequently, grammatical complexity and correctness measures are related to vocabulary development as measured with standardized vocabulary tests. Narratives are a complex genre that makes a great demand on the morphological/syntactic skills of children. It has been demonstrated that children showed clear morphological and syntactic difficulties in a narrative task and more difficulties than in conversation (Blankenstijn & Scheper, 2003). A narrative task prompts the children to express semantic relations that are difficult for them, and which they may avoid in spontaneous speech (Reilly, et al. 2004; Wetherell, Botting, & Conti-Ramsden, 2007). Using narratives to study grammatical development can therefore be considered as an ecologically valid way to study grammatical development in children with SLI.

Method

Participants

Participants were 30 monolingual Dutch children with SLI, 30 age-matched (CA) control children, and 30 language age-matched (LA) control children. At the three time points, the children with SLI had a mean age (*SD* in months) of 8;5 (1.7), 9;4 (1.6), and 10;4 (1.9) years respectively. The LA control children were aged 6;5 (2.4), 7;5 (2.4), and 8;5 (2.4) years in that order. The CA control children were aged 8;5 (3.5), 9;5 (3.5), and 10;5 (3.5) years respectively. Informed written consent was obtained from all parents. The data from the SLI group originated from a previous study by Van Weerdenburg et al. (2006). The children with SLI (21 boys, 9 girls) were all enrolled in special education for children with severe language impairments. The children in the LA control group (16 boys, 14 girls) and the CA control group (17 boys, 13 girls) were recruited from four different primary schools in the central part of the Netherlands.

As can be seen in Table 1, all children obtained non-verbal IQ scores within the normal range, measured with the Raven Colored Progressive Matrices (Raven CPM; van Bon, 1986). The children with SLI scored -1.1 *SD* below the age-normed mean on the sequential memory tasks from the Dutch version of the Kaufman Assessment Battery for Children (Kaufman ABC; Kaufman & Kaufman, 1983). In the first test session (T1), the language abilities of all participants were assessed with the Dutch standardized Language Proficiency Test All Children (LPT, *Taaltoets Alle Kinderen*, TAK; Verhoeven & Vermeer, 2001). The CA control group scored significantly higher on all LPT subtests compared to both the SLI and LA groups. The SLI group did not differ from the two years younger LA group on any of the LPT subtests. However, the SLI group did perform poorly on the Kaufman ABC sequential memory tasks compared to both control groups ($p < .001$).

Table 1. *Age, Raven Coloured Progressive Matrices Non-verbal IQ, Kaufman ABC Sequential Memory, and Scores on LPT Subtests in the three Groups at T1*

	Max. score	SLI (*n* = 30) *M (SD)*	LA (*n* = 30) *M (SD)*	CA (*n* = 30) *M (SD)*
Age in months		101 (1.7)	77 (2.4)	101 (3.5)
Raven CPM[a]		6.04 (1.98)	6.32 (1.79)	6.81 (1.75)
Kaufman ABC Seq. memory[b]		84.27 (9.38)	96.03 (11.24)	100.03 (10.67)
LPT Receptive vocabulary[c]	96	75.73 (9.27)	77.33 (6.85)	88.10 (4.44)
LPT Sentence comprehension I[c]	42	37.67 (2.25)	36.90 (2.17)	40.37 (1.81)
LPT Sentence comprehension II[c]	42	36.37 (3.07)	35.40 (2.51)	38.27 (1.91)
LPT Morphology[c]	24	17.43 (5.02)	16.57 (5.56)	22.80 (1.42)

Note. [a]standard score (range +/- 1 *SD* = 3.0 - 7.0), [b]quotient score, [c]raw score. LPT Sentence comprehension I measures function words, LPT Sentence comprehension II measures syntactic patterns, and LPT Morphology measures production of noun plurals and past participles.

Materials

The Storytelling tasks of the LPT (Figures 1 and 2) were used to elicit narratives. Two monochrome picture series show a sequence of events that form a coherent story each. The children are instructed to look at the pictures carefully and then tell the story in such a way that someone who cannot see the pictures will be able to understand the story in full. The investigator does not ask questions, but may encourage the children to continue if they stop midway. Both male and female characters, as well as plural referents appear in the LPT stories. This variety in characters increases the likelihood of observing a range of morphosyntactic errors in pronouns (e.g., case, gender, and number), determiners (e.g., gender) and subject-verb agreement. In the first story all characters are introduced in the first picture. The second story might be more taxing for the children, because some characters are introduced later on or are reintroduced. In this story, a shop attendant and a clown are acting in the background and mentioning them is not necessary for a complete and comprehensible narration. The narratives from the SLI group were recorded on audio cassette and later digitized. The stories told by the control groups were digitally recorded on a laptop.

The vocabulary skills of the SLI group were tested with three LPT vocabulary subtests, administered at ages 8, 9 and 10. The LPT receptive vocabulary test requires the child to choose one of four pictures that matches an orally presented target word. The LPT expressive vocabulary test uses pictures that have to be named after an orally presented short lead-in sentence. Both tests can be regarded as measuring vocabulary breadth. Vocabulary depth was measured with the LPT word definitions test. In this task the meaning of orally presented words has to be defined or explained.

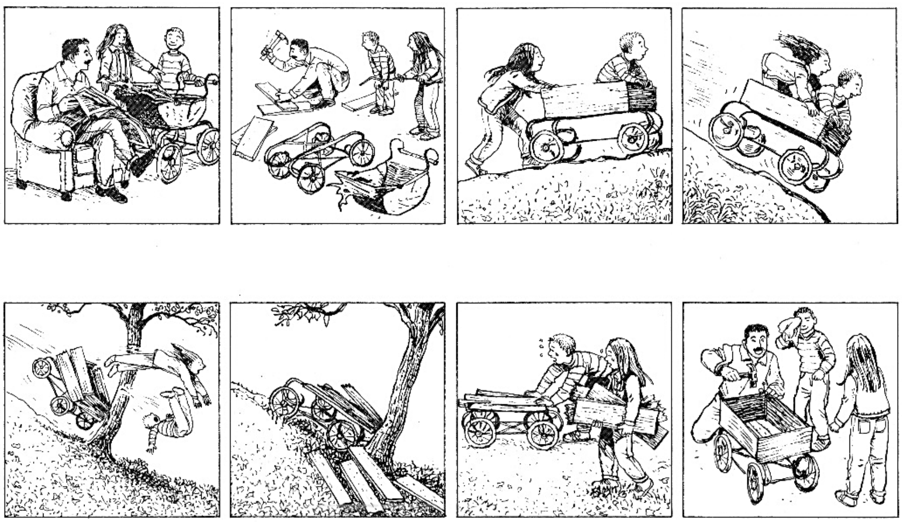

Figure 1. LPT Storytelling task 1. Copyright 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

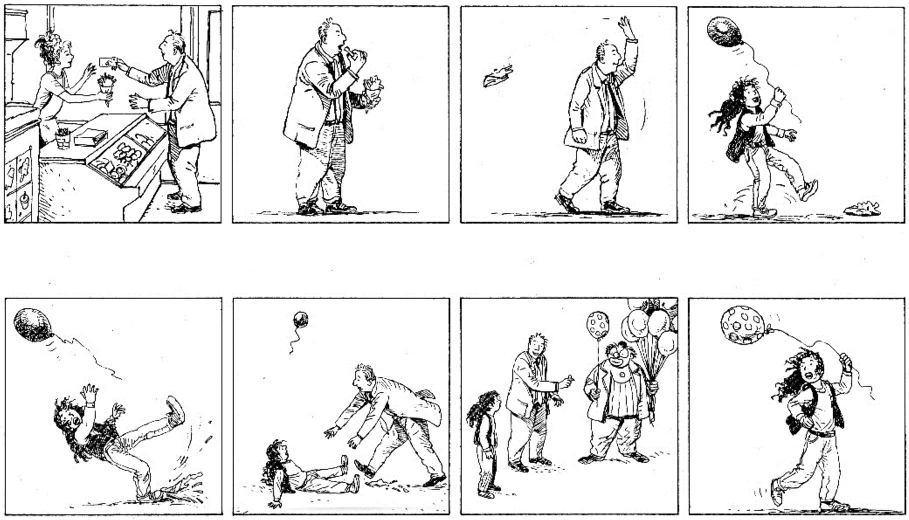

Figure 2. LPT Storytelling task 2. Copyright 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

Procedure

All children were tested at three points in time with an interval of 12 months during a period of two years. The narratives were orthographically transcribed using the CHAT format from the CHILDES system (MacWhinney, 2000). Each transcript contained Storytelling tasks 1 and 2 in this fixed order. The basic unit of analysis was the T-unit, defined as a single main clause plus any subordinate clause or non-clausal structure that is attached to it or embedded in it (Hunt, 1970). Coordinate clauses were transcribed and counted as separate utterances, except in cases where the coreferential subject of the second clause is omitted (e.g., "The children went down the hill and bumped into the tree"). Sentences with quoted speech, where the quote forms a full clause containing a subject and a verb, were also transcribed as separate utterances.

Additionally, grammatical complexity and grammatical errors were coded on separate dependent tiers in the CHAT files. The reliability of the transcriptions was checked by re-transcribing 10% of the files by either trained speech-language pathologists (for the transcripts of the children with SLI) or trained master students in linguistics (for the transcripts of the control groups). The point-to-point reliability at word level reached 95%.

There were no missing data in the three groups. All children told all stories at all three time points. Furthermore, there were also almost no unanalyzable utterances. When utterances were intelligible or otherwise unanalyzable (false starts or utterances that were broken off) these were excluded from the analysis.

Analysis

In order to answer the first research question, grammatical development was analyzed. A distinction was made between grammatical complexity and correctness. The operationalizations of the dependent measures are presented in Table 2. Because the narratives varied in length, percentages had to be computed for most variables, in order to make the measures comparable across children and time points.

The first step in the analysis concerned grammatical complexity. To this end, five different measures were computed. First, mean length of T-units (*MLTU*) in words was used as a general measure of grammatical complexity. Second, the *percentage of complex sentences* was computed. This measure contained all sentences with subordinate clauses (i.e., all forms of adverbial, nominal and relative clauses), coordinated sentences with conjunction reduction, direct speech and infinitival clauses. The third and fourth measures were *percentage of subordinate clauses* and *percentage of relative clauses*. Finally, the *percentage of complex-compound sentences* was computed.

The second step in the analysis of grammatical development was to examine grammatical correctness. The operationalizations of the grammatical correctness measures can be found in Table 2. The first measure was the *percentage of T-units correct* (i.e., error-free T-units) and was used as a general measure of grammatical correctness. However, one T-unit can contain several grammatical errors. Therefore, all grammatical errors were tallied and classified into two broad categories: *percentage of verb-related errors* and *percentage of non-verb related errors*. The measure *percentage of verb-related errors* contained all observed errors in verb morphology (e.g., agreement, tense, auxiliaries, and participles), verb placement and verb argument structure. Results on the separate verb-related error types are presented in the Appendix. The measure *percentage of non-verb related errors* contains all remaining grammatical errors. Most errors concerned determiners, pronouns, prepositions, and (pronominal) adverb '*er*' (there). Dutch children with SLI are known to omit this (pronominal) adverb in obligatory contexts (Bol & Kuiken, 1990). The measure *percentage of non-verb related errors* was not further subdivided, because many different error types were identified, and percentages of individual error types were too low for a meaningful quantitative analysis.

Finally, the measure *percentage of dummy auxiliaries* was included in the analysis of grammatical development. Overuse of these dummy verbs has been reported frequently in studies on Dutch SLI (de Jong, 1999; de Jong et al., 2013; Orgassa, 2009). Using a dummy auxiliary does not render a sentence ungrammatical, therefore counts of dummy verbs were not included in the grammatical correctness measures. The prolonged and frequent use of dummy auxiliaries can be argued to reflect an immature stage of grammatical development, which makes dummy auxiliaries a developmental feature worth investigating in Dutch children with SLI.

A remark has to be made on the status and interpretation of the grammatical complexity and correctness measures. Low percentages of various types of complex sentences do not necessarily mean that the children are not able to produce these constructions. Using complex sentences in a narrative is optional, and children may choose not to use complex syntax. On the other hand, the grammatical correctness measures do not have this inherent optionality, because correct verb inflection and determiner selection are imperative. In fact, we expect TD children to eventually reach a stage where they produce (almost) error-free sentences, approaching 100% grammatical correctness.

Table 2. *Operationalizations of the Grammatical Complexity and Correctness Measures*

Grammatical complexity:	
MLTU	Mean length of T-units in words (speech disfluencies, such as filled pauses, interjections, and repetitions are excluded from this count).
Complex sentences	Total number of complex sentences (= sum of all sentences containing subordinate, infinitival and reduced clauses, conjunction reduction, and direct speech) divided by the total number of T-units.
Subordinate clauses	Total number of subordinate clauses divided by the total number of T-units.
Relative clauses	Total number of relative clauses divided by the total number of T-units.
Complex-compound clauses	Total number of sentences containing two or more independent clauses combined with one or more dependent clauses, divided by the total number of T-units
Grammatical correctness:	
T-units correct	Number of error-free T-units divided by the total number of T-units.

Verb-related errors	Sum of all errors related to verbs (defined below) divided by clauses containing a(t least one) verb: - errors in subject-verb agreement (*hij loop*: he walk). - errors in verb-second placement (*dan de man komt:* then the man comes). - tense errors (present or past tense adverb with an incorrectly tensed verb: *toen*PAST *valt*PRESENT *het meisje*: then falls the girl). - root infinitives (*ballon geven:* balloon give). - past tense verb overregularisations (*hij brengde*: he bringed). - omissions and substitutions of aspect auxiliaries *zijn/hebben* (be/have) with a past participle (*toen Ø ze naar de clown gelopen*: then Ø they to the clown walked). - past participle errors: deletion of pre- and/or suffix or use of wrong suffix (*hij heeft het meisje (ge-)pak(-t) /gepakken*: he has the girl take). - verb argument structure errors: subject and object omissions were divided by the number of clauses where a subject or object was expected and obligatory (instances of allowed subject or object drop were not counted as errors).
Non-verb related errors	Sum of all remaining errors divided by clauses containing a(t least one) verb: all errors in word order, deletion of nouns, substitution and omission errors in determiners, prepositions, pronouns (case, gender and number), conjunctions, omission of *er* (there), and errors in adjectival inflection.
Dummy auxiliaries	Number of dummy auxiliaries divided by clauses containing a(t least one) verb.

To answer the first research question concerning grammatical development, mixed model analyses of variance (ANOVA) were used to examine differences between groups and at different time points, with Time (T1, T2, T3) as within-subjects factor and Group (SLI, LA, CA) as between-subjects factor. Significance level was set at 0.05. In order to analyze differences between groups, subsequent one-way ANOVA with post-hoc Bonferroni correction was used. GLM repeated measures ANOVA was used to test differences across time within the separate groups. The assumption of sphericity was checked for all variables with Mauchly's test of sphericity. Whenever this assumption was violated, the Greenhouse-Geisser corrected values are reported.

To answer the second research question, the first step was to analyze the development of vocabulary in the SLI group across time. GLM repeated measures ANOVA was used to test differences across time in the SLI group on the three LPT vocabulary tests. The second step was to examine the relation between vocabulary development and grammatical development with a Pearson two-tailed bivariate correlation analysis. The scores on the three LPT vocabulary tests at T1, T2, and T3 were correlated with the grammatical complexity and correctness measures, and with the use of dummy auxiliaries at the three time points.

Results

Grammatical Complexity

This first section of the results concerns grammatical complexity in the narratives of the three groups at the three time points. The descriptives of the grammatical complexity measures at the three measurement points are presented in Table 3.

Table 3. *Age, MLTU and Grammatical Complexity Measures (Percentages of Total Number of T-units in the Narratives) for the three Groups at the three Time Points*

	SLI (n = 30) *M (SD)*			LA (n = 30) *M (SD)*			CA (n = 30) *M (SD)*	
T1	T2	T3	T1	T2	T3	T1	T2	T3
Age (years; *SD* in months)								
8;5	9;4	10;4	6;5	7;5	8;5	8;5	9;5	10;5
(1.7)	(1.6)	(1.9)	(2.4)	(2.4)	(2.4)	(3.5	(3.5)	(3.5)
Mean length of T-units (in words)								
6.34	6.66	6.79	6.35	6.59	6.84	6.72	7.41	7.46
(0.78)	(0.94)	(0.79)	(0.78)	(0.86)	(0.74)	(0.98)	(1.14)	(0.84)
Complex sentences								
8.6	16.2	13.8	6.0	9.3	11.8	12.6	19.3	21.7
(6.6)	(13.2)	(8.0)	(6.8)	(9.3)	(9.7)	(10.4)	(13.9)	(10.9)
Subordinate clauses								
2.9	5.5	7.2	3.2	4.3	4.3	3.7	7.1	9.5
(3.1)	(5.4)	(5.0)	(4.2)	(5.5)	(5.4)	(4.9)	(7.3)	(5.1)
Relative clauses								
0.5	1.4	1.2	1.0	0.8	1.1	1.7	2.4	2.1
(1.4)	(2.0)	(2.1)	(2.6)	(1.7)	(2.0)	(3.1)	(3.2)	(2.9)
Complex-compound clauses								
0.1	1.1	0.4	0.1	0.2	0.8	1.0	1.4	2.2
(0.7)	(2.0)	(1.3)	(0.6)	(1.0)	(1.9)	(2.3)	(2.3)	(2.7)

MLTU

For *MLTU*, no significant interaction was found between Time and Group ($F_{(4,174)}$ = .949, p = .437, partial η^2 = .021). However, there was a significant main effect of Group ($F_{(2,87)}$ = 7.93, p = .001, partial η^2 = .154), as well as of Time ($F_{(2,174)}$ = 15.83, p < .001, partial η^2 = .154). The trajectories for *MLTU* and group differences are presented in Figure 3(a). This schematic outline, with an arbitrary scale on the y-axis, sketches the developmental trajectories of the three groups. Group differences are indicated by brackets and asterisks. One-way ANOVAs yielded significant differences between groups at T2 ($F_{(2,87)}$ = 6.38, p = .003) and T3 ($F_{(2,87)}$ = 6.51, p = .002). Post-hoc tests revealed that the SLI group had a significantly lower *MLTU* than the CA controls at T2 (p = .013) and at T3 (p = .005). Between the SLI group and the LA controls no significant differences were found at any of the time points. The LA group had a lower *MLTU* than the CA controls at T2 (p = .005) and at T3 (p = .011). Repeated measures (RM) ANOVA for the separate groups revealed that in the SLI group *MLTU* increased significantly between T1-T2 ($F_{(1,29)}$ = 4.31, p = .047 partial η^2 = .129), and between T1-T3 ($F_{(1,29)}$ = 8.85, p = .006 partial η^2 = .234). The LA control group only showed a significant increase in *MLTU* between T1-T3, $F_{(1,29)}$ = 9.13, p = .005, partial η^2 = .239). In the CA group, *MLTU* increased significantly between T1-T2 ($F_{(1,29)}$ = 14.47, p = 001, partial η^2 = .333), and between T1-T3 ($F_{(1,29)}$ = 16.24, p = 002, partial η^2 = .275).

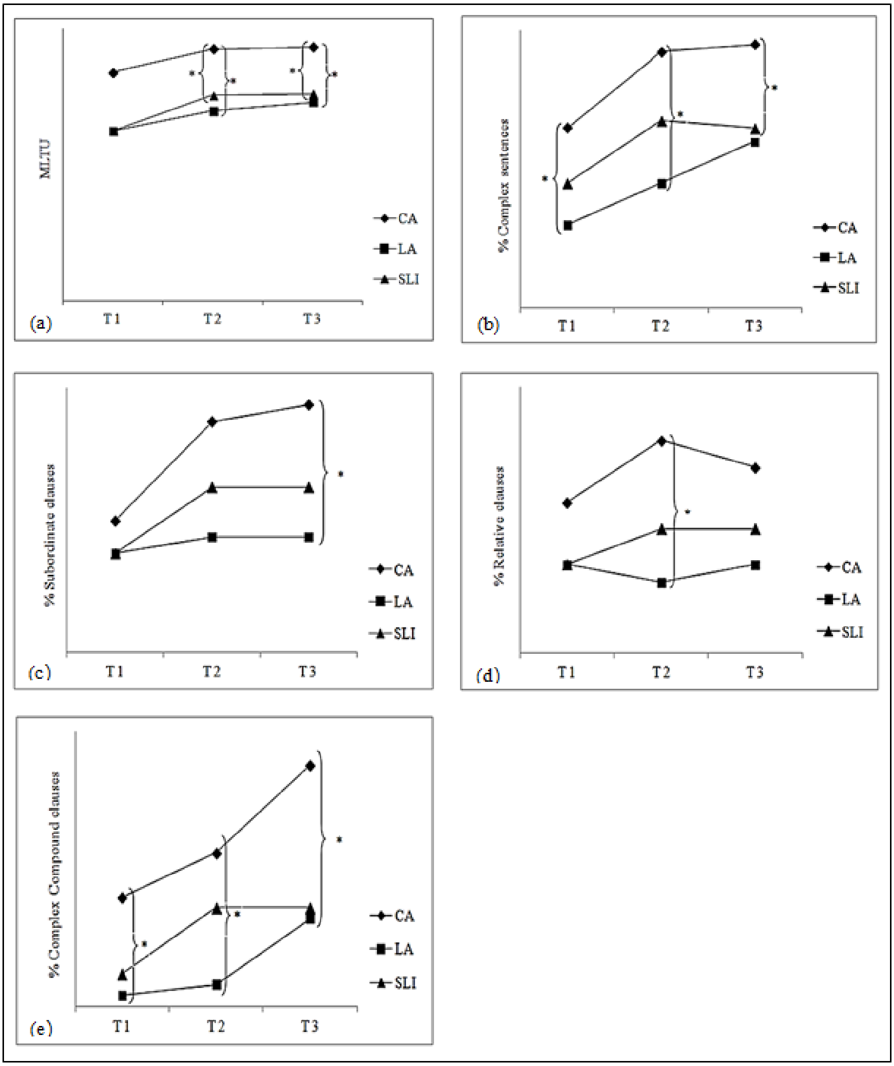

Figure 3. Development of markers of grammatical complexity at T1, T2, and T3 in the three groups: (a) MLTU, (b) Complex sentences, (c) Subordinate clauses, (d) Relative clauses, and (e) Complex-compound clauses (schematic outline, * = significant difference between groups).

Complex sentences

The measure *percentage of complex sentences* represents the number of complex sentences divided by the number of T-units in the narratives. There was no significant interaction between Time and Group ($F_{(4,174)}$ = 1.05, p = .382, partial

$\eta^2 = .024$). However, there was a significant main effect of Time ($F_{(2,174)} = 14.85$, $p < .001$, partial $\eta^2 = .146$) and of Group ($F_{(2,87)} = 11.90$, $p < .001$, partial $\eta^2 = .215$).

The growth trajectories of the percentages of complex sentences can be seen in Figure 3(b). One-way ANOVAs revealed significant differences between groups at T1 ($F_{(2,87)} = 5.07$, $p = .008$), T2 ($F_{(2,87)} = 5.17$, $p = .008$), and T3 ($F_{(2,87)} = 8.85$, $p < .001$). Post-hoc testing showed that the CA controls used more complex sentences than the LA controls at T1 ($p = .007$), T2 ($p = .007$), and T3 ($p < .001$). The children with SLI did not differ from the LA controls at any time point. The SLI group had fewer complex sentences than the CA group at T3 ($p = .006$). RM ANOVAs for the separate groups revealed that in the SLI group, the use of complex sentences increased between T1-T2 ($F_{(1,29)} = 8.14$, $p = .008$, partial $\eta^2 = .219$), and between T1-T3 ($F_{(1,29)} = 8.72$, $p = .006$, partial $\eta^2 = .231$). In the LA control group, this increase was significant between T1-T3 ($F_{(1,29)} = 8.49$, $p = .007$, partial $\eta^2 = .227$). In the CA group, the use of complex sentences increased between T1-T2 ($F_{(1,29)} = 5.73$, $p = .023$, partial $\eta^2 = .165$), and between T1-T3 ($F_{(1,29)} = 14.11$, $p = .001$, partial $\eta^2 = .327$).

Subordinate clauses

The *percentage of subordinate clauses* was computed by dividing the total number of subordinate clauses by the number of T-units. No interaction between Time and Group was found ($F_{(4,174)} = 2.08$, $p = .086$, partial $\eta^2 = .046$). There was a significant main effect of Time ($F_{(2,174)} = 15.41$, $p < .001$, partial $\eta^2 = .150$) and of Group ($F_{(2,87)} = 4.55$, $p = .013$, partial $\eta^2 = .095$). Figure 3(c) illustrates the changes over time for *percentage of subordinate clauses* in the three groups. One-way ANOVAs revealed that only at T3 significant differences between groups were found ($F_{(2,87)} = 7.60$, $p = .001$). Post-hoc testing showed that at T3, the CA controls used more subordinate clauses than the LA controls ($p = .001$). RM ANOVAs for the separate groups showed that use of subordinate clauses increased in the SLI group between T1-T2 ($F_{(1,29)} = 5.37$, $p = .028$, partial $\eta^2 = .156$), and between T1-T3 ($F_{(1,29)} = 18.27$, $p < .001$, partial $\eta^2 = .387$). In the LA control group no significant differences between time points were found. In the CA control group, the use of subordinate clauses increased between T1-T2 ($F_{(1,29)} = 5.30$, $p = .029$, partial $\eta^2 = .155$), and T1-T3 ($F_{(1,29)} = 20.50$, $p < .001$, partial $\eta^2 = .414$).

Relative clauses

The measure *percentage of relative clauses* was calculated by dividing counts of relative clauses by the number of T-units in the narratives. Figure 3(d) illustrates the changes over time with respect to percentages of relative clauses in the three groups. We found no significant interaction between Time and Group ($F_{(4,174)}$ = .58, p = .675, partial η^2 = .013). No significant main effect of Time was found ($F_{(2,174)}$ = 1.08, p = .341, partial η^2 = .012), but there was a significant main effect of Group ($F_{(2,87)}$ = 4.18, p = .019, partial η^2 = .088). One-way ANOVAs revealed that the difference between groups was only significant at T2 ($F_{(2,87)}$ = 3.36, p = .039). Post-hoc testing indicated that the LA group used fewer relative clauses than the CA group at T2 (p = .037). RM ANOVAs for the separate groups showed that use of relative clauses increased in the SLI group between T1-T2 ($F_{(1,29)}$ = 4.19, p = .05, partial η^2 = .126) and between T1-T3 ($F_{(1,29)}$ = 6.06, p = .02, partial η^2 = .173). In the LA and CA controls no significant differences between time points were found.

Complex-compound clauses

The measure *percentage of complex-compound clauses* was computed by dividing complex-compound clauses by the number of T-units. No significant interaction was found between Time and Group ($F_{(3.6,157.6)}$ = 1.55, p = .195, partial η^2 = .034). There was a significant main effect of Group ($F_{(2,87)}$ = 10.43, p < .001, partial η^2 = .193), as well as of Time ($F_{(1.8,157.6)}$ = 3.55, p = .035, partial η^2 = .039). In Figure 3(e) the developmental trajectories are illustrated. One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)}$ = 3.99, p = .022), T2 ($F_{(2,87)}$ = 3.31, p = .041), and T3 ($F_{(2,87)}$ = 6.11, p = .003). Post-hoc tests revealed that the SLI group used significantly fewer complex-compound clauses than CA controls at T3 (p = .004), although at T1 this difference almost reached significance (p = .051). The LA controls did not differ from the SLI group at all three time points. The LA group performed more poorly than the CA controls at T1 (p = .048), at T2 (p = .045), and at T3 (p = .030). RM ANOVAs for the separate groups revealed that the use complex-compound clauses increased in the SLI group between T1-T2 ($F_{(1,29)}$ = 6.50, p = .016 partial η^2 = .183). In the LA and CA group no significant changes across time were found.

Grammatical Correctness

The descriptives of the grammatical correctness measures and the use of dummy auxiliaries calculated from the narratives are presented in Table 4. The results on the separate measures will be reported separately in the following paragraphs.

Table 4. *Grammatical Correctness Measures and Dummy Auxiliaries (Percentages) at the three Time Points*

	SLI (n = 30) *M (SD)*			LA (n = 30) *M (SD)*			CA (n = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
T-units correct[a]									
	66.4	71.5	77.4	81.4	91.1	92.8	91.1	95.7	94.1
	(14.7)	(13.1)	(14.6)	(13.1)	(7.1)	(8.6)	(7.3)	(5.1)	(6.1)
Verb-related errors[b]									
	16.1	12.4	7.7	9.3	3.3	4.0	3.5	1.7	1.9
	(10.9)	(8.1)	(6.4)	(8.8)	(3.9)	(4.9)	(4.1)	(3.0)	(3.0)
Non-verb related errors[b]									
	22.0	18.8	19.1	10.0	5.2	3.3	6.0	2.4	3.5
	(14.9)	(13.7)	(13.2)	(8.6)	(6.0)	(5.7)	(5.5)	(2.8)	(4.5)
Dummy auxiliaries[b]									
	15.2	12.6	10.4	16.7	11.7	8.3	8.7	7.3	5.4
	(11.7)	(10.9)	(11.6)	(13.4)	(11.0)	(8.2)	(9.4)	(7.3)	(7.9)

Note. [a]Total number of correct T-units divided by the total number of T-units produced in the narratives. [b]Total number of verb-related errors, non-verb related errors, or dummy auxiliaries divided by the total number of clauses containing a(t least one) verb produced in the narratives.

T-units correct

The *percentage of T-units correct* was computed by dividing the number of error-free T-units by the sum of T-units. Figure 4(a) illustrates the differences between groups and changes over time. There was a significant interaction between Group and Time ($F_{(4,174)}$ = 3.44, p = .010, partial η^2 = .073). There also was a significant main effect of Time ($F_{(2,174)}$ = 27.10, p < .001, partial η^2 = .238) and of Group ($F_{(2,87)}$ = 56.77, p < .001, partial η^2 = .566). One-way ANOVAs showed that significant differences between groups were found at T1 ($F_{(2,87)}$ = 31.70, p < .001), T2 ($F_{(2,87)}$ = 60.22, p < .001), and T3 ($F_{(2,87)}$ = 24.05, p < .001). Post-hoc tests revealed that at all time points the SLI group had fewer correct T-units than the LA and CA control groups (p < .001 for all comparisons). The LA group had fewer correct T-units than the CA group at T1 (p = .008). At T2 and T3 this difference did not reach significance (p = .156 and p = 1.0 respectively). RM ANOVAs indicated that percentages of correct T-units increased in the SLI group between T1-T2 ($F_{(1,29)}$ = 5.71, p = .024, partial η^2 = .164), and between T2-

T3 ($F_{(1,29)}$ = 5.40, p = .027, partial η^2 = .157). In the LA controls, percentages of correct T-units increased significantly between T1-T2 ($F_{(1,29)}$ = 14.38, p = .001, partial η^2 = .332), and between T1-T3 ($F_{(1,29)}$ = 34.33, p < .001, partial η^2 = .542). In the CA group, the increase between T1-T2 was significant ($F_{(1,29)}$ = 11.08, p = .002, partial η^2 = .276). No significant differences were found between T2-T3 ($F_{(1,29)}$ = 1.93, p =.175, partial η^2 = .062), and between T1-T3 ($F_{(1,29)}$ = 4.08, p =.053, partial η^2 = .123). The CA groups seems to have reached ceiling from age 8 upwards.

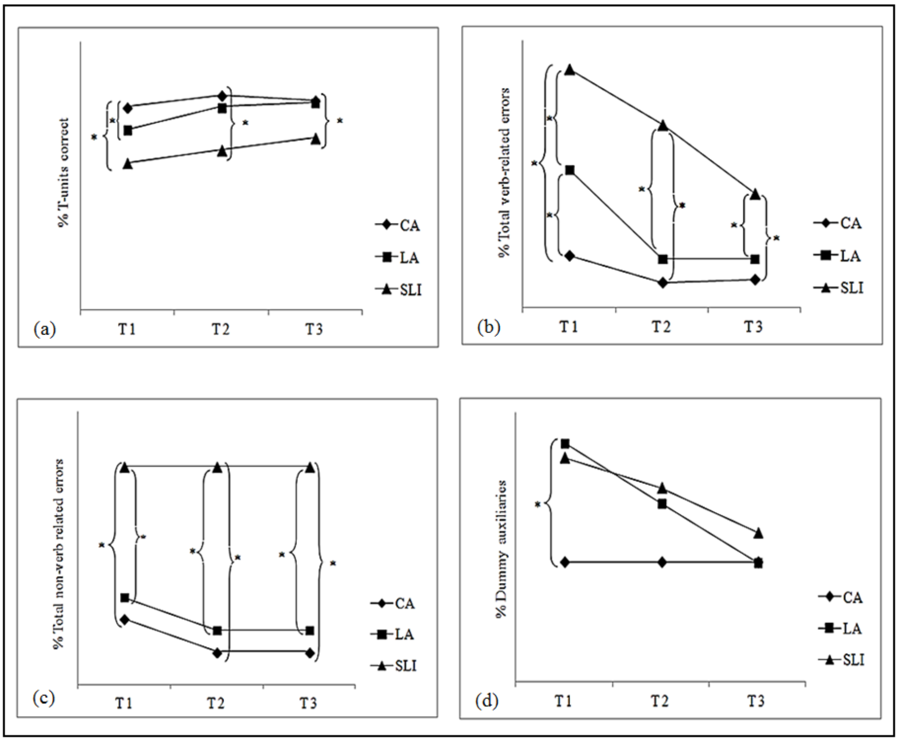

Figure 4. Development of markers of grammatical correctness at T1, T2, and T3 in the three groups: (a) Correct T-units, (b) Verb-related errors, (c) Non-verb related errors, and (d) Dummy auxiliaries (schematic outline, * = significant difference between groups).

Verb-related errors

The measure *percentage of verb-related errors* consisted of all verb-related errors divided by the total number of T-units. Figure 4(b) illustrates the differences between groups and changes over time. There was a significant interaction between Group and Time ($F_{(3.46,150.56)} = 4.48$, $p = .002$, partial $\eta^2 = .093$). There also was a significant main effect of Group ($F_{(2,87)} = 31.98$, $p < .001$, partial $\eta^2 = .424$) as well as of Time ($F_{(1.73,150.56)} = 22.32$, $p < .001$, partial $\eta^2 = .204$). One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)} = 16.95$, $p < .001$), T2 ($F_{(2,87)} = 33.34$, $p < .001$) and T3 ($F_{(2,87)} = 10.26$, $p < .001$). Post-hoc tests revealed that at all time points, the SLI group made more verb-related errors than the CA group ($p < .001$ for all comparisons) and the LA group ($p = .007$, $p < .001$, $p = .016$ respectively). At T1, the difference between LA and CA controls was significant ($p = .027$). At T2 and T3, the difference between the TD groups did not reach significance ($p = .801$, $p = .333$ respectively). RM ANOVAs for the separate groups revealed that verb-related errors decreased across time in the SLI group between T1-T2 ($F_{(1,29)} = 4.93$, $p = .034$, partial $\eta^2 = .145$), and between T2-T3 ($F_{(1,29)} = 10.59$, $p = .003$, partial $\eta^2 = .268$). In the LA controls, verb-related errors decreased between T1-T2 ($F_{(1,29)} = 11.30$, $p = .002$, partial $\eta^2 = .280$), and between T1-T3 ($F_{(1,29)} = 11.35$, $p = .002$, partial $\eta^2 = .281$). In the CA group, verb-related errors decreased between T1-T2 ($F_{(1,29)} = 5.00$, $p = .033$, partial $\eta^2 = .147$), but not between T2-T3 ($F_{(1,29)} = 0.08$, $p =.773$, partial $\eta^2 = .003$). The difference between T1-T3 was also not significant ($F_{(1,29)} = 3.55$, $p =.070$, partial $\eta^2 = .109$).

Non-verb related errors

The measure *percentage of non-verb related errors* consisted of all non-verb related errors divided by the total number of T-units. There was no significant interaction between Group and Time ($F_{(3.56,154.86)} =.974$, $p = .417$, partial $\eta^2 = .022$). However, there was a significant main effect of Group ($F_{(2,87)} = 39.17$, $p < .001$, partial $\eta^2 = .474$) as well as of Time, $F_{(1.78,154.86)} = 10.44$, $p < .001$, partial $\eta^2 = .107$). Developmental trajectories are presented in Figure 4(c). One-way ANOVAs revealed significant differences between groups at T1 ($F_{(2,87)} = 19.10$, $p < .001$), T2 ($F_{(2,87)} = 30.12$, $p < .001$) and T3 ($F_{(2,87)} = 32.40$, $p < .001$). Post-hoc tests revealed that at all time points, the SLI group made more non-verb related errors than both LA and CA control groups ($p < .001$ for all comparisons). There were no significant differences between the LA and CA control groups at any of the time points ($p = .408$, $p = .668$, and $p = 1$ respectively). RM ANOVAs for the separate groups revealed that *percentages of non-verb related errors* did not decrease significantly in the SLI group between any of the time points T1-T2, T2-T3, and T1-T3. The LA and CA controls showed a different developmental pattern. In the LA group, *percentages of non-verb related errors* decreased significantly between T1-T2 ($F_{(1,29)} = 9.76$, $p = .004$, partial $\eta^2 =$

.252) and between T1-T3 ($F_{(1,29)}$ = 17.27, p < .001, partial η^2 = .373). In the CA control group, the decrease was significant between T1-T2 ($F_{(1,29)}$ = 12.39, p = .001, partial η^2 = .252), and between T1-T3 ($F_{(1,29)}$ = 8.01, p = .008, partial η^2 = .216).

Dummy auxiliaries

The measure *percentage of dummy auxiliaries* was computed by dividing the number of dummy auxiliaries by the number of clauses containing a(t least one) verb. There was no significant interaction between Group and Time ($F_{(4,174)}$ = .79, p = .536, partial η^2 = .018). There was a significant main effect of Group ($F_{(2,87)}$ = 4.70, p = .012, partial η^2 = .097) as well as of Time ($F_{(2,174)}$ = 10.07, p < .001, partial η^2 = .104). The use of dummy auxiliaries in the groups is presented in Figure 4(c). One-way ANOVAs yielded a significant difference between groups at T1 ($F_{(2,87)}$ = 3.97, p = .022), but not at T2 and T3. Post-hoc tests revealed that at T1, the LA group used more dummy auxiliaries than the CA controls (p = .029). The difference between the SLI group and the control groups was not significant at any time point. RM ANOVAs for the separate groups revealed that in the SLI group the use of dummy auxiliaries decreased significantly between T1-T3 ($F_{(1,29)}$ = 4.94, p = . 034, partial η^2 = .145). In the LA control group the use of dummy verbs also decreased significantly between T1-T3 ($F_{(1,29)}$ = 10.33, p = .003, partial η^2 = .263). In the CA group, the use of dummy verbs did not change significantly across time.

Relations between Vocabulary and Grammatical Complexity and Correctness

To answer the second research question, first, the development of vocabulary knowledge in the SLI group, as measured with the three LPT vocabulary tests, was analyzed with RM ANOVAs. Means and standard deviations of the raw scores on the LPT vocabulary tests are presented in Table 5. For *receptive vocabulary*, there was a significant main effect of Time ($F_{(1.45,42.17)}$ = 89.56, p < .001, partial η^2 = .755). Scores increased significantly between T1-T2 ($F_{(1,29)}$ = 65.51, p < .001, partial η^2 = .693) and between T2-T3 ($F_{(1,29)}$ = 56.21, p < .001, partial η^2 = .660). For *expressive vocabulary*, there was a significant main effect of Time ($F_{(2,58)}$ = 28.02, p = .001, partial η^2 = .491). Scores increased significantly between T1-T2 ($F_{(1,29)}$ = 13.14, p = .001, partial η^2 = .312) and between T2-T3 ($F_{(1,29)}$ = 15.89, p < .001, partial η^2 = .354). For *word definitions*, there was a significant main effect of Time ($F_{(1.53,44.49)}$ = 32.92, p < .001, partial η^2 = .532). Scores increased significantly between T1-T2 ($F_{(1,29)} = 21.58$, $p < .001$, partial $\eta^2 = .427$) and between T2-T3 ($F_{(1,29)} = 23.37$, $p < .001$, partial $\eta^2 = .446$). In sum, the scores of the SLI group on all the three vocabulary tests increased significantly over time.

Table 5. *Raw scores of the SLI group on the LPT Vocabulary Tests at the three Time Points*

	Max. score	T1 *M (SD)*	T2 *M (SD)*	T3 *M (SD)*
LPT Receptive vocabulary	96	75.73 (9.27)	82.57 (7.27)	86.80 (6.06)
LPT Expressive vocabulary	60	40.40 (8.99)	45.03 (8.24)	48.67 (6.22)
LPT Word definitions	45	18.63 (8.23)	23.80 (6.74)	28.00 (6.26)

Subsequently, the scores on the LPT vocabulary tests at T1, T2, and T3 were correlated with the grammatical complexity and correctness measures at these time points. Table 6 shows the correlations between grammatical complexity and the vocabulary measures. Significant positive correlations were found between *MLTU* and *receptive vocabulary* as well as between MLTU and *expressive vocabulary* at the three time points. At T2, *receptive vocabulary*, *expressive vocabulary*, and *word definitions* correlated with the measure *percentage of complex-compound clauses.* No significant correlations were found between the three vocabulary tests and *percentages of complex sentences, subordinate clauses and relative clauses.*

Table 6. *Correlations of the LPT Vocabulary Tests with MLTU and the Grammatical Complexity Measures in the SLI group at the three Time Points*

	Receptive vocabulary			Expressive vocabulary			Word definitions		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
	Mean length of T-units								
T1	.36	.30	.27	.49**	.34	.30	.20	.25	.20
T2	.42*	.40*	.51**	.42*	.34	.34	.18	.23	.22
T3	.16	.33	.38*	.42*	.50**	.48**	.13	.27	.22
	Complex sentences								
T1	.03	.05	.10	-.09	-.09	-.25	.15	.20	-.03
T2	.36	.24	.26	.31	.27	.20	.04	.11	.21
T3	.17	.24	.29	.22	.20	.31	.09	.26	.18
	Subordinate clauses								
T1	.19	.11	.18	.16	.10	-.07	.18	.20	-.04
T2	.22	.15	.08	.11	.09	.06	.01	.03	.14
T3	.25	.32	.35	.22	.34	.29	.12	.31	.24
	Relative clauses								
T1	.03	.15	.25	.29	.14	.23	.16	.09	.01
T2	.12	.15	.05	-.18	-.01	-.06	-.16	.09	.03
T3	.22	.31	.29	.36	.31	.36	.26	.31	.28
	Complex-compound clauses								
T1	.07	.06	.19	.14	.18	.16	.24	.26	.12
T2	.48**	.41*	.31	.37*	.39*	.23	.31	.39*	.31
T3	.14	.15	.24	.22	.09	.14	.06	.22	.14

Note. Pearson correlation coefficients, two-tailed.* $p < .05$, ** $p < .01$.

The correlation analysis of the vocabulary tests with the grammatical correctness measures and the use of dummy auxiliaries showed more and stronger (negative) correlations, as indicated in Table 7. When scores on the vocabulary tests increase, the error percentages and the percentages of dummy verbs decrease. With respect to the different measurement points in time, dummy auxiliaries begin to correlate with the vocabulary tests at T1, non-verb-related errors at T2, and verb-related errors at T3.

Table 7. *Correlations of the LPT Vocabulary Tests with the Grammatical Correctness Measures and Dummy Auxiliaries in the SLI group at the three Time Points*

	Receptive vocabulary			Expressive vocabulary			Word definitions		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
				Verb-related errors					
T1	-.35	-.32	-.36	-.16	-.23	-.22	-.22	-.28	-.26
T2	-.30	-.31	-.17	-.13	.04	-.14	-.17	-.28	-.35
T3	-.37*	-.44*	-.35*	-.38*	-.34	-.36*	-.41*	-.44*	-.21
				Non-verb related errors					
T1	-.16	-.21	-.14	-.56**	-.28	-.39*	-.31	-.20	-.10
T2	-.50**	-.60**	-.50**	-.57**	-.53**	-.63**	-.37*	-.47**	-.38*
T3	-.48**	-.53**	-.45*	-.39*	-.48**	-.30	-.43*	-.48**	-.37*
				Dummy auxiliaries					
T1	-.44*	-.57**	-.61**	-.35	-.55**	-.55**	-.18	-.32	-.48**
T2	-.33	-.43*	-.50**	-.30	-.48**	-.52**	-.22	-.29	-.40*
T3	-.46*	-.58**	-.61**	-.31	-.64**	-.57**	-.45*	-.42*	-.46*

Note. Pearson correlation coefficients, two-tailed.* $p < .05$, ** $p < .01$

Discussion

In this study, grammatical development and vocabulary knowledge were investigated longitudinally in Dutch school-age children with SLI between the ages 8 and 10 years. Grammatical development was assessed in narratives using a wide range of grammatical complexity and correctness measures. Subsequently, correlation analysis was used to examine associations between vocabulary knowledge and grammatical complexity and correctness. The first research question concerned the grammatical development in older school-age children with SLI. Two aspects of this development were investigated, i.e., grammatical complexity and correctness.

Regarding the grammatical complexity measures, results showed that the SLI group performed similar to the LA group on MLTU at all three time points. At age 10, the children with SLI produced fewer complex sentences and complex-compound sentences than their CA peers. At all three time points, the SLI group did not differ from the LA and CA groups on subordinate and relative clause use. These results partly confirm the findings from other studies. Norbury and Bishop (2003) and Reilly et al. (2004) also reported fewer complex sentences in school-age children with SLI. Furthermore, a limited use of complex-compound clauses by children with SLI was also found by Marinellie (2004). However, several studies demonstrated poor performance on relative clauses in children with SLI compared to control groups (Contemori & Garraffa, 2010; Hakansson & Hansson, 2000; Hesketh, 2006; Koch et al., 2013; Novogrodsky & Friedmann, 2006; Schuele and colleagues, 2000; 2001; 2005; Stavrakaki, 2001), which contrasts with our findings. The reasons for not finding such group differences here may be related to differences in the age of the participants and to differences in methodology. For instance, most of the aforementioned studies investigated participants that were considerably younger (Contemori & Garraffa, 2010; Hakansson & Hansson, 2000; Koch et al., 2013). Stavrakaki (2001) only looked at relative clause comprehension, whereas most studies investigated relative clause production, often with elicited imitation or sentence completion tasks (Hesketh, 2006; Koch et al., 2013; Novogrodsky & Friedmann, 2006). In a narrative task, as used in our study, producing complex sentences, such as ones with relative clauses is optional. Nevertheless, the children with SLI produced relative clauses and other complex syntax from age 8 years onward and continuous progress on the grammatical complexity measures was also observed. As their linguistic knowledge of complex syntax appears to be intact, processing limitations best explain the fact that the 10-year-old children with SLI produced complex syntax to a lesser extent than their age peers, as corroborated by some recent studies (Coco, Garraffa, & Branigan, 2012; Hestvik, Schwartz, & Tornyova, 2010).

The analysis of grammatical correctness showed that the children with SLI performed poorly on all correctness measures, and were outperformed by the LA and CA groups at all three time points. Although the SLI group produced more correct T-units and verb-related errors decreased over time, non-verb related errors did not. In contrast to Moyle et al. (2011), who found that verb morphology alone was not a useful clinical marker of SLI in school-age children, in our study verb-related errors continued to distinguish between children with SLI and TD peers. For non-verb related errors, group differences were even larger and different developmental patterns were found, with decreasing error percentages in the LA and CA groups and stagnation in the SLI group.

Two possible explanations for this stagnation in non-verb related errors come to mind. The first explanation concerns the 'interface' characteristics of these non-verb related grammatical elements. The correct use of pronouns, determiners and (pronominal) adverb '*er*' (there) is not exclusively guided by grammatical rules, but also by discourse/pragmatics-related constraints. These functional elements may therefore be regarded as more complex than verb inflection paradigms, which are supposedly more straightforward and stable. The complexity of these non-verb related elements renders these functional elements more difficult to learn or to process, in both cases leading to more errors. The second explanation is related to the perceptual salience of the grammatical elements at stake. Dutch determiners, most pronouns, and (pronominal) adverb '*er*' (there) are monosyllabic words, mostly unstressed, and are often pronounced with unstressed, schwa-like vowels. According to the surface hypothesis, children with SLI have difficulties acquiring grammatical items with a low saliency (Leonard, 1989). Therefore, phonologically weak morphemes representing abstract grammatical features are harder to identify and consequently more difficult for children with SLI. These two explanations for the observed stagnation and high error rates of non-verb related grammatical elements are not mutually exclusive, and may even reinforce each other.

The examination of the use of dummy auxiliaries yielded some unexpected results. At T1, the CA group already demonstrated adequate performance on the LPT and the grammatical complexity and correctness measures. Because an overuse of dummy verbs was regarded as a sign of delayed grammatical development, we expected low percentages of dummy verbs in the CA group, higher percentages in the SLI group, and no differences between the SLI and LA groups. However, group differences did not reach significance, presumably because of the high variance in the groups. Still, the percentages of dummy verbs were almost twice as high in the SLI group compared to the CA group, and equaled those in the LA group. The SLI and LA groups also demonstrated similar trajectories for the use of dummy auxiliaries, both showing a decrease over time. Therefore, we tentatively conclude that the children with SLI are similar to the LA group with respect to dummy verbs.

Interestingly, the decrease in dummy verbs in the SLI group paralleled a decrease in verb agreement and verb placement errors. It has been proposed that dummy verbs are used to sidestep "movement" and inflection of the lexical verb (de Jong, 1999; de Jong et al., 2013; Orgassa, 2009). Obviously, the older school-age children with SLI in the present study gradually acquire the verb inflection paradigms, which would render the use of dummy auxiliaries superfluous, thus leading to a fade-out of dummy verbs. However, another explanation may be found in the word-finding difficulties often observed in many children with SLI (Dockrell & Messer, 2004; Kail & Leonard, 1986; Leonard & Deevy, 2004; McGregor, et al., 2013). Children with SLI may use a dummy verb as a stalling device, offering the children more time to retrieve a lexical verb. Using a dummy verb places the lexical verb in sentence final position, thus gaining extra time for its retrieval. From this point of view, the gradual fade-out of dummy verbs in the SLI group between the ages 8 and 10 could be associated with increasing efficiency of lexical retrieval. In fact, the progress on the LPT word definition test suggests an increasing strength of semantic representations in the lexicon. These stronger mental representations would then facilitate lexical retrieval (Kail & Leonard, 1986).

To conclude this section on grammatical development, a positive result is that grammatical skills continuously improved in the SLI group, notably with regard to verb-related errors and complex syntax. However, compared to their CA peers, the SLI group continued to lag behind in syntactic complexity and grammatical correctness. The picture that arises is one of a double deficit in grammatical development. The SLI group clearly demonstrates knowledge of complex syntax, but presumably as a result of processing limitations, the children with SLI use complex syntax to a smaller extent than their TD peers. In addition, stagnation was observed for non-verb related errors. The underlying problem appears to be a learnability issue. Verb inflection paradigms seem to be easier to learn than other functional categories, such as determiner-noun combinations, prepositions, pronouns, and (pronominal) adverb '*er*'. On the other hand, performance limitations were seen in the production of complex sentences.

Associations between vocabulary and grammar

Our second research question concerned the relation between vocabulary development and grammatical development. We investigated to what extent vocabulary knowledge could be associated with grammatical development in children with SLI. Correlations between vocabulary and grammatical complexity were limited to MLTU and to complex-compound clauses at age 9. However, vocabulary knowledge turned out to be strongly correlated with grammatical correctness and the use of dummy auxiliaries. Thus, when vocabulary scores increased, the rates of grammatical errors and dummy verbs decreased. These results are in line with Marchman et al. (2004), who found that production of correctly inflected past tense was predicted by vocabulary and with McGregor et al. (2012), who concluded that sparse lexicons were strongly associated with syntactic deficits. In the present study, vocabulary was associated with grammatical correctness and the use of dummy auxiliaries.

Although our study did not directly investigate procedural and declarative learning, as in Lum et al. (2012), our findings appear to lend some support to the DP model of lexicon and grammar (Ullman, 2001; 2004) and the PDH (Ullman

& Pierpont, 2005). The first finding concerns the correlation analyses that showed that vocabulary measures and complex sentences were mostly not correlated. According to the DP model, complex syntax is rule-governed and subserved by procedural memory. Therefore, it is to be expected that complex syntax is not strongly associated with lexical knowledge subserved by declarative memory and this is in accordance with our results. However, vocabulary was strongly associated with grammatical correctness and the use of dummy auxiliaries. At T3, vocabulary correlated with verb-related errors. Inspection of these errors (see Appendix) shows that more than half of these concerned verb argument structure which depends on memorized lexical knowledge (Ullman, 2004; Ullman & Pierpont, 2005). Therefore, errors in verb argument structure are likely to correlate with vocabulary. Furthermore, vocabulary and non-verb related errors were strongly correlated at T2 and T3. The majority of these non-verb related errors were found in determiners, pronouns, prepositions, and (pronominal) adverb '*er*'. Most of these functional elements can be regarded as lexicalized grammatical representations stored in declarative memory, and thus expected to correlate with lexical knowledge. Lum et al. (2012) found similar correlations between declarative memory and lexical abilities. Lastly, dummy auxiliaries correlated with vocabulary at all three time points. According to the PDH, declarative memory can compensate for grammatical deficits in SLI by explicit rule learning. Examples would be "use verb stem+*ing* for ongoing actions", and "use verb stem+*ed* for past tense". In a similar vein, Dutch children with SLI could use the rule "use dummy verb *gaat/gaan*+infinitive" for present tense and "use *ging/gingen*+infinitive" for past tense to compensate for difficulties with verb inflection. The combined gradual decrease of dummy verbs and errors in verb morphology observed in our study would suggest that grammatical rules are gradually and implicitly abstracted by the procedural memory system (Ullman, 2004).

A second finding from our study that agrees with the PDH, is the fact that the children with SLI performed poorly on the Kaufman sequential memory tasks. According to the DP model, implicit sequence learning and processing sequences are subserved by the procedural memory system, which is supposed to be impaired in children with SLI. These poor sequential memory skills may thus be indicative of a procedural memory deficit.

Interestingly, the PDH also posits that the heterogeneity in children with SLI can partly be explained by the severity of the procedural deficit and the relative strength of the declarative memory system. This compensatory role of declarative memory matches the associations found between lexical knowledge and grammatical development in our study. The children that demonstrate a well-developed vocabulary make fewer grammatical errors and are less dependent on dummy auxiliaries.

Although the present study tested associations between vocabulary knowledge and grammatical development, we did not investigate any causal relationships. Further research is needed to investigate the concrete relationship between vocabulary and grammar in children with SLI. Additional research could also help to reach a better understanding of the functions of dummy auxiliaries in Dutch SLI. Furthermore, intervention studies may be used to examine claims made by the PDH about procedural deficits and the compensatory role of declarative memory. Such studies could examine the learning mechanisms of children with SLI, for instance by comparing interventions based on explicit rule learning to interventions based on imitation of chunks and exemplars.

In conclusion, the grammatical skills of the children with SLI were found to improve over the years with regard to complex sentences and verb-related errors, although stagnation was observed for non-verb related errors. On the complexity measures, performance of the SLI group was intermediate between the LA and CA groups. The 10-year-old children with SLI lagged behind in complex syntax compared to the CA group. On the grammatical correctness measures, the SLI group was outperformed by both LA and CA groups at all three time points. Individual grammatical correctness scores were strongly associated with vocabulary knowledge, most markedly so for non-verb related errors and the use of dummy auxiliaries.

The picture that arises is one of a double deficit, with competence limitations especially for non-verb related errors and processing difficulties limiting the use of complex syntax. The results of the correlation analysis lends some support to the PDH (Ullman & Pierpont, 2006).

Clinical implications

For diagnostic purposes, it is important to know that MLTU and grammatical correctness measures computed from narratives continue to differentiate between older school-age children with SLI and typically developing control groups. In addition to standardized tests, the analysis of narrative tasks can provide valuable diagnostic information on grammatical complexity and grammatical correctness. As vocabulary knowledge turns out to be strongly associated with grammatical correctness, vocabulary expansion should receive intensive and continuous attention in therapy and education, for instance by deploying reading activities. Furthermore, intervention aimed at improving grammar in older school-age children with SLI should include remediation of difficulties with functional elements, such as determiners, pronouns and prepositions. Because for these grammatical elements stagnation was observed, specific intervention for these functional categories is advisable. Finally, we conclude that school-age children with SLI are still developing their language skills. No patterns of total stagnation were observed between the ages 8 and 10 years for grammatical complexity, correctness and vocabulary. Consequently, no substantial evidence was found for current policies that use concepts such as critical or sensitive periods to justify reductions in educational or therapeutic resources for older children with SLI.

Acknowledgments

This study was supported by the Royal Auris Group, Gouda, The Netherlands. We thank Cito for their permission to use the images of the LPT Storytelling tasks. We also thank the children, parents and staff of all the schools that participated; and Irene Sormani, Marij van Ewijk, Henaly Leijenhorst, Marjolein van der Horst, and Merel van Witteloostuijn for their help with the transcriptions.

References

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

Blankenstijn, C., & Scheper, A. (2003). *Language development in children with psychiatric impairment*. Utrecht: LOT Dissertation Series 82. (Doctoral dissertation, University of Amsterdam). Retrieved from http://www.lotpublications.nl/publish/articles/000542/bookpart.pdf.

Bol, G. W., & Kuiken, F. (1990). Grammatical analysis of developmental language disorders: A study of the morphosyntax of children with specific language disorders, with hearing impairment and with Down's syndrome. *Clinical Linguistics and Phonetics*, *4*, 77-86.

Coco, M. I., Garraffa, M., & Branigan, H. P. (2012). Subject Relative Production in SLI Children during Syntactic Priming and Sentence Repetition. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012), Sapporo, Japan, August, 2012* (pp. 228–233).

Colozzo, P., Gillam, R. B., Wood, M., Schnell, R. D., & Johnston, J. R. (2011). Content and form in the narratives of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 54*, 1609–1627.

Contemori, C., & Garraffa, M. (2010). Comparison of modalities in SLI syntax: A study on the comprehension and production of non-canonical sentences. *Lingua, 120*, 1940–55.

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42*, 1195–1204.

Curtiss, S., Katz, W., & Tallal, P. (1992). Delay versus deviance in the language acquisition of language-impaired children. *Journal of Speech and Hearing Research*, *35,* 373–383.

Dionne, G., Dale, P. S., Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development, 74*, 394–412.

Dockrell, J. E., & Messer, D. (2004). Lexical acquisition in the early school years. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 35–52). Philadelphia: John Benjamins.

Dollaghan, C. A. (1987). Fast mapping in normal and language-impaired children. *Journal of Speech and Hearing Disorders*, *52*, 218- 222.

De Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure*. Groningen: Grodil. (Doctoral dissertation, University of Groningen). Retrieved from: http://irs.ub.rug.nl/ppn/297976826.

De Jong, J. Blom, E. Orgassa, A. (2013). Dummy auxiliaries in children with SLI - a study on Dutch, in monolinguals and bilinguals. In E. Blom, I. van de Craats, & J. Verhagen (Eds), *Dummy auxiliaries in first and second language acquisition* (pp. 251-278). Berlin: De Gruyter Mouton.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working Memory and Language*. Hove: Lawrence Erlbaum Associates Ltd.

Gillam, R., & Johnston, J. (1992). Spoken and written language relationships in language/learning impaired and normally achieving school-age children. *Journal of Speech and Hearing Research, 35*, 1303–1315.

Håkansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language, 27*, 313-333.

Hesketh, A. (2006). The use of relative clauses by children with language impairment. *Clinical Linguistics and Phonetics*, *20*, 539–546.

Hestvik, A., Schwartz, R. G., & Tornyova, L. (2010). Relative clause gap-filling in children with specific language impairment. *Journal of Psycholinguistic Research*, *39*, 443–456.

Kail, R., & Leonard, L. (1986). Word-finding abilities in language impaired children. *ASHA Monographs*, no. 25.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children: Interpretative Manual.* Circle Pine, MN: American Guidance Service.

Keij, B., Cornips, L., Van Hout, R., Hulk, A., & Emmerik, J. van (2012). Knowing versus producing: The acquisition of grammatical gender and the definite determiner in Dutch by L1-TD, L1-SLI, and eL2 children. *Linguistic Approaches to Bilingualism*, *2*, 379–403.

Koch, C., Schuler, R., Friedmann, N., & Schulz, P. (2013, April). *Are relative clauses impaired in German-speaking children with SLI?* Poster session presented at the meeting of Linguistic Evidence, Berlin.

Leonard, L. B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10*, 179–202.

Leonard, L. B. (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.

Leonard, L. B., & Deevy, P. (2004). Lexical deficits in specific language impairment. In L. Verhoeven, & H. van Balkom (Eds), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 209-233). Mahwah, NJ: Erlbaum.

Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, *48*, 1138–1154.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk (3rd Ed.).* Mahwah, NJ: Erlbaum.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339–366.

Marchman, V. A., Saccuman, C., & Wulfeck, B. (2004). Productive use of the English past tense in children with focal brain injury and specific language impairment. *Brain and Language*, *88*, 202-214.

Marinellie, S. A. (2004). Complex syntax used by school-age children with specific language impairment (SLI) in child-adult conversation. *Journal of Communication Disorders*, *37*, 517–533.

McGregor, K. K. (2009). Semantics in child language disorders. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 365-387). New York, NY: Psychological Press.

McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders*, *42*, 35–47.

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language and Communication Disorders*, *48*, 307–319.

Moyle, M., Karasinski, C., Ellis Weismer, S., &. Gorman, B. K. (2011). Grammatical morphology in school-age children with and without language impairment: A discriminant function analysis. *Language, Speech, and Hearing Services in Schools*, *42*, 550–561.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.

Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *International Journal of Speech-Language Pathology, 8*, 364–375.

Norbury, C. F., & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders*, *38*, 287–313.

Oetting, J.B., & Hadley, P. A. (2009). Morphosyntax in child language disorders. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 341-364). New York, NY: Psychological Press.

Orgassa, A. (2009). *Specific language impairment in a bilingual context. The acquisition of Dutch by Turkish-Dutch learners*. Utrecht: LOT Dissertation Series 220. (Doctoral dissertation, University of Amsterdam). Retrieved from: http://www.lotpublications.nl/publish/articles/003604/bookpart.pdf.

Paradis, J. (2010). Comparing typically developing children and children with specific language impairment. In E. Blom and S. Unsworth (Eds.), *Experimental methods in language acquisition* (pp. 223-243). Amsterdam: John Benjamins.

Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language*, *88*, 229–247.

Rice, M.L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy and J. Schaeffer (Eds.), *Language competence across populations. Toward a definition of specific language impairment* (pp. 63-95). Mahwah: Lawrence Erlbaum Associates.

Schuele, M. C., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. *Clinical Linguistics and Phonetics*, *19*, 295–318.

Schuele, C.M., & Nicholls, L. (2000). Relative clauses: evidence of continued linguistic vulnerability in children with specific language impairment. *Clinical Linguistics and Phonetics*, *14*, 563–585.

Schuele, C. M., & Tolbert, L. (2001). Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics and Phonetics*, *15*, 257–274.

Schwartz, R.G. (2009). Specific language impairment. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 3-43). New York, NY: Psychological Press.

Sheng, L., & McGregor, K. (2010). Lexical-semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 146–159.

Stark, R.E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders, 46,* 114–180.

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language, 77*, 419–431.

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

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment. In H. Tager- Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 361–382). Cambridge, MA: MIT Press.

Tomblin, J., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research, 49*, 1193–1208.

Trauner, D., Wulfeck, B., Tallal, P., & Hesselink, J. (1995). Neurologic and MRI Profiles of language impaired children. Technical Report No. CND-9513, Center for Research in Language, University of California at San Diego.

Ullman, M. T. (2001). A neurocognitive perspective on language: the declarative/procedural model. *Nature Reviews Neuroscience*, *2*, 717–726.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92,* 231-270.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*, 399–433.

Van Bon, W. H. J. (1986). *Coloured Progressive Matrices: Manual of Dutch Norms.* London: Lewis.

Van Kampen, N. J. (1997). *First Steps in Wh-movement* (Doctoral dissertation, Utrecht University). Delft: Eburon.

Van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry*, *47*, 176–189.

Verhoeven, L. & Vermeer, A. (2001). *Taaltoets Alle Kinderen* [LPT, Language Proficiency Test for All Children]. Arnhem: Cito.

Wetherell, D., Botting, N., & Conti-Ramsden, G. (2007). Narrative in adolescent specific language impairment (SLI): A comparison with peers across two different narrative genres. *International Journal of Language and Communication Disorders, 42*, 583-605.

Wijnen, F., & Verrips, M. (1998). The acquisition of Dutch syntax. In S. Gillis & A. De Houwer (Eds), *The acquisition of Dutch* (pp. 223-299). Amsterdam: John Benjamins.

Appendix

Percentages of verb-related errors and the subtypes of verb errors (total counts of errors divided by the total counts of clauses with a(t least one verb)) in the three groups at the three time points.

	SLI *M (SD)*			LA *M (SD)*			CA *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Total verb errors	16.1 (10.9)	12.4 (8.1)	7.7 (6.4)	9.3 (8.8)	3.3 (3.9)	4.0 (4.9)	3.5 (4.1)	1.7 (3.0)	1.9 (3.0)
Agreement errors	4.4 (5.9)	4.0 (4.4)	1.5 (2.7)	2.9 (5.7)	1.2 (2.0)	0.8 (1.6)	0.4 (1.2)	0.5 (1.4)	0.6 (1.7)
Tense errors	4.6 (6.0)	2.1 (2.9)	1.3 (2.7)	1.1 (2.4)	0.8 (1.7)	1.0 (2.1)	0.8 (2.0)	0.1 (0.5)	0.3 (1.0)
Overregularisations	1.7 (2.7)	1.5 (2.9)	1.1 (2.2)	1.9 (4.3)	0.1 (0.6)	1.1 (2.9)	0.3 (1.8)	0.3 (1.1)	0.1 (0.4)
Verb-second errors	0.8 (1.6)	0.4 (1.2)	0.2 (0.8)	0.4 (1.7)	0.1 (0.6)	0.4 (1.1)	0.0 (0.0)	0.0 (0.0)	0.3 (1.4)
Root infinitives	0.1 (0.3)	0.1 (0.5)	0.1 (0.4)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Subject omissions	2.0 (4.6)	1.4 (2.1)	1.0 (1.6)	0.8 (2.0)	0.2 (0.9)	0.3 (1.3)	0.5 (1.4)	0.3 (1.1)	0.4 (1.7)
Object omissions	3.8 (5.8)	2.8 (4.5)	3.5 (5.2)	1.3 (4.3)	0.2 (1.1)	0.3 (1.5)	0.8 (2.6)	0.0 (0.0)	0.6 (2.4)

4 Speech disruptions in school-age children with SLI: a developmental perspective [3]

Abstract

Purpose: This study examined the development of speech disfluencies (i.e., stalls and revisions) longitudinally in children with specific language impairment (SLI) and typically developing peers.

Method: Participants were 30 monolingual Dutch children with SLI, 30 language-matched (LA) children, and 30 age-matched (CA) children. Speech disfluencies were analyzed for frequencies, types, silent pause duration, and syntactic distribution in narratives collected at three points in time (T1, T2, and T3) with 12 month intervals, during a two-year period. The SLI and CA groups were aged 8 years at T1, and the LA group was 2 years younger.

Results: The SLI group exhibited more stalls than the CA and LA groups, but the groups did not differ on revisions. Frequencies and types of disfluencies in the SLI group resembled those in the LA control group more than those of the CA control group. The distribution of disfluencies followed the same pattern in all three groups: highest frequencies at utterance-initial position, followed by clause-initial, phrase-initial, and word-initial position. However, the children with SLI produced more disfluencies than CA peers at word-initial positions, suggesting that difficulty with lexical retrieval may contribute to their disfluency. Over time, silent pause rates steadily decreased in the SLI group, but no changes were observed for other types of disfluencies.

Conclusions: The findings suggests that, although some improvement in fluency is seen, sentence formulation continues to be challenging for school-age children with SLI. The higher disfluency rates in the SLI group reflect their compromised expressive language skills.

Keywords: speech disfluency, specific language impairment (SLI), narratives.

---

[3] To be submitted (Authors: Rob Zwitserlood[a,c], Marjolijn van Weerdenburg[b], Frank Wijnen[a], and Ludo Verhoeven[b], [a]Utrecht Institute of Linguistics (OTS), University Utrecht, The Netherlands, [b]Behavioural Science Institute, Radboud University Nijmegen, The Netherlands, [c]Koninklijke Auris Groep, The Netherlands).

Speech disruptions are a normal phenomenon in spontaneous speech. A speaker may experience problems at all three stages of the language production process, namely conceptualization, sentence formulation, or articulation (Levelt, 1989). Such difficulties may be reflected in self-correcting, pausing and other disruptions of the steady flow of language production (Levelt, 1989, Rispoli, 2003). Extensive research on such phenomena in adult language production has indicated that pauses, repetitions or repairs reflect the strategies that speakers adopt to solve breakdowns in sentence planning and production (Clark 2002; Levelt, 1989; Postma & Kolk, 1993).

Although speech disfluencies have been studied extensively in adults, researchers have also become increasingly interested in speech disfluencies in typically developing (TD) young children as a source of information on the development of language production processes. More recently, the study of speech disruptions in clinical populations, such as children with specific language impairment (SLI), has gained interest in order to provide insight in the nature of language (production) difficulties, i.e., the level or type of processing that is impaired. Children with SLI exhibit syntactic deficits as well as lexical deficits, notably poor lexical retrieval. Furthermore, working memory limitations and processing problems have also been reported frequently in these children (for an overview, see Leonard, 1998; Schwartz, 2009). Because of these linguistic and cognitive weaknesses, it is to be expected that children with SLI are prone to breakdowns in speech production, leading to higher disfluency rates than observed in TD children. However, the evidence presented in the literature is not unambiguous.

Speech disfluencies: types and functions

Different terms have been coined for interruptions of fluent speech, such as 'hesitation phenomena', 'mazes', 'speech disfluencies', and 'speech disruptions'. In the present study, the terms 'speech disfluencies' and 'speech disruptions' will be used interchangeably. In the literature on stuttering, speech disfluencies are often divided into normal and stutter-like disfluencies (Yairi & Ambrose, 1992). Normal disfluencies are observed in all speakers and include silent and filled pauses, interjections, whole-word repetitions, phrase repetitions, and revisions. Stutter-like disfluencies include tensed pauses, blocks, part-word repetitions and prolongations.

Another way to categorize speech disruptions is based on their (apparent) function. Rispoli (2003) proposed a division of speech disruptions into stalls and revisions. Stalls are speech disruptions that interrupt sentence production but are not associated with a change of linguistic structure (e.g., pauses and repetitions). Rispoli suggests that the function of stalls is to buy time in the sentence formulation process. Stalls arise because of something that has not yet been articulated. They can result from difficulties in forming concepts, activating syntactic frames or retrieving syntactic and semantic information from the lexicon (Levelt, 1989; Postma & Kolk, 1993). Stalls can take the form of silent pauses (*no phonation*), filled pauses (*uh, um*), interjections (*well, you know*), part-word repetitions (*b-bi-bicycle)* and whole-word repetitions (*we go-we go to the garden*). In addition to gaining extra time for sentence formulation, stalls have been argued to serve communicative purposes (Clark, 2002). Producing filled pauses and interjections allows speakers to keep the floor while formulating their utterance. Word and phrase repetitions would indicate that the speaker attempts to keep the syntactic units intact, and thus tries to accommodate the listener.

The other class of speech disruptions are revisions, which also interrupt sentence production, but involve modifications of already produced speech. Revisions can be phonological, grammatical, or semantic repairs, which take place when an error has been produced. The error has to be detected by the speaker, who subsequently also chooses to correct the speech error. Examples of revisions are given in Example (1).

(1) Yesterday, the <frish> [*phonological*] fisherman <steers> [*grammatical*] steered his boat into the <whales> waves [*semantic*].

Revisions are often accompanied by stalls, as the speaker needs time to execute the repairs, and buys time by inserting silent pauses, filled pauses, and whole-word repetitions, as illustrated in Example (2).

(2) Yesterday, the <frish> [*silent pause*] fisherman <steers> [*uh*] steered his boat into the <whales> [*into the*] waves.

Speech disfluencies in TD children

Various studies on speech disfluencies in TD children have demonstrated a relationship between speech fluency and language development. In a cross-sectional study on speech fluency in narratives of seven age groups ranging from 5;10 to 18;1 years, the frequency and duration of silent pauses were found to decrease with age (Kowal, O'Connell, & Sabin, 1975). It was suggested that younger children need more time to plan an utterance compared to the older children. Wijnen (1990) analysed disfluencies in a TD Dutch child between the ages of 2;4 and 2;11 years. The frequency of disfluencies increased first and subsequently declined. In the first months of the observation period, disfluencies were also distributed randomly over sentences, but in the final months disfluencies were concentrated at phrase and sentence boundaries. Wijnen argued that the decline of speech disruptions was related to an increasing use of a few syntactic frames. The frequent use of these frames led to automaticity in sentence planning, which reduced the likelihood of disruptions. The distribution of speech disruptions before sentences and phrases, where sentence planning takes place, was regarded as reflecting the emergence of an adult-like sentence formulation system. Colburn and Mysak (1982a; 1982b) examined early language development and fluency longitudinally in four children between the ages of 2 and 3 years. They also reported that novel syntactic structures provoked more speech disruption than structures that had already been used for some time, even though the pattern of total frequency of disfluency varied widely among these four children by age or Mean Length of Utterance (MLU) level. They concluded that these new structures had not been fully practiced and placed a burden on sentence formulation abilities (Colburn & Mysak, 1982a, 1982b).

In a cross-sectional design, Rispoli and Hadley (2001) studied disfluency behaviour in children from 2;6 to 4;0 years. Longer and more complex sentences contained more disfluencies, but as syntactic competence increased over time, as indicated by Index of Productive Syntax scores, these complex sentences became more automated and less prone to trigger disfluencies. The idea that disfluencies are associated with sentence planning is supported by Rispoli (2003), who found that frequencies of speech disruptions were highest at clause onset. Over 70% of all disfluencies occurred within the first three words of a clause. Rispoli, Hadley, and Holt (2008) examined frequencies of stalls and revisions longitudinally in twenty children between 1;9 and 2;9 years. Results indicated that revision rate increased with age, but no such developmental trend was found for stalls. In contrast, stall rate was related to sentence length whereas revision rate was not. It was argued that the increase in revision rate reflected the growing ability of children to monitor and repair their language output.

To conclude, evidence from studies on speech fluency in TD children suggests that speech disfluency is related to age and language experience, syntactic complexity, and sentence length. Generally speaking, disfluency decreases with age in young children, and this has been associated with maturation and automatization of utterance production processes.

Speech disfluencies in children with SLI

The association between speech fluency and language development in TD children has prompted researchers to investigate this relationship in children with language impairments. Because of their compromised language functioning, these children were expected to demonstrate more disfluencies than TD children. In fact, several studies have reported a higher incidence of speech disfluencies in children with weak language skills. For instance, Nettelbladt and Hansson (1999) investigated speech disruptions in Swedish preschoolers with SLI and compared them to MLU-matched TD controls. These control children had a phonological impairment, but no grammatical impairment. The children with SLI produced more (part-word) repetitions and silent pauses. The repetitions affected both lexical words and function words in equal measure, whereas in the control children mainly function words were affected. Boscolo, Bernstein Ratner, and Rescorla (2002) reported that 9-year-old children with a history of expressive SLI (H-SLI) were significantly less fluent than their TD peers. The H-SLI group produced more stutter-like disfluencies, but no group differences were found for normal disfluencies, such as whole-word repetitions and revisions. A study by Guo, Tomblin, and Samelson (2008) investigated speech disruptions in narratives of 10-year-old children with SLI in a three-group design with age-matched (CA) and language-matched (LA) control groups. Results showed that the children with SLI produced more silent pauses compared to CA controls and also had more speech disruptions at phrase boundaries. The SLI group did not differ from the LA group. The authors suggested that speech disruptions were related to language ability. Higher disruption rates at phrase boundaries would reflect lexical and syntactic deficits in children with SLI. Further support for a relation between disfluencies and language impairment was presented by Finneran, Leonard, and Miller (2009). They examined disfluencies in 8-year-old children with SLI with a structural priming task eliciting single sentences. Such a highly controlled sentence production task was argued to have lower processing demands than a narrative task. However, the SLI group produced a significantly higher rate of speech disruptions as compared with TD control children, even when sentences were grammatically accurate.

Other researchers did not find a relationship between speech disfluencies and developmental language impairment (Lees, Anderson & Martin, 1999; MacLachlan & Chapman, 1988; Scott & Windsor, 2000). According to Guo et al. (2008), this inconsistency may have been caused by methodological differences. Different disfluency taxonomies were used in different studies. For instance, some studies did not take silent pauses into account, or only pauses longer than 2 seconds. Disfluency rates were computed by dividing speech disruptions either by number of utterances, or by number of intended words (i.e., total number of words without filled pauses, interjections, repetitions, and revisions). The studies that found an association between disfluencies and language impairment adjusted the frequency of disfluencies by the total number of intended words. According to Dollaghan and Campbell (1992), this correction is more sensitive than other methods to normalize disfluency rates in clinical populations. Furthermore, it should be noted that some studies only used CA control groups, and others used LA control groups, matched on either MLU, on vocabulary scores or on language composite measures. The numbers of participants in the studied groups were often small, and the ages and age ranges of the participants varied considerably. Moreover, most studies used spontaneous speech, whereas some used narratives or sentence production tasks. These different elicitation methods could lead to different results (Blankenstijn & Scheper, 2003). It was found that children with SLI produced more disfluencies in narratives as compared with spontaneous speech (MacLachlan & Chapman, 1988). In future research, more uniformity with respect to methodology, such as sample size, disfluency taxonomy, and elicitation method would certainly help to elucidate the relation between speech disfluency and impaired language abilities.

To date, longitudinal studies on the development of disfluency in children with SLI are almost non-existent. To the best of our knowledge, only Hall (1996) longitudinally studied speech disruptions in children with SLI. This longitudinal study followed up on Hall, Yamashita, and Aram (1993), who identified a subgroup of ten children from a sample of sixty preschoolers with SLI as a high disfluency group. The children in this subgroup were older than the other participants and had better developed lexical skills than morphosyntactic skills. It was suggested that this asynchronous development of lexical and syntactic abilities placed these children at risk for fluency breakdown. Hall (1996) re-examined nine children of this subgroup between the ages 7 and 9 years. Fluency improved with age and language development, although in some of the children stutter-like disfluencies increased over time.

In summary, it has been shown that in normal language development speech disfluency decreases as a function of age. A viable interpretation is that, as children become more experienced language users, sentence formulation becomes more automatized, leading to a decrease in speech disfluencies. However, in children with SLI the results are not so straightforward. Not all studies found higher disfluency rates in children with SLI as compared with age-matched TD peers. Some studies found qualitative differences, where children with SLI produced more stutter-like disfluencies than their peers. The higher disfluency rates in children with SLI have been related to their weaker language skills. However, whether the syntactic deficits and lexical deficits of children with SLI contribute equally to speech disfluency remains unclear. Perhaps, these contributions change over time, when grammatical abilities gradually improve, but lexical deficits, such as word-finding difficulties persist (Messer & Dockrell, 2006).

The present study

The present study aims to gain more insight into the relation between speech disfluency and impaired language development by examining the development of speech disruptions longitudinally in school-aged children with SLI. In part, the present study replicates Guo et al. (2008), in which frequency, types, and distribution of disfluencies across syntactic positions were examined in narratives, using a three group design (i.e., SLI, LA, and CA groups). The children with SLI and the CA controls in Guo et al. were 10-year-olds, and the LA controls were two years younger. The groups in our study were assessed with a narrative task at T1, T2, and T3, when the SLI and CA groups were aged 8, 9, and 10 years and the LA group was aged 6, 7, and 8 years, respectively. The results from our study at T3 can thus be compared to Guo et al. However, Guo et al.'s categorization of speech disruptions into silent and vocal hesitations was not followed. Instead, we adopted Rispoli's (2003) division into stalls and revisions, because this functionally based division allows for a better understanding of speech processing (difficulties).

In the present study the following research questions were addressed: (1) Do children with SLI produce more disfluencies than the TD control children? (2) Does fluency in the SLI group increase over time, as observed in TD children? (3) Can we relate disfluencies to specific components of the sentence formulation process? Given the fact that children with SLI are delayed in their syntactic and lexical development compared to CA peers, we expected to find more disfluencies in the SLI group than in the CA group. With respect to the first research question, we expected the SLI group to have more silent pauses and (stutter-like) part-word repetitions than the CA group. The children with SLI need more time for sentence planning and can have difficulties in accessing the phonological form of words (Guo et al., 2008). On the other hand, we expected the SLI and LA groups to perform similarly because the SLI group has a language level comparable to the LA group. With respect to the second research questions, fluency was expected to improve over time in all three groups, because of increasing language experience and automaticity. Finally, with respect to the third research question, children with SLI were expected to have more disfluencies than the control groups at utterance, clause and phrase onset, because they have difficulties with sentence formulation and, because sentence planning takes place at these locations (Guo et al, 2008; Rispoli, 2003; Wijnen, 1990). Furthermore, a higher disfluency rate at word initial (i.e., non-phrase or clause-initial) positions was also expected, because of the lexical retrieval problems often reported in children with SLI.

Method

Participants

Participants were 30 monolingual Dutch children with SLI, 30 LA control children, and 30 CA control children. At the three time points, the children with SLI had a mean age (standard deviations in months) of 8;5 (1.7), 9;4 (1.6), and 10;4 (1.9) years, respectively. The LA control children were aged 6;5 (2.4), 7;5 (2.4), and 8;5 (2.4) years, in that order. The CA control children were aged 8;5 (3.5), 9;5 (3.5), and 10;5 (3.5) years, respectively. Informed written consent was obtained from all parents. The data from the SLI group originated from a previous study by Van Weerdenburg, Verhoeven, and van Balkom (2006). The children with SLI (21 boys, 9 girls) were all enrolled in special education for children with language impairments and were diagnosed by a team of specialists (i.e., speech therapist, psychologist, educationalist, and ENT specialist) as having SLI at all three measurement points. The children in the LA group (16 boys, 14 girls) and the CA group (17 boys, 13 girls) were recruited from four different primary schools in the central part of the Netherlands. Table 1 presents the group means and standard deviations (*SD*) at first measurement for nonverbal IQ (Raven Coloured Progressive Matrices; van Bon, 1986), sequential memory (Dutch version of the Kaufman Assessment Battery for Children; K-ABC; Kaufman & Kaufman, 1983), and subtests of the Dutch Language Proficiency Test for All Children (LPT; Taaltoets Alle Kinderen, TAK; Verhoeven & Vermeer, 2001). All children obtained non-verbal IQ scores within the normal range. The children with SLI scored -1.1 *SD* below the standardized mean on the sequential memory tasks. The CA group scored significantly higher on all LPT subtests as compared with the SLI group and the LA group. The SLI group did not differ from the LA group on any of the LPT subtests. However, the SLI group did perform poorly on the Kaufman ABC sequential memory tasks compared to the LA and CA control groups ($p < .001$).

Table 1. *Background Measures of Participants at the first Time Point*

	Max. score	SLI (*n* = 30) *M (SD)*	LA (*n* = 30) *M (SD)*	CA (*n* = 30) *M (SD)*
Age in months		101 (1.7)	77 (2.4)	101 (3.5)
Raven CPM[a]		6.04 (1.98)	6.32 (1.79)	6.81 (1.75)
K-ABC Sequential memory[b]		84.27 (9.38)	96.03 (11.24)	100.03 (10.67)
LPT Receptive vocabulary[c]	96	75.73 (9.27)	77.33 (6.85)	88.10 (4.44)
LPT Sentence comprehension I[c]	42	37.67 (2.25)	36.90 (2.17)	40.37 (1.81)
LPT Sentence comprehension II[c]	42	36.37 (3.07)	35.40 (2.51)	38.27 (1.91)
LPT Morphology[c]	24	17.43 (5.02)	16.57 (5.56)	22.80 (1.42)

Note. [a]Standard score (range +/- 1 *SD* = 3.0 - 7.0); [b]Quotient score; [c]Raw score. LPT Sentence comprehension I measures function words, LPT Sentence comprehension II measures syntactic patterns, and LPT Morphology measures production of noun plurals and past participles.

Materials

The narrative tasks of the LPT are two story generation tasks based on sequences of eight monochrome pictures each (see Figures 1 and 2). The children were instructed to look at the pictures carefully and then tell the story in such a way that someone who cannot see the pictures, will be able to understand the story in full. The investigator did not intervene, but would encourage the children to continue if they stopped midway. In the first story all characters are introduced in the first picture. The second story might be more taxing as it forces the storyteller to introduce or reintroduce some characters at a later stage. The narratives from the SLI group were recorded on audio cassette and later digitized. The stories told by the control groups were digitally recorded on a laptop.

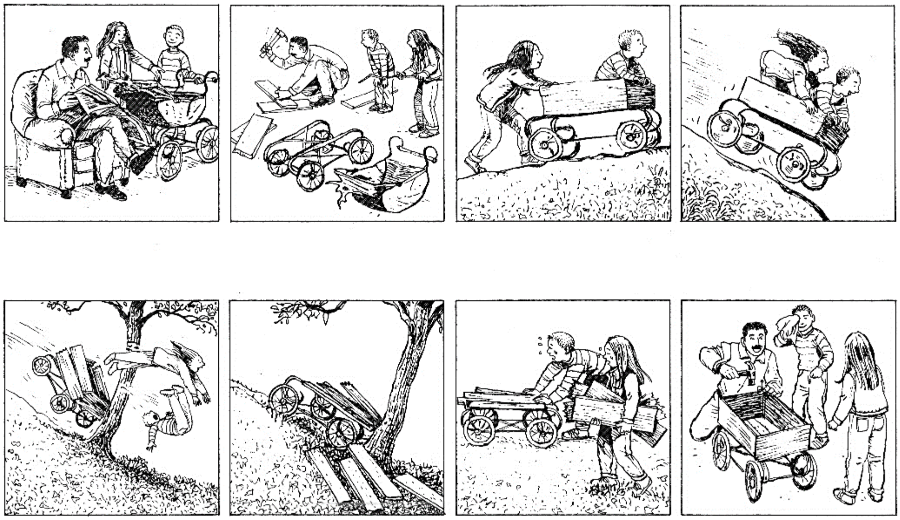

Figure 1. LPT Storytelling task 1. Copyright 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

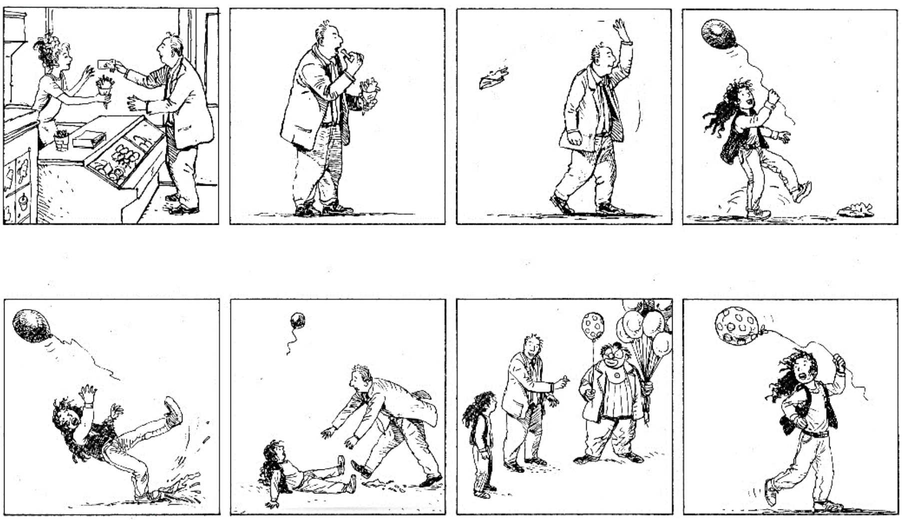

Figure 2. LPT Storytelling task 2. Copyright 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

Procedure

All children were tested at three points in time separated by 12-month intervals, in a period of two years. The narratives were orthographically transcribed using the CHAT format from the CHILDES system (MacWhinney, 2000). Each transcript contained the narratives from Storytelling tasks 1 and 2 told in this fixed order. The basic unit of analysis was the T-unit, which was defined as a single main clause plus any subordinate clause or non-clausal structure that is attached to, or embedded in it (Hunt, 1970). Coordinate clauses were transcribed and counted as separate utterances, except in cases where the coreferential subject of the second clause was elided (e.g., "*The children went down the hill and bumped into the tree*"). Sentences containing quotations, where the quote forms a full clause containing a subject and a verb, were also transcribed as separate utterances. Silent pauses were detected and their duration was measured using Audacity® software and subsequently coded on the main tier. Other speech disruptions were coded on an additional dependent tier. The reliability of the transcriptions was checked by re-transcribing 10% of the files from each group by either trained speech-language pathologists (the transcripts of the children with SLI) or trained master students in linguistics (the transcripts of the control groups). The point-to-point reliability at word level reached 95% and was 92% for identifying the types and syntactic positions of the speech disruptions.

Analysis

The background measures of the narratives (i.e., number of utterances, number of intended words, and MLU in words), were computed with the CLAN programs from the CHILDES system (MacWhinney, 2000). Intended words were the words produced in the narratives without filled pauses, interjections, repetitions, and words that were revised. All utterances produced by the children were analyzed, including the utterances that were partly unintelligible or abandoned halfway. The first step of the analysis concerned the frequency and types of speech disfluencies. Speech disruption rates were computed by dividing counts for all speech disruption types by the total number of intended words (Dollaghan & Campbell, 1992). The composite measure *total disfluencies* comprised all stalls and revisions. The measure *total stalls* contained all tallied disfluencies excluding the revisions. The different stall categories were filled pauses, interjections, part-word repetitions, whole-word repetitions, and silent pauses. Filled pauses were all occurrences of *uh* and *um*. Interjections were comments that did not add meaning to the original utterance (*<well> they <what's it called> bumped into the tree*). Part-word repetitions were repetitions of phonemes and syllables *(<wa> walked*). Whole-word repetitions were repetitions of whole words or phrases without emphatic meaning (*they walk <walk> back to the <to the> house*). Silent pauses were all silences longer than 250 ms, because according to Goldman-Eisler (1968), pauses shorter than 250 ms are related to articulation rather than to cognitive speech planning acts. The measure *total revisions* contained all overt modifications of speech output. The revisions could address phonological, semantic, or grammatical errors. Revisions were not subdivided further because the percentages of these subtypes were too low for a meaningful quantitative analysis.

The second step of the analysis concerned the durations of silent pauses. Following Guo et al. (2008), silent pauses longer than 250 ms were divided in four duration categories: (a) 250-500 ms, (b) 500-1000 ms, (c) 1000-2000 ms, or (d) over 2000 ms. Frequencies of pauses in the four duration categories were computed by dividing the number of pauses in each duration category by the number of intended words. Although these duration categories are arbitrary, the choice for identical categories as in Guo et al. ensures that results from both studies can be compared.

The third step of the analysis concerned the distribution of speech disruptions at four syntactic positions, namely utterance-initial, clause-initial, phrase-initial and word-initial. In each narrative, the numbers of utterances, clauses, phrases and intended words were computed. Disfluency rates at each syntactic position were computed by dividing the total number of disfluencies before each syntactic position by the number of possible contexts for each syntactic position. Disruptions before an utterance were always assigned to utterance onset. Disfluencies at the onset of a clause were always at a sentence-internal location where two clauses were joined. Phrases always contained a head and its modifier (e.g., noun phrase, verb phrase, prepositional phrase) and a word was a bare head. Example (3) shows how silent pauses (SP) could be distributed in an utterance. In this example, SP1 is in utterance initial position, SP2 is located before the verb phrase, SP3 is situated before a (subordinate) clause, SP4 is in word initial position, and SP5 is located before a prepositional phrase.

(3) [SP1] Ze [SP2] gingen huilen [SP3] omdat de [SP4] kar [SP5] tegen de boom geknald was.
They went cry because the cart against the tree crashed was.
'They started to cry because the cart had crashed into the tree.'

Mixed model analyses of variance (ANOVA) were used to examine differences between groups and at different time points, with Time (T1, T2, T3) as within-subjects factor and Group (SLI, LA, CA) as between-subjects factor. Significance level was set at 0.05. In order to analyze differences between groups, subsequent one-way ANOVAs with post-hoc Bonferroni correction were used.

GLM repeated measures ANOVAs were used to test differences across time within the separate groups. The assumption of sphericity was checked for all variables with Mauchly's test of sphericity. Whenever this assumption was violated, the Greenhouse-Geisser corrected values are reported.

Results

Background measures of the narratives

Table 2 summarizes the background measures of the narratives at T1, T2, and T3). Oneway ANOVA showed that no group differences for *MLU in words (MLUw)* were found at T1 ($F_{(2,87)} = 1.23$, $p = .272$) and T3 ($F_{(2,87)} = 2.24$, $p = .111$). At T2, the difference was significant ($F_{(2,87)} = 4.0619.95$, $p = .021$), and post-hoc tests showed that the CA group had a higher *MLUw* than the LA group ($p = .041$). For *number of words,* group differences were significant at T1 ($F_{(2,87)} = 7.86$, $p = .001$), T2 ($F_{(2,87)} = 5.25$, $p = .007$), and T3 ($F_{(2,87)} = 5.61$, $p = .005$). The SLI group used more words than the LA group at all three time points ($p = .001$, $p = .007$, $p = .004$, respectively) and more than the CA group at T1 ($p = .024$). For *utterances* group differences were significant T1 ($F_{(2,87)} = 12.13$, $p < .001$), T2 ($F_{(2,87)} = 9.15$, $p < .001$), and T3 ($F_{(2,87)} = 9.01$, $p < .001$). At all measurement points, the SLI group used more *utterances* than the LA group ($p < .001$, $p = .004$, $p < .001$, respectively) and the CA group ($p = .001$, $p < .001$, $p = .004$, respectively). The LA and CA groups did not differ significantly on *number of words* and *utterances.*

Table 2. *Background Measures of the Narratives*

	SLI (n = 30) *M (SD)*			LA (n = 30) *M (SD)*			CA (n = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
MLUw	6.19 (0.86)	6.53 (0.95)	6.64 (0.80)	6.23 (0.97)	6.50 (0.93)	6.74 (0.82)	6.56 (0.97)	7.20 (1.31)	7.13 (1.18)
Words	181.50 (66.24)	202.57 (57.64)	218.00 (57.64)	132.27 (37.83)	164.13 (40.50)	172.50 (52.45)	146.87 (38.83)	174.87 (42.18)	189.27 (49.22)
Utt.	29.50 (11.03)	31.40 (9.27)	32.97 (7.83)	21.03 (3.78)	25.50 (5.71)	25.53 (6.87)	22.17 (4.56)	24.40 (4.56)	26.80 (7.02)

Note. MLUw = mean length of utterance in words; Words = total number of intended words (i.e., all words produced in the narratives without filled pauses, interjections, repetitions and words that were revised); Utt. = total number of utterances.

Frequencies and types of disfluencies

The first step of the analysis concerned the frequencies and different types of disfluencies produced in the narratives by the three groups at the three time points. Table 3 displays the means and standard deviations of the speech disfluency rates by group, by time points, and by specific disfluency type.

Table 3. *Speech Disfluency Rates by Group, Time and Type*

	SLI (n = 30) *M (SD)*			LA (n = 30) *M (SD)*			CA (n = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Total disfluencies	.33	.35	.31	.24	.26	.26	.25	.23	.23
	(.11)	(.09)	(.10)	(.09)	(.07)	(.07)	(.07)	(.06)	(.07)
Stalls	.31	.31	.28	.22	.24	.23	.23	.21	.21
	(.10)	(.10)	(.09)	(.08)	(.06)	(.07)	(.07)	(.06)	(.05)
Revisions	.03	.03	.03	.02	.02	.03	.03	.02	.02
	(.01)	(.02)	(.02)	(.02)	(.02)	(.02)	(.01)	(.02)	(.01)
Different types of stalls									
Filled pauses	.03	.03	.03	.02	.02	.02	.03	.03	.03
	(.02)	(.03)	(.03)	(.02)	(.02)	(.02)	(.03)	(.02)	(.03)
Interjections	.00	.01	.01	.00	.00	.00	.00	.00	.01
	(.01)	(.01)	(.01)	(.00)	(.00)	(.00)	(.01)	(.01)	(.01)
Part-word repetitions	.01	.01	.01	.00	.00	.00	.01	.00	.00
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.00)	(.00)
Whole-word repetitions	.03	.03	.03	.03	.02	.02	.03	.02	.02
	(.02)	(.02)	(.02)	(.02)	(.01)	(.02)	(.01)	(.02)	(.02)
Silent pauses	.24	.24	.21	.17	.20	.18	.16	.16	.15
	(.08)	(.06)	(.06)	(.06)	(.06)	(.05)	(.05)	(.04)	(.05)

Note. Speech disfluency rates are counts of disfluencies divided by the number of intended words.

Total disfluencies

For *percentages of total disfluencies*, no significant interaction was found between Time and Group ($F_{(4,174)}$ = 2.20, p = .071, partial η^2 = .048). There was a significant main effect of Group ($F_{(2,87)}$ = 18.97, p < .001, partial η^2 = .304), but not of Time ($F_{(2,174)}$ = 1.50, p = .227, partial η^2 = .017). The trajectories for *percentages of total disfluencies* and group differences are presented in Figure 3(a). This schematic outline, with an arbitrary scale on the y-axis, sketches the developmental trajectories of the three groups. Group differences are indicated by brackets and asterisks. One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)} = 9.94$, $p < .001$), T2 ($F_{(2,87)} = 19.95$, $p < .001$), and T3 ($F_{(2,87)} = 7.58$, $p = .001$). Post-hoc tests revealed that the SLI group had more disfluencies than the LA and the CA groups at T1 and T2 (all p's smaller than .002) $p < .001$). At T3, the SLI group and LA group did not differ significantly ($p = .073$), but the SLI group had more disfluencies than the CA group ($p = .001$). Differences between LA and CA groups were not significant at any of the three time points. Repeated measures (RM) ANOVAs for the separate groups revealed that in the SLI group *total disfluencies* decreased significantly between T2 and T3 ($F_{(1,29)} = 6.23$, $p = .019$ partial $\eta^2 = .177$), but not between T1 and T2 ($F_{(1,29)} = 1.71$, $p = .202$ partial $\eta^2 = .056$). In the LA and CA groups no differences across time were found.

Stalls

For *percentages of stalls*, no significant interaction between Time and Group was found ($F_{(3.70,160.9)} = 1.23$, $p = .300$, partial $\eta^2 = .028$). There also was no significant main effect of Time ($F_{(1.85,160.9)} = 0.81$, $p = .440$, partial $\eta^2 = .009$), but there was a main effect of Group ($F_{(2,87)} = 17.56$, $p < .001$, partial $\eta^2 = .288$). The trajectories for *percentages of stalls* and group differences are presented in Figure 3(b). One-way ANOVAs revealed significant differences between groups at T1 ($F_{(2,87)} = 10.10$, $p < .001$), T2 ($F_{(2,87)} = 13.67$, $p < .001$) and T3 ($F_{(2,87)} = 7.23$, $p = .001$). Post-hoc testing revealed that the SLI group had more stalls than the LA controls at T1, T2, and T3 ($p < .001$, $p = .001$, $p = .030$ respectively). The SLI group also used more stalls than the CA group at T1, T2, and T3 ($p = .001$, $p < .001$, $p = .001$ respectively). There were no significant differences between the two control groups at any time point. RM ANOVAs for the separate groups revealed that the use of stalls did not change significantly in any of the groups across time.

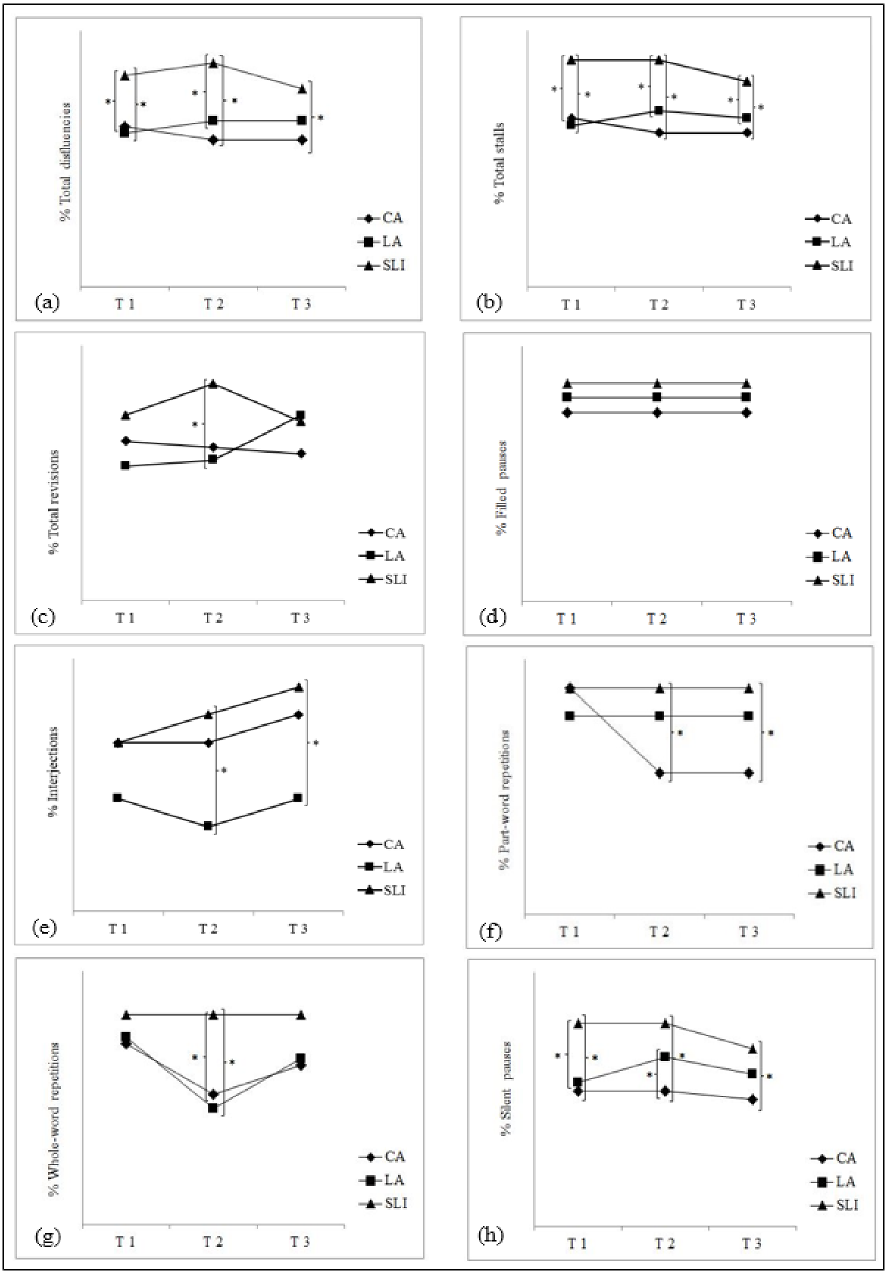

Figure 3. Disfluencies in the Groups at the three Time Points: (a) Total disfluencies, (b) Total stalls, (c) Total revisions, (d) Filled pauses, (e) Interjections, (f) Part-word repetitions, (g) Whole-word repetitions, and (h) Silent pauses (schematic outline, * = significant difference between groups).

Revisions

For *percentages of revisions*, no interaction between Time and Group was found ($F_{(4,174)}$ = 2.18, p = .073, partial η^2 = .048). There was no significant main effect of Time ($F_{(2,174)}$ = .37, p = .695, partial η^2 = .004), but the main effect of Group was significant ($F_{(2,87)}$ = 3.97, p = .022, partial η^2 = .084). The trajectories for *percentages of revisions* and group differences are presented in Figure 3(c). One-way ANOVAs revealed significant differences between groups at T2 ($F_{(2,87)}$ = 4.13, p = .019). Post-hoc testing showed that at T2, the SLI group used more revisions than the LA controls (p = .023). At other time points, and between the LA and CA groups no significant differences were found. RM ANOVAs for the separate groups yielded no significant differences over time in the SLI and CA groups. Only in the LA group, revisions increased between T2 and T3 ($F_{(1,29)}$ = 5.18, p = .030, partial η^2 = .151).

Filled pauses

For *percentages of filled pauses*, we found no significant interaction between Time and Group ($F_{(4,174)}$ = .71, p = .586, partial η^2 = .016). No significant main effect of Time was found ($F_{(2,174)}$ = 1.10, p = .336, partial η^2 = .012), and also no significant main effect of Group ($F_{(2,87)}$ = 2.61, p = .079, partial η^2 = .057). The trajectories for *percentages of filled pauses* and group differences are presented in Figure 3(d).

Interjections

For *percentages of interjections*, no significant interaction between Group and Time was found ($F_{(3.69,160.43)}$) = .26, p = .902, partial η^2 = .006). There was no significant main effect of Time ($F_{(1.84,160.43)}$) = 2.36, p = .097, partial η^2 = .026), but the main effect of Group was significant ($F_{(2,87)}$ = 5.58, p = .005, partial η^2 = .114). The trajectories for *percentages of interjections* and group differences are presented in Figure 3(e). One-way ANOVAs returned significant differences between groups at T2 ($F_{(2,87)}$ = 4.09, p = .020) and T3 ($F_{(2,87)}$ = 3.26, p = .043). Post-hoc tests revealed that at T2 and T3, the SLI group used more interjections than the LA group (p = .016, and p = .043 respectively). The LA and CA groups did not differ at any of the time points, and the SLI group did not differ from the CA controls. RM ANOVAs for the separate groups revealed no significant changes over time in the use of interjections.

Part-word repetitions

For *percentages of part-word repetitions*, a significant interaction between Group and Time was found ($F_{(4,174)}$ = 3.13, p = .017, partial η^2 = .067). There was no significant main effect of Time ($F_{(2,174)}$ = 2.33, p = .100, partial η^2 = .026). The main effect of Group was significant ($F_{(2,87)}$ = 4.06, p = .021, partial η^2 = .085). The trajectories for *percentages of part-word repetitions* and group differences are presented in Figure 3(f). One-way ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)}$ = 3.31, p = .041), T2 ($F_{(2,87)}$ = 3.61, p = .031), and T3 ($F_{(2,87)}$ = 4.12, p = .020). Post-hoc tests revealed that at T1, group differences were not significant, because of the conservative Bonferroni correction. At T2 and T3, the SLI group used more part-word repetitions than the CA controls (p = .027 and p = .021 respectively). The LA and CA groups did not differ at any time point, and the SLI group did not differ from the LA group. RM ANOVAs for the separate groups indicated that part-word repetitions decreased significantly only in the CA group between T1 and T2 ($F_{(1,29)}$ = 11.88, p = .002, partial η^2 = .291), and T1 and T3 ($F_{(1,29)}$ = 13.27, p = .001, partial η^2 = .314).

Whole-word repetitions

For *percentages of whole-word repetitions*, no significant interaction between Group and Time was found ($F_{(4,174)}$ = 1.45, p = .218, partial η^2 = .032). There was no significant main effect of Time ($F_{(2,174)}$ = 2.39, p = .095, partial η^2 = .027), but the main effect of Group was significant ($F_{(2,87)}$ = 6.82, p = .002, partial η^2 = .136). The trajectories for *percentages of whole-word repetitions* and group differences are presented in Figure 3(g). One-way ANOVAs returned significant differences between groups at T2 ($F_{(2,87)}$ = 12.67, p < .001). Post-hoc tests revealed that at T2, the SLI group used more whole-word repetitions than the LA and CA groups (p < .001, and p < .001 respectively). The two TD groups did not differ at any time point. RM ANOVAs for the separate groups indicated that whole-word repetitions decreased significantly in the LA group between T1 and T2 ($F_{(1,29)}$ = 5.79, p = .023, partial η^2 = .166), and in the CA group between T1 and T2 ($F_{(1,29)}$ = 4.62, p = .040, partial η^2 = .137).

Silent pauses

For *percentages of silent pauses*, no significant interaction was found between Time and Group ($F_{(4,174)}$ = 2.26, p = .064, partial η^2 = .049). There was a significant main effect of Group ($F_{(2,87)}$ = 16.43, p < .001, partial η^2 = .274), as well as of Time ($F_{(2,174)}$ = 3.89, p = .022, partial η^2 = .043). The trajectories for *percentages of silent pauses* and group differences are presented in Figure 3(h). One-way

ANOVAs yielded significant differences between groups at T1 ($F_{(2,87)}$ = 12.59, p < .001), T2 ($F_{(2,87)}$ = 14.06, p < .001), and T3 ($F_{(2,87)}$ = 7.25, p = .001). Post-hoc tests revealed that the SLI group used significantly more silent pauses than the LA group at T1 (p < .001) and T2 (p = .019), but not at T3 (p = .310). Compared to the CA group, the SLI group used more silent pauses at T1 (p < .001), T2 (p < .001), and T3 (p = .001). The LA group only differed from the CA controls at T2 (p = .044). RM ANOVAs for the separate groups revealed that in the SLI group silent pauses decreased between T2 and T3 ($F_{(1,29)}$ = 8.15, p = .008, partial η^2 = .219), and between T1 and T3 ($F_{(1,29)}$ = 7.08, p = .013, partial η^2 = .196) . In the LA group silent pauses decreased between T1 and T2 ($F_{(1,29)}$ = 4.69, p = .039, partial η^2 = .139), but not between T1 and T3. In the CA group the use of silent pauses did not change over time.

Duration of silent pauses

The second step of the analysis concerned the examination of silent pause duration. Table 4 displays the means and standard deviations of the silent pause rates in the four duration categories by group and the three time points.

Table 4: *Silent Pause Rates by Group and Duration Category at the three Time Points*

	SLI (*n* = 30) *M (SD)*			LA (*n* = 30) *M (SD)*			CA (*n* = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
250-500 ms	.07	.08	.07	.06	.07	.07	.06	.05	.06
	(.02)	(.03)	(.03)	(.03)	(.03)	(.04)	(.03)	(.02)	(.03)
500-1000 ms	.09	.08	.09	.07	.08	.08	.07	.07	.06
	(.03)	(.02)	(.03)	(.03)	(.03)	(.02)	(.02)	(.03)	(.02)
1000-2000 ms	.06	.06	.04	.03	.04	.03	.03	.04	.03
	(.04)	(.03)	(.03)	(.02)	(.02)	(.03)	(.03)	(.02)	(.02)
>2000 ms	.02	.02	.01	.02	.01	.00	.01	.01	.00
	(.02)	(.02)	(.01)	(.02)	(.02)	(.01)	(.02)	(.01)	(.01)

Note. Silent pause rates are the counts of silent pauses in the four duration categories divided by the number of intended words.

Pauses 250-500 ms

For *percentages of silent pauses 250-500 ms* no significant interaction between Group and Time was found ($F_{(4,174)}$ = 2.16, p = .075, partial η^2 = .047). There was a significant main effect of Time ($F_{(2,174)}$ = 3.62, p = .029, partial η^2 = .040). The main effect of Group was significant ($F_{(2,87)}$ = 4.83, p = .010, partial η^2 = .100).

The trajectories for silent pauses 250-500 ms and group differences are presented in Figure 4(a). One-way ANOVAs returned significant differences between groups at T2, ($F_{(2,87)}$ = 9.17, p < .001). Post-hoc tests revealed that at T2, the SLI group and the LA group used more *pauses 250-500 ms* than the CA group (p < .001 and p = .029 respectively). RM ANOVAs for the separate groups revealed that *pauses 250-500 ms* increased significantly in the SLI group between T1 and T2 ($F_{(1,29)}$ = 7.20, p = .012, partial η^2 = .199), and decreased between T2 and T3 ($F_{(1,29)}$ = 4.75, p = .038, partial η^2 = .199). In the LA group, a significant increase was found between T1 and T2 ($F_{(1,29)}$ = 5.18, p = .030, partial η^2 = .151). In the CA group, no significant changes over time were found.

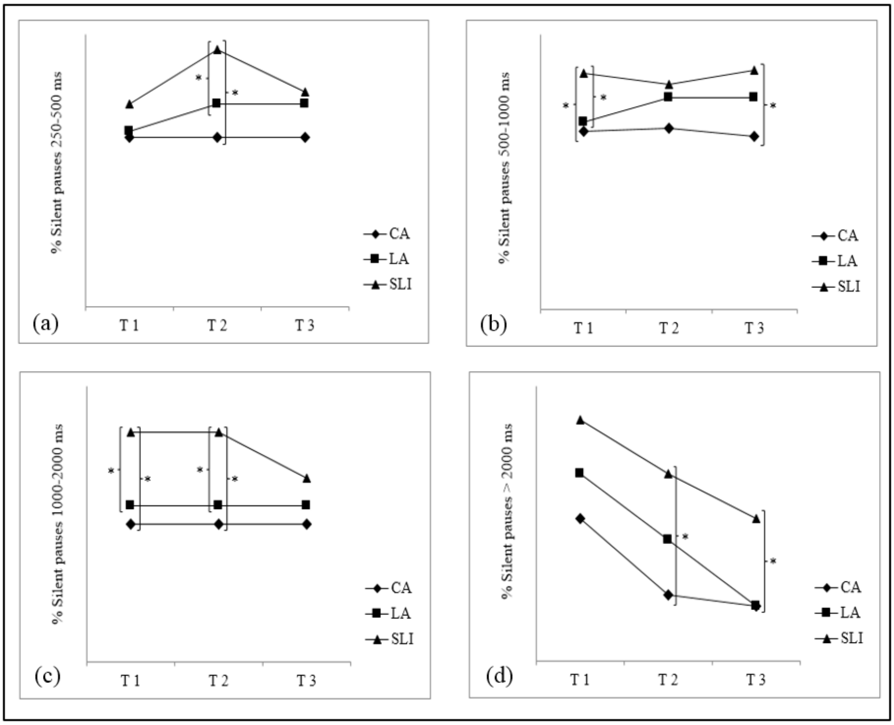

Figure 4. Silent Pauses in the Groups in 4 Duration Categories: (a) 250-500 ms, (b) 500-1000 ms, (c) 1000-2000 ms, and (d) > 2000 ms (* = significant difference between groups).

Pauses 500-1000 ms

For *percentages of silent pauses 500-1000ms* no significant interaction between Group and Time was found ($F_{(4,174)}$ = 1.05, p = .385, partial η^2 = .023). There was no significant main effect of Time ($F_{(2,174)}$ = .34, p = .715, partial η^2 = .004). The main effect of Group was significant ($F_{(2,87)}$ = 7.10, p = .001, partial η^2 = .140). The trajectories for silent pauses 500-1000 ms and group differences are presented in Figure 4(b). One-way ANOVAs revealed significant differences between groups at T1, ($F_{(2,87)}$ = 5.07, p = .008) and at T3 ($F_{(2,87)}$ = 5.79, p = .004). Post-hoc tests showed that at T1, the SLI group used more pauses 500-1000 ms than the LA and the CA group (p = .038 and p = .013). At T3, the SLI group used more pauses 500-1000 ms than the CA group (p = .003). The LA and CA groups did not differ at any time point. RM ANOVAs for the separate groups revealed no significant changes in the three groups over time.

Pauses 1000-2000 ms

For silent pauses 1000-2000 ms a significant interaction between Group and Time was found ($F_{(4,174)}$ = 2.84, p = .026, partial η^2 = .061). There was a significant main effect of Time ($F_{(2,174)}$ = 5.62, p = .004, partial η^2 = .061) and of Group ($F_{(2,87)}$ = 11.40, p < .001, partial η^2 = .208). The trajectories for silent pauses 1000-2000 ms and group differences are presented in Figure 4(c). One-way ANOVAs returned significant group differences at T1, ($F_{(2,87)}$ = 10.97, p < .001) and T2 ($F_{(2,87)}$ = 7.57, p = .001). Post-hoc tests revealed that the SLI group used more pauses 1000-2000 ms than the LA and CA group at T1 (p = .001 and p < .001 respectively) and at T2 (p = .004 and p = .003 respectively). The LA and CA groups did not differ at any time point. RM ANOVAs for the separate groups revealed a significant decrease in the SLI group between T1 and T3 ($F_{(1,29)}$ = 17.19, p < .001, partial η^2 = .372) and between T2 and T3 ($F_{(1,29)}$ = 9.17, p = .005, partial η^2 = .240).

Pauses > 2000 ms

For *percentages of silent pauses > 2000 ms* no significant interaction between Group and Time was found ($F_{(3.14,136.49)}$ = .54, p = .666, partial η^2 = .012). There was a significant main effect of Time ($F_{(1.57,136.49)}$ = 13.10, p < .001, partial η^2 = .131). The main effect of Group was significant ($F_{(2,87)}$ = 4.25, p = .017, partial η^2 = .089). The trajectories for silent pauses > 2000 ms and group differences are presented in Figure 4(b). One-way ANOVAs returned significant differences between groups at T2, ($F_{(2,87)}$ = 5.11, p = .008) and T3 ($F_{(2,87)}$ = 4.99, p = .009). Post-hoc tests revealed that at T2, the SLI group used more pauses >2000 ms than the CA group (p = .006). At T3, the SLI group used more pauses >2000 ms than the LA and CA groups (p = .027 and p = .019 respectively). The two control groups did not differ at any time point. RM ANOVAs for the separate groups revealed a significant decrease in the SLI group between T1 and T3 ($F_{(1,29)}$ = 6.96, p = .013, partial η^2 = .193). In the LA group a significant decrease was found between T1 and T2 ($F_{(1,29)}$ = 5.04, p = .033, partial η^2 = .148), between T2 and T3 ($F_{(1,29)}$ = 6.72, p = .015, partial η^2 = .188), and between T1 and T3 ($F_{(1,29)}$ = 14.48, p = .001, partial η^2 = .333). In the CA group a significant decrease was found between T1 and T3 ($F_{(1,29)}$ = 4.33, p = .046, partial η^2 = .130).

Speech disruptions by syntactic position

The last step of the analysis concerned the distribution of speech disfluencies at the different syntactic positions, Table 5 presents the speech disfluency rates by group, by time points, and by syntactic position, namely utterance initial, clause initial, phrase initial, and word initial.

Table 5. *Speech Disruption Rates by Group and Syntactic Position at the three Time Points*

	SLI (*n* = 30) *M (SD)*			LA (*n* = 30) *M (SD)*			CA (*n* = 30) *M (SD)*		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Utterance	1.16	1.21	1.13	1.00	.97	1.05	.97	1.03	.99
	(.31)	(.27)	(.35)	(.39)	(.29)	(.29)	(.29)	(.30)	(.20)
Clause	.38	.63	.41	.13	.13	.26	.18	.25	.52
	(.58)	(.82)	(.53)	(.39)	(.29)	(.45)	(.31)	(.37)	(.60)
Phrase	.14	.16	.13	.08	.15	.11	.11	.09	.08
	(.08)	(.12)	(.09)	(.05)	(.16)	(.06)	(.05)	(.05)	(.06)
Word	.06	.06	.05	.02	.03	.03	.03	.03	.03
	(.04)	(.03)	(.03)	(.02)	(.02)	(.03)	(.02)	(.02)	(.02)

Note. Speech disruption rates at the different syntactic positions are the counts of disfluencies found at the onset of utterances, clauses, phrases and words divided by the possible syntactic positions.

Utterance initial

For the *utterance initial disfluencies*, no significant interaction between Group and Time was found ($F_{(4,174)}$ = .72, p = .577, partial η^2 = .016). There was no significant main effect of Time ($F_{(2,174)}$ = .17, p = .841, partial η^2 = .002), but the main effect of Group was significant ($F_{(2,87)}$ = 5.99, p = .004, partial η^2 = .121).

One-way ANOVAs showed that significant differences between groups were found at T2 ($F_{(2,87)}$ = 5.33, p = .007). Post-hoc tests revealed that at T2, the SLI group produced more utterance initial disfluencies than the LA controls (p = .007). The LA and CA groups did not differ from each other at any of the time points. RM ANOVAs for the separate groups revealed no changes in the three groups over time.

Clause initial

For the *clause initial disfluencies*, a significant interaction between Group and Time was found ($F_{(4,174)}$ = 2.54, p = .042, partial η^2 = .055). There was a significant main effect of Time ($F_{(2,174)}$ = 3.13, p = .046, partial η^2 = .035) and for Group ($F_{(2,87)}$ = 5.69, p = .005, partial η^2 = .116). One-way ANOVAs showed significant differences between groups at T2 ($F_{(2,87)}$ = 6.78, p = .002) Post-hoc tests revealed that at T2, the SLI group produced more clause initial disfluencies than the LA and CA controls (p = .002, and p = .024 respectively). The LA and CA groups did not differ from each other at any of the time points. RM ANOVAs for the separate groups revealed that clause initial disfluencies increased significantly in the CA group between T2 and T3 ($F_{(1,29)}$ = 4.56, p = .041, partial η^2 = .136) and between T1 and T3 ($F_{(1,29)}$ = 8.00, p = .008, partial η^2 = .216).

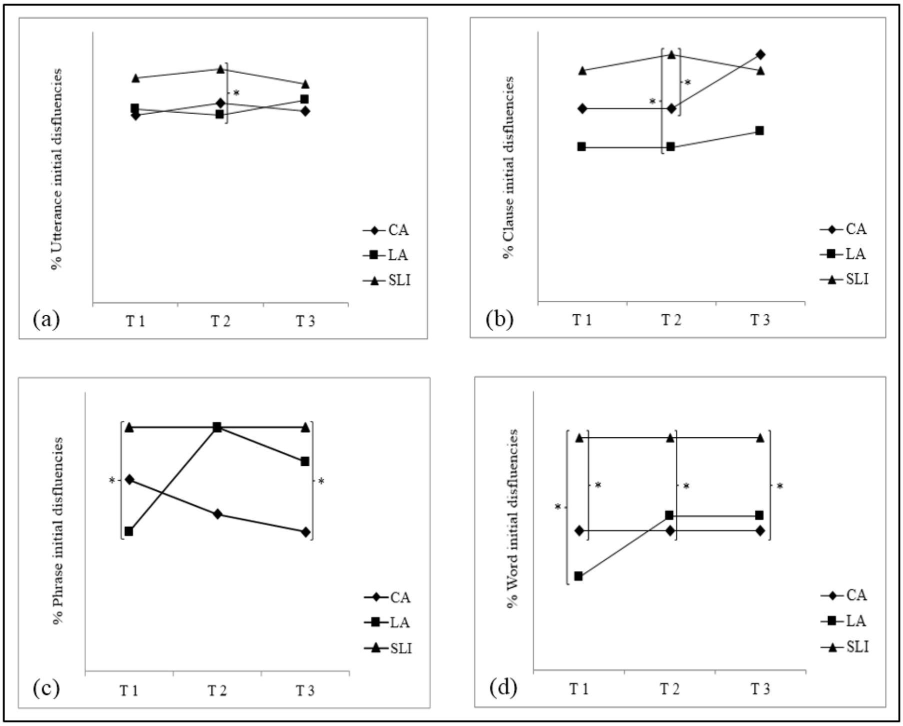

Figure 5. Speech Disruptions by Syntactic Position: (a) Utterance initial, (b) Clause initial, (c) Phrase initial, and (d) Word initial (schematic outline, * = significant difference between groups)

Phrase initial

For the *phrase initial disfluencies*, no significant interaction between Group and Time was found ($F_{(3.13,135.98)}$ = 2.28, p = .080, partial η^2 = .050). There was no significant main effect of Time ($F_{(1.56,135.98)}$ = 1.91, p = .161, partial η^2 = .021), but the main effect of Group was significant ($F_{(2,87)}$ = 6.48, p = .002, partial η^2 = .130). One-way ANOVAs showed that significant differences between groups were found at T1, ($F_{(2,87)}$ = 6.47, p = .002), T2 ($F_{(2,87)}$ = 3.19, p = .046), and T3 ($F_{(2,87)}$ = 4.24, p = .018). Post-hoc tests revealed that at T1, the SLI group used more *phrase initial disfluencies* than the LA group (p = .002). At T3, the SLI group used more *phrase initial disfluencies* than the CA group (p = .014). The LA and CA groups did not differ from each other at any test round. RM ANOVAs for the separate groups revealed that *phrase initial disfluencies* did not change over time in the SLI group. In the LA group, a significant increase was found between T1 and T2 ($F_{(1,29)}$ = 4.35, p = .046, partial η^2 = .130). In the CA group *phrase initial* *disfluencies* decreased between T1 and T2 ($F_{(1,29)} = 5.37$, $p = .028$, partial $\eta^2 = .156$), and between T1 and T3 ($F_{(1,29)} = 5.49$, $p = .026$, partial $\eta^2 = .159$).

Word initial

For the *word initial disfluencies*, no significant interaction between Group and Time was found ($F_{(4,174)} = 1.99$, $p = .098$, partial $\eta^2 = .044$). There was no significant main effect of Time ($F_{(4,174)} = 1.53$, $p = .221$, partial $\eta^2 = .017$). The main effect of Group was significant ($F_{(2,87)} = 19.05$, $p < .001$, partial $\eta^2 = .305$). One-way ANOVAs showed that significant differences between groups were found at T1 ($F_{(2,87)} = 14.21$, $p < .001$), T2 ($F_{(2,87)} = 15.34$, $p < .001$), and T3 ($F_{(2,87)} = 4.11$, $p = .020$). Post-hoc tests revealed that at T1, the SLI group used more *word initial disfluencies* than the LA ($p < .001$) and CA groups ($p = .001$). At T2, the SLI group used more *word initial disfluencies* than the LA ($p < .001$) and CA group ($p < .001$). At T3, the SLI group used more *word initial disfluencies* than the CA group ($p = .031$). The LA and CA groups did not differ from each other at the three time points. RM ANOVAs for the separate groups revealed that *word initial disfluencies* did not change over time in the SLI group and the CA group. In the LA group, a significant increase was found between T1 and T2 ($F_{(1,29)} = 5.32$, $p = .028$, partial $\eta^2 = .155$) and between T1 and T3 ($F_{(1,29)} = 5.50$, $p = .026$, partial $\eta^2 = .159$).

Discussion

In this study, we investigated speech disruptions longitudinally in school-age children with SLI and LA and CA control groups. We were interested in the differences in frequencies, types and syntactic distribution of disfluencies between the three groups and in the development of fluency over time.

Our first research question concerned the differences in disfluencies between the groups. We expected the SLI group to have more (total) disfluencies than the CA group, but to perform similarly to the LA group. Our expectations were only partly confirmed, as the SLI group produced more disfluencies than *both* control groups at T1 and T2. At T3, the SLI group had more disfluencies than the CA group, but equaled the LA group. The overall difference between the SLI group and the two control groups appears to be carried by the frequency of stalls; the SLI group had more stalls than both the LA and CA groups at each time point, but this result was not found for revisions. The SLI group used revisions to the same extent as the CA group, but outnumbered the LA group at

T2. Similar results were also found by other researchers (Boscolo et al., 2002; Guo et al., 2008) and suggest that the children with SLI are able to monitor their speech and make repairs just as well as their peers. Moreover, no group differences were found for filled pauses and whole-word repetitions. For interjections, the SLI group also performed similarly to the CA group, but produced more interjections than the LA group at T2 and T3. In addition, at all three time points, the SLI group had more silent pauses than the CA group, but only at T1 when compared to the LA group.

Silent and filled pauses, interjections and whole-word repetitions are all regarded as normal disfluencies. It may be the case that the children with SLI and the TD control groups used filled pauses and interjections as a means to keep the floor and whole-word repetitions to accommodate the listener, as suggested by Clark (2002). However, an obvious alternative explanation is that the normal (non-stutter-like) disfluencies were mostly used to buy time for sentence formulation (Rispoli, 2003). The narrative elicitation task used in the present study was targeted at producing a monologue and thus the setting lacked a true conversation partner. Furthermore, the higher silent pause rates in the SLI group compared to TD peers also suggest that children with SLI need more time to prepare their utterances, which may be due to processing difficulties (e.g., retrieval of lexical information and construction of sentence frames).

With respect to stutter-like disfluencies, we found that the SLI group used more part-word repetitions at T2 and T3 than the CA controls, but not when compared to the LA controls. Several studies have found higher rates of stutter-like disfluencies in children with SLI (Boscolo, et al., 2002; Hall, 1996; Nettelbladt & Hansson, 1999). It has been suggested that pressure from language remediation may be a factor leading to increased stutter-like disfluencies (Merits-Patterson & Reed, 1981). However, the children with SLI in our study did not differ from the LA controls on these stutter-like disfluencies, which makes a 'pressure' or 'awareness' explanation less viable. An alternative explanation could be that treatment aimed at learning new language skills would cause an increase of these disfluencies. The children with SLI would then go through a phase of increased disfluency in the process of mastering new structures, similar to what has been observed in young TD children (Colburn & Mysak, 1982a, 1982b; Hall, 1977). A narrative task is more demanding than conversational speech (Blankenstijn & Scheper, 2003; MacLachlan & Chapman, 1988). Possibly, narrative production is equally taxing for children with weaker language skills or less language experience, i.e., both the SLI and LA groups, and is therefore leading to more stutter-like disfluencies. According to Guo et al. (2008) part-word repetitions arise from difficulties in retrieving the phonological information of lexical items. This explanation matches proposals that children with SLI have weaker phonological representations of lexical items (Leonard & Deevy, 2004; Messer & Dockrell, 2006). At this point, further elaboration on the use of stutter-like disfluencies is not appropriate, because we only examined part-word repetitions and did not investigate tensed pauses, blocks, or prolongations. In summary, this study confirmed that children with SLI produced more disfluencies than TD children. For the most part, frequencies and types of disfluencies in the SLI group resembled those in the LA control group more than those of the CA control group.

Our second research question concerned the development of fluency over time. We investigated whether fluency improved in the children with SLI, as has been observed in TD children (Kowal, et al., 1975, Wijnen, 1990) and in children with SLI (Hall, 1996). We predicted that fluency would increase in all three groups because of advances in linguistic skills over time. This prediction was only partly confirmed. On the measure total disfluencies, no changes over time were found in the control groups. However, in the SLI group, total disfluencies decreased significantly between T2 and T3, which resulted solely from a decrease in longer silent pauses (i.e., 1000-2000 ms and > 2000 ms), because no changes were found for the other disfluency measures. Furthermore, in all three groups, silent pauses > 2000 ms significantly decreased. In the LA group, revisions increased between T2 and T3, suggesting that monitoring still develops in younger TD children. In addition, in the CA group part-word repetitions and phrase-initial disfluencies decreased between T1 and T3. This result was not observed in the SLI and LA groups, suggesting that a decrease of these stutter-like and phrase-initial disfluencies might reflect further linguistic sophistication in the older TD children. Across the board, the developmental changes in disfluency were rather limited, with most improvement observed in the SLI group. It appears that between the ages 8 and 10 years, the children with SLI slightly improve in lexical retrieval and the construction of sentence frames.

Our third research question concerned the distribution of disfluencies at specific syntactic positions. We expected the SLI group, because of their syntactic deficits, to have more disfluencies at utterance onset, clause onset, and at phrase onset, where preparation of utterance structure (grammatical encoding) takes place. In addition, higher word-initial disfluency rates were also expected in the SLI group because of poor lexical retrieval. In all three groups, the highest frequencies of disfluencies were found at utterance-initial position, followed by clause-initial, phrase-initial, and word-initial positions. This result indicates that for all groups, forming multi-word syntactic units is more challenging than forming units consisting of just one word. This result is in accordance with the literature outlined earlier. Clearly children with SLI and TD children did not differ in this respect. However, for word-initial disfluencies we found that the children with SLI outnumbered the CA group at all three time points and the LA group at T1. Furthermore, when we inspect the proportions of disfluencies at the syntactic positions (Table 5) and compute ratios for utterance-initial/word-initial positions, for the SLI group, these ratios are 19.3 (T1), 20.2 (T2), and 22.6 (T3); for the LA group 50, 32.3, and 35; and for the CA group 32.3, 34.3, and 33, respectively. These ratios illustrates that for the SLI group, roughly around every 20 utterance-initial disfluencies a word-initial disfluency occurs, although this number is approximately 32 or higher in the control groups. In other words, word-initial disfluencies are proportionally higher in the SLI group than in the LA and CA control groups. These word-initial disfluencies may be regarded as reflecting problems with lexical retrieval, because of weak semantic and/or phonological representations (Leonard & Deevy, 2004; Messer & Dockrell, 2006). It has been argued that speech disfluencies are characteristic of word-finding difficulties, together with using empty words and substitutions (German, 1987; German & Newman, 2004, German & Simon, 1991). However, the theoretical construct of word-finding difficulties has not yet been clearly defined and validated (Messer & Dockrell, 2006; Tingley, Kyte, Johnson, & Beitchman, 2003). Nonetheless, lexical deficits have been shown to persist in children with SLI (McGregor, Oleson, Bahnsen, & Duff, 2013) and word-finding difficulties can be regarded as one of the hallmarks of school-age children with SLI (Seiger-Gardner & Schwartz, 2008). Because sentence planning is not regarded to occur at word-initial positions, difficulties with lexical retrieval seem the most obvious explanation for the higher frequencies of word-initial disfluencies.

Although Guo et al. (2008) and the present study investigated disfluencies in different languages, both studies can be compared because of similar research design and participant age. In Guo et al. (2008), differences between the 10-year-old SLI group and CA group were restricted to silent pauses 500-1000 ms and disfluencies before phrases. No differences were found between the SLI and LA groups. However, we found differences between the 10-year-old children with SLI and both TD groups. In our study, the SLI group outnumbered the CA group in total disfluencies, stalls, silent pauses (i.e., 500-1000 ms and > 2000 ms), part-word repetitions, and phrase-initial and word-initial disfluencies. When compared to the LA group, the children with SLI produced more stalls, interjections, and more pauses > 2000 ms. With respect to syntactic distribution, similar to Guo et al., we found no differences between the SLI group and the LA group.

These differences in results between both studies might be attributed to differences in methodology. Both studies used a different narrative task with a different protocol. In our study, two story generation tasks were used, both consisting of eight pictures each. Guo et al. (2008) used one set of three pictures, and another three picture set was used for practice. In the practice phase, key elements of the practice story were identified and a prewritten model story was read out to the child. When the child failed to name the key elements in the following test story, a full description of these elements was given before storytelling started. In our study, no such practice or help was included in the protocol. Our narrative tasks may have been more taxing, leading to more speech disfluencies in the SLI group as compared to both control groups.

It is also possible that the children in Guo et al (2008) were less severely impaired than the Dutch children with SLI in the present study. The admission criteria for the Dutch schools for children with severe language disorders are rather strict. For instance, Dutch children with SLI not only have to score -1.5 *SD* on standardized tests, their language difficulties also have to be demonstrably resistant to therapy and their academic achievement has to be below the 10th percentile. Yet another possibility for the differences in results may be that disfluency types and distribution may vary across languages and speech disfluency is partly language-specific. However, this option is very unlikely and there is no evidence to support this in the literature.

One final observation that needs to be discussed was the striking differences in narrative length between the SLI and TD groups. The SLI group used more utterances and more words in their narratives than the control groups. Similar findings were reported by other researchers for Dutch speaking children with SLI (de Jong, 1999; Steenge, 2006), whereas studies on English speaking children often reported the opposite (Guo et al., 2008; Reilly, Losh, Bellugi, & Wulfeck, 2004). Possibly, the children with SLI used more compound sentences, which were transcribed as separate utterances, thus inflating the total number of utterances. Another possibility is that a training effect caused these longer stories in the SLI group. Story generation materials are widely used by speech therapists to improve narrative skills, whereas TD children generally do not practice their storytelling skills. Perhaps the children with SLI also used less specific vocabulary and were poorly at establishing cohesion and coherence, which would then lead to extra words and sentences to form a comprehensive story. However, the differences in story length cannot account for the difference between the results from Guo et al. and our study, because all speech disruption rates were adjusted by dividing all counts of speech disruptions by numbers of intended words and possible syntactic positions.

Conclusion

The disfluency patterns of the children with SLI resemble those of the LA control group more than those of the CA control group. The higher disfluency rates in the SLI group are interpreted as related to their (impaired) expressive language level. The distribution of disfluencies at syntactic positions follows the same pattern in all three groups. However, the children with SLI produce more disfluencies at word-initial positions, suggesting that problems with lexical retrieval may contribute to their disfluency. Over time, the percentage of silent pauses steadily decreases in the SLI group, but no changes are observed for the other types of disfluencies. This suggests that, although some improvement in speech fluency is seen, sentence formulation continues to be challenging for school-age children with SLI. This study adds to the existing literature suggesting that disfluency during expressive language tasks reflects the syntactic and lexical deficits of children with SLI.

Clinical implications

The administration of standardized language tests and analysis of spontaneous speech or narratives are generally seen to be complementary. Taken together, they provide the best clinical picture for accurate diagnosis and selection of targets for intervention. In older school-age children with SLI, an additional investigation of speech disfluencies may be considered to assess subtle language difficulties that go beyond grammatical complexity and grammatical correctness (Finneran et al., 2009). Such an analysis may also help to diagnose word-finding difficulties.

However, the coding and analysis of speech disfluencies is a laborious task. In the present study, no differences between children with SLI and age peers were found for filled pauses, interjections, whole-word repetitions, and revisions. Therefore, it may suffice to restrict the disfluency analysis to silent pauses larger than 250 ms and part-word repetitions. Currently, software for automatic pause detection in speech samples is already available (De Jong & Wempe, 2009). Rapid advances in speech technology will probably facilitate assessment of fluency in clinical populations in the near future.

Limitations of the present study and future research

This study did not include a deeper investigation of revision types. Future studies examining different revision categories could reveal whether children with SLI use the same types of revisions as TD children, and whether their monitoring system is sensitive to the same errors as observed in TD children. We also did not examine stutter-like disfluencies in depth, as we only analyzed part-word repetitions. In a future study, other stutter-like disfluencies such as prolongations, tensed pauses, and blocks should also been taken into account in order to arrive at a deeper understanding why children with SLI use more stutter-like disfluencies than their peers.

Furthermore, studies that compare disfluency in different tasks, such as narratives, sentence repetition and sentence elicitation could inform researchers on possible task effects of these different elicitation methods. Another issue demanding more research are the contributions of syntactic deficits and lexical deficits (i.e., word-finding difficulties) to disfluency in children with SLI. Possibly, there is a trade-off over time, when grammatical abilities of children with SLI gradually improve, but lexical retrieval does not. Grammatical and lexical processes are difficult to disentangle in narratives or discourse. Research on speech disfluency that also takes word frequency, as well as semantic and phonological complexity of words into account, may further our understanding of the contribution of lexical access and retrieval in speech disfluencies of children with SLI.

Acknowledgments

This study was supported by the Royal Auris Group, Gouda (The Netherlands). We thank Cito for their permission to use the images of the LPT Storytelling tasks. We also thank the children, parents and staff of all the schools that participated; and Irene Sormani, Marij van Ewijk, Henaly Leijenhorst, Marjolein van der Horst, and Merel van Witteloostuijn for their help with the transcriptions.

References

Blankenstijn, C., & Scheper, A. (2003). *Language development in children with psychiatric impairment*. Utrecht: LOT Dissertation Series 82. (Doctoral dissertation, University of Amsterdam). Retrieved from http://www.lotpublications.nl/publish/articles/000542/bookpart.pdf.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gersbacher (Ed.), Handbook of psycholinguistics (pp. 945–984). San Diego, CA: Academic Press.

Boscolo, B., Bernstein Ratner, N., & Rescorla, L. (2002). Fluency of school-aged children with a history of specific expressive language impairment: An exploratory study. *American Journal of Speech-Language Pathology, 11,* 41-49.

Clark, H. H. (2002). Speaking in time. *Speech communication, 36*, 5-13.

Colburn, N., & Mysak, E. (1982a). Developmental disfluency and emerging grammar I: Dysfluency characteristics in early syntactic utterances. *Journal of Speech and Hearing Research, 25,* 414-420.

Colburn, N., & Mysak, E. (1982b). Developmental disfluency and emerging grammar II: Co-occurrence of disfluency with specified semantic-syntactic structures. *Journal of Speech and Hearing Research, 25,* 421-427.

De Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure.* Groningen: Grodil. (Doctoral dissertation, University of Groningen). Retrieved from: http://irs.ub.rug.nl/ppn/297976826.

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*, 385–90.

Dockrell, J. E., Messer, D., George, R., Wilson, G. (1998). Children with word-finding difficulties-Prevalence, presentation and naming problems. *International Journal of Language and Communication Disorders*, *33,* 445–454.

Dollaghan, C.A., & Campbell, T.F. (1992). A procedure for classifying disruptions in spontaneous language samples. *Topics in Language Disorders, 12,* 56-68.

Finneran, D. A., Leonard, L. B., & Miller, C. A. (2009). Speech disruptions in the sentence formulation of school-age children with specific language impairment. *International Journal of Language and Communication Disorders, 44,* 271-286.

German, D. (1987). Spontaneous language profiles of children with word-finding problems. *Language, Speech, and Hearing Services in Schools*, *18*, 217–230.

German, D. J., & Newman, R. S. (2004). The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research, 47*, 624–636.

German, D., & Simon, E. (1991). Analysis of children's word-finding skills in discourse. *Journal of Speech and Hearing Research*, *34*, 309–316.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* New York: Academic Press.

Guo, L., Tomblin, J. B., & Samelson, V. (2008). Speech disruptions in the narratives of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 51,* 722-738.

Hall, N. E. (1996). Language and fluency in child language disorders: changes over time. *Journal of Fluency Disorders, 21*, 1-32.

Hall, N. E., Yamashita, T. S., & Aram, D. M. (1993). Relationship between language and fluency in children with developmental language disorders. *Journal of Speech and Hearing Research, 36,* 568-579.

Hall, P. K. (1977). The occurrence of disfluencies in language-disordered school-age children. *Journal of Speech and Hearing Disorders*, *42*, 364–369.

Hunt, K. W. (1970). Syntactic maturity in school children and adults. *Monograph of the Society for Research in Child Development, 134,* 1-67.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children: Interpretative Manual.* Circle Pine, MN: American Guidance Service.

Kowal, S., O'Connell, D., & Sabin, E. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research, 4*, 195-207.

Lees, R., Anderson, H., & Martin, P. (1999). The influence of language disorder on fluency: A pilot study. *Journal of Fluency, 24,* 227–238.

Leonard, L. B. (1998). *Children with Specific Language Impairment*. Cambridge: MIT Press.

Leonard, L. B., & Deevy, P. (2004). Lexical deficits in specific language impairment. In L. Verhoeven, & H. van Balkom (Eds), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 209-233). Mahwah, NJ: Erlbaum.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Messer, D., & Dockrell, J. (2006). Children's naming and word-finding difficulties: Descriptions and explanations. *Journal of Speech, Language, and Hearing Research*, *49*, 309–324.

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language and Communication Disorders*, *48*, 307–19.

MacLachlan, B. G., & Chapman, R. S. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders*, *53*, 2–7.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Merits-Patterson, R., & Reed C. G. (1981). Disfluencies in the speech of language-delayed children. *Journal of Speech and Hearing Research*, *46*, 55–58.

Nettelbladt, U., & Hansson, K. (1999). Mazes in Swedish pre-school children with specific language impairment. *Clinical Linguistics and Phonetics, 13*, 483-497.

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36,* 472-487.

Reilly, J., Losh, M., Bellugi, U. & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language, 88*, 229-247.

Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *Journal of Speech, Language, and Hearing Research, 46*, 818-830.

Rispoli, M., & Hadley, P. (2001). The leading-edge: the significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research, 44,* 1131-1143.

Rispoli, M., Hadley, P., & Holt, J. (2008). Stalls and revisions: a developmental perspective on sentence production. Journal of Speech, Language, and Hearing Research, *51,* 953-966.

Seiger-Gardner, L., & Schwartz, R. G. (2008). Lexical access in children with and without specific language impairment: A cross-modal picture-word interference study. *International Journal of Language and Communication Disorders*, *43*, 528–551.

Schwartz, R. G. (2009). Specific language impairment. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 3-43). New York, NY: Psychological Press.

Steenge, J. (2006). *Bilingual children with specific language impairment: Additionally disadvantaged?* (Doctoral dissertation). Nijmegen: EAC, Research Centre on Atypical Communication, Radboud University Nijmegen. Retrieved from: http://repository.ubn.kun.nl/handle/2066/29807.

Tingley, S. J., Kyte, C. S., Johnson, C. J., & Beitchman, J. H. (2003). Single-word and conversational measures of word-finding proficiency. *American Journal of Speech-Language Pathology*, *12*, 359–68.

Van Bon, W. H. J. (1986). Coloured Progressive Matrices: Manual of Dutch Norms. London: Lewis.

Van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry, 47*, 176-189.

Verhoeven, L. & Vermeer, A. (2001). *Taaltoets Alle Kinderen* [LPT, Language Proficiency Test for All Children]. Arnhem: Cito. Verhoeven, L. & Vermeer, A. (2001). *Taaltoets Alle Kinderen* [LPT, Language Proficiency Test for All Children]. Arnhem: Cito.

Wijnen, F. (1990). The development of sentence planning. *Journal of Child Language, 17,* 651-675.

Wijnen, F., & Verrips, M. (1998). The acquisition of Dutch syntax. In S. Gillis, & A. De Houwer (Eds.), *The acquisition of Dutch* (pp. 223-299). Amsterdam: John Benjamins.

Yairi, E., & Ambrose, N. (1992). A longitudinal study of stuttering in young children: A preliminary report. *Journal of Speech and Hearing Research*, *35*, 755-760.

5 Enhancing complex syntax in children with SLI: a metalinguistic and multimodal approach [4]

Abstract

Background: Currently, most research on effective treatment of morphosyntax in children with specific language impairment (SLI) pertains to younger children. In the last two decades, several studies have provided evidence that intervention for older school-age children with SLI can be effective. These metalinguistic intervention approaches teach grammatical rules explicitly and use shapes and colors as two-dimensional visual support. Reading or writing activities form a substantial part of these interventions. However, some children with SLI are poor readers and the two-dimensional approaches used so far, might be considered as somewhat static.

Aims: The present study examined the effectiveness of a combined metalinguistic and multimodal approach in older school-age children with SLI. The intervention was adapted to suit poor readers and targeted the improvement of relative clause production, because relative clauses still pose difficulties for older children with SLI.

Methods & Procedures: Participants were 12 monolingual Dutch children with SLI (mean age 11;2). All children visited a special school for children with speech and language disorders in the Netherlands. A quasi-experimental multiple-baseline design was chosen to evaluate the effectiveness of the intervention. A set of tasks was constructed to test relative clause production and comprehension. Two balanced versions were alternated in order to suppress a possible learning effect from multiple presentations of the tasks. After 3 monthly baseline measurements, the children received individual treatment with a protocolled intervention program twice a week during 5 weeks. The tests were repeated directly post-therapy and at a retention measurement 3 months later. During the intervention

[4] Submitted (Authors: Rob Zwitserlood[a,c], Frank Wijnen[a], Marjolijn van Weerdenburg[b], and Ludo Verhoeven[b], [a]Utrecht Institute of Linguistics (OTS), University Utrecht, The Netherlands, [b]Behavioural Science Institute, Radboud University Nijmegen, The Netherlands, [c]Koninklijke Auris Groep, The Netherlands).

program, the speech therapist delivering the treatment remained blind to the test results.
Outcomes & Results: No significant changes were found during the baseline measurements. However, measurement directly post-therapy showed that 5 hours of intervention produced significant improvement on the relative clause production tasks, but not on the relative clause comprehension task. The gains were also maintained 3 months later.
Conclusions & Implications: The motor and tactile/kinesthetic dimensions of the 'Metataal' metalinguistic intervention approach are a valuable addition to the existing metalinguistic approaches. This study supports the evidence that grammatical skills in older school-age children with SLI can be remediated with direct intervention using a metalinguistic approach. The current tendency to diminish direct intervention for older children with SLI should be reconsidered.
Keywords: specific language impairment, intervention, metalinguistic, Dutch, multimodal, school-age children with SLI.

Specific language impairment (SLI) is a neurodevelopmental language disorder that affects approximately 7% of kindergarten children (Tomblin et al., 1997) and can persist into adolescence and adulthood (Aram, Ekelman, & Nation, 1984; Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). Children with SLI can experience difficulties in many areas of language, but problems with verb morphology and complex syntax are generally considered to be core symptoms of the disorder (Leonard, 1998). Persisting language impairment has severe effects both on academic achievements (Simkin & Conti-Ramsden, 2006) and psychosocial development (Clegg, Hollis, Mawhood, & Rutter, 2005). Because of the impact of language impairment on child development, the general consensus is that speech and language problems should be identified and treated at an early age. In fact, it has been shown that it is possible to remedy the language difficulties of young children with a highly structured parent-based language intervention group program (Buschmann et al., 2009; Ward, 1999). Furthermore, early language oriented intervention can also prevent the development of learning difficulties (Gillon, 2000) and social-emotional problems (Robertson & Ellis Weismer, 1999). However, the downside of this focus on the treatment of language disorders in young children is that, to date, research on effective treatment for older children with SLI (i.e., late elementary school) has remained scarce. Moreover, the existing evidence of effective intervention for older children has not been disseminated extensively (Bishop, 2009). Possibly, the unfamiliarity with this evidence has led to the general assumption that speech therapy for older children with SLI probably is not very effective anymore. The prevailing idea is that older children with SLI will have to learn to cope with their language problems. The current tendency in many countries is to even further reduce therapy resources for older children with SLI.

If older children with SLI receive treatment, it usually focuses on the acquisition of compensation strategies, rather than remediation of grammar and vocabulary. Of course, learning to cope with a language disorder and making use of compensation strategies are important and valuable skills for older children with SLI. However, impoverished grammatical abilities impede effective communication, and this remains the case in older school-age children with SLI. Addressing grammatical skills in language intervention would therefore seem useful also for this age group. To date, the development of intervention techniques and materials for this age group has remained limited. As a consequence, evidence for effective treatment for older children has remained sparse, which in turn leads to a further decline in therapeutic options for these children. The current situation can only change if more evidence for the effectiveness of intervention for older children is produced. The present study aims to contribute to this body of evidence by investigating a multimodal and metalinguistic intervention approach in Dutch children with SLI from nine years upwards.

Metalinguistic intervention: rationale and effectiveness

In the past two decades, several studies have provided evidence for the effectiveness of intervention for older children and adolescents with SLI (Ebbels, 2007; Ebbels & van der Lely, 2001; Ebbels, van der Lely, & Dockrell, 2007; Hirschman, 2000; Levy and Friedman, 2009). These studies focused on the remediation of morphosyntax and complex syntax (i.e., complex sentences) by using a metalinguistic approach, and thus differed considerably from the interventions commonly used with younger children. Therapy approaches for young (pre-school) children mostly train morphosyntactic skills implicitly, using techniques like recasting, elicited imitation and modelling (Fey & Proctor-Williams, 2000). The idea is that young children with SLI are able to learn the selected grammatical targets when they are presented with an increased frequency and salience. However, difficulties with morphosyntax and complex syntax often persist in older school-age children with SLI. For these children, speech therapy using recasting, imitation and modelling techniques has not succeeded in remediating these persisting grammatical problems.

Metalinguistic intervention sets out to remediate problems in morphosyntax and complex syntax by the explicit teaching of grammatical rules. The rationale for metalinguistic intervention is the notion that older children with SLI have attained a level of cognitive development and language experience that enables them to reflect on language. Furthermore, they usually have acquired reading and writing skills. Both these metalinguistic and academic skills can be utilized to teach grammatical rules explicitly. A hallmark of the existing metalinguistic intervention practices is the presence of some form of visual support, such as writing, shapes or colors. Spoken language is very transitory and a visual representation offers children with SLI compensation for their limited capacity to process and store verbal information (Bishop, 1992, 1994; Ellis Weismer & Evans, 2002). Children with SLI are generally faced with substantial phonological short-term memory difficulties (Gathercole & Baddeley, 1990, 1993; Montgomery, 2000, 2003), but visuo-spatial short-term memory appears to be an area of relative strength (Henry, Messer & Nash, 2012). There is evidence that the provision of visual information can support working memory performance in children with SLI (Quail, Williams & Leitão, 2009).

The literature on metalinguistic intervention approaches ranges from case studies to group studies in various clinical settings and using varied research designs. Studies on metalinguistic intervention for older children with SLI report favorable results for a wide array of different grammatical targets. For instance, metalinguistic classroom therapy during one school year, directed at improving the use of complex sentences and utilizing written language and illustrations, improved the use of complex sentences in oral and written language of 9- and 10-year-old children with SLI (Hirschman, 2000). In this study, the children who performed most poorly on complex sentences turned out to be the ones that benefited the most from therapy. The efficacy of the Shape Coding system, which uses two-dimensional shapes, colors and arrows to identify phrases, words, and morphology (Bryan, 1997; Ebbels, 2007; Lea, 1970) was investigated in a number of different studies. The children improved in their use of passives and wh-questions (Ebbels & van der Lely, 2001), the comprehension of dative forms and comparative questions, as well as the use of past tense in written work (Ebbels, 2007). In a randomized controlled trial, argument structure was targeted (Ebbels, van der Lely & Dockrell, 2007). This study compared three different types of intervention (i.e., Shape Coding, verb semantic therapy and a usual care control therapy). Both the Shape Coding approach and the semantic therapy yielded significant progress, which was also maintained after three months. Finally, Levy and Friedman (2009) reported significant improvement in a case study of a 12-year-old boy with severe SLI. Written and oral exercises were used to teach syntactic movement explicitly in various syntactic structures, such as relative clauses, object questions, topicalization sentences, and sentences with verb movement. Colors depicting the verb argument structure were used as support. In this study, progress was maintained 10 months post-treatment.

In a recent study, Bolderson, Dosanjh, Milligan, Pring and Chiat (2011) investigated whether younger children with SLI could also benefit from metalinguistic intervention. Results showed that an intervention program based on Colorful Semantics (Bryan, 1997) significantly improved verb argument structure in 5- and 6-year-old children with SLI.

In conclusion, although to this day, only one study investigated the effectiveness of metalinguistic therapy with a randomized controlled trial design (Ebbels, van der Lely & Dockrell, 2007), evidence is accumulating that metalinguistic intervention can further grammatical skills in both younger and older school-age children with SLI. An important asset of metalinguistic therapy methods, from a cost-effectiveness point of view, is that positive results apparently can be obtained with a relatively limited quantity of therapy.

Multimodal learning

The metalinguistic approaches discussed so far, all used two-dimensional visual support systems consisting of different shapes and colors. Written language also formed an important component of the interventions used with older children with SLI. However, not all children with SLI have well-developed reading skills, and employing reading activities in therapy may place an extra burden on these children.

For the present study, we searched for a more dynamic, multimodal approach that would not only make use of the visual and auditory channels, but that would also employ the tactile/kinesthetic and motor systems. Sensory-motor engagement raises the level of active participation in the children and may also enhance their enjoyment and motivation for therapy. An intervention approach extended with a tactile/kinesthetic and motor component could effectively support language learning. According to Shams and Seitz (2008), the human brain has evolved to develop, learn and operate optimally in multimodal environments. A multimodal training approach would therefore be more effective for learning. Multimodal training approaches are not new at all, and were introduced in regular education about a century ago by Montessori (1912, 1949). In fact, what Montessori achieved through careful observation of child development is now being substantiated by modern research. The processing of information entering the cognitive system through multiple channels helps to circumvent the limited processing capabilities of the individual channels (Clark & Paivio, 1991). Consequently, a greater total of information can be processed when shared between multiple senses (Birsh, 2005).

A multimodal and metalinguistic intervention approach in which language is represented by manipulable items was found in the 'Grammar in Form and Color' program. This method was originally developed for Danish by Thyme to treat grammatical problems in children with severe hearing problems (K. Thyme, personal communication, December 8, 2010). 'Grammar in Form and Color' uses Lego® bricks as abstract representations of grammatical structures. Every word class is represented by a brick in a specific color. Function words (e.g. determiners, prepositions) are represented by smaller bricks, and content words (e.g. nouns and verbs) by larger ones. With a limited set of colors and shapes all targets related to morphosyntax and grammatical complexity can be expressed. For instance, plural forms of nouns are depicted by piling up two bricks, and subject-verb agreement is represented by giving the noun and the inflected verb an equal number of bricks. Written language can be used as additional support, but it does not constitute a necessary part of the program. By manipulating the bricks and laying them out sequentially on a base plate, children actually learn to (literally) build and expand sentences. Thus, in addition to the auditory and visual channels, tactile/kinesthetic and motor channels are also deployed in the explicit learning of grammatical rules.

Thyme's intervention program was never published, but was adapted for Dutch by van Geel (1973). Unfortunately, interest in the approach gradually waned. One of the reasons might be the absence of systematic studies on its effectiveness. Furthermore, the program had not attained its full development. Although the system was already rather elaborate, it lacked a system for more complex syntax, such as codes for subordinate clauses and their complementizers.

Relative clause use was selected as the intervention target for this study. Therefore, the existing program was expanded with new codes to represent relative clauses and relative pronouns. Because of the changes to the original intervention method, the new version was renamed 'Metataal', a compound of the Dutch words for 'metalinguistic' and 'language'.

Relative clauses and (Dutch) SLI

The comprehension and production of relative clauses still poses difficulties for older school-age children with SLI (Friedmann & Novogrodsky, 2004; Håkansson & Hansson, 2000; Novogrodsky & Friedmann, 2006; Schuele & Nicholls, 2000; Stavrakaki, 2001). Several explanations have been proposed to account for this phenomenon. From a nativist perspective, it is claimed that the children have a deficit in grammatical knowledge (Van der Lely, Rosen, & McClelland, 1998). The children with SLI construct grammars that fail to represent long-distance dependencies. Alternatively, from a processing point of view, Hestvik, Schwartz, and Tornyova (2010) argued that the production of relative clauses in children with SLI is hampered by factors such as sentence length and verbal working memory demands. In their study, children with SLI did perform poorly on an automatic on-line gap-filling task, but reached normal levels on a relative clause comprehension task. This finding is not compatible with a deficient grammatical knowledge account. According to Hestvik et al. (2010), problems in relative clause production of children with SLI therefore should be contributed to impaired processing mechanisms.

In order to appreciate the difficulties that children with SLI experience with relative clauses, some elaboration on (Dutch) relative clauses is appropriate at this point. A relative clause can be defined as a subordinate clause that modifies a (head) noun or a noun phrase in an adjacent main clause. The classification of relative clauses is based on two structural features. The first feature is the syntactic role of the head, i.e., the main clause element that is modified by the relative clause, which mostly is either subject (S) or object (O). The second feature is the syntactic role of the gap, i.e., the element that is gapped or relativized inside the relative clause, which again is the subject or object of the relative clause. The taxonomy put forward by Goodluck and Tavakolian (1982) has been widely used in studies on relative clauses and involves 4 different relative clause types. The sentences in (1) to (4) give an example of each type.

(1) The boy (S) [that (S) drank the juice] fell ill. (SS)
(2) Mary kissed the boy (O) [that (S) brought the flowers]. (OS)
(3) The boy (S) [*that* (O) John kicked] ran away. (SO)
(4) John kicked the boy (O) [*that* (O) Mary knew]. (OO)

The sentences in (1) and (2) can be classified as subject relative clauses, because the relative pronoun takes the subject role. Sentences in the examples (3) and (4) are usually defined as object relative clauses, because the relative pronoun takes the object role. Relative clauses can also be classified on the basis of their position within the sentence, and can either be center-embedded (examples 1 and 3) or right-branching (examples 2 and 4). In the center-embedded versions, there is a distance (non-local dependency) between the subject and the finite verb of the main clause. This distance is dependent on the length of the relative clause. From a processing view, sentences containing (longer) center-embedded clauses are harder to process because of these non-local dependencies.

Dutch and English relative clauses are different in a number of grammatical aspects. A first difference concerns the relative pronoun. In English, the relative pronoun in object relative clauses (in italics in examples 3 and 4) is optional, whereas in Dutch realization of the relative pronoun is obligatory. Furthermore, in Dutch relative clauses gender agreement between the head noun and the relative pronoun is required. The relative pronoun can either take common gender form *'die'*, or neuter gender form *'dat'*. A second difference between Dutch and English relative clauses relates to their verb placement requirements. Dutch is a so-called SOV + verb-second language. The finite verb always takes second position in main clauses, and in subordinate (relative) clauses the finite verb appears clause final. A third contrast is that, because of these verb placement requirements, Dutch relative clauses with animate subjects and objects cannot be disambiguated by word order, as is the case in English. Such clauses remain ambiguous between a subject relative clause reading and an object relative clause reading, as can be seen in example (5). Only when the animate subject and object differ in number, these clauses are syntactically disambiguated by subject-verb agreement (example 6).

(5) Het konijntje SING (S), dat (S or O) de jager SING (S or O) ziet SING, zit in het gras.
The rabbit (S), that (S) sees the hunter, sits in the grass. (SS), *or:*
The rabbit (S), (that) (O) the hunter sees, sits in the grass. (SO)
(6) Het konijntje SING (S), dat (O) de jagers PLU (S) zien PLU, zit in het gras.
The rabbit (that) the hunters see, sits in the grass.

To sum up, (Dutch) relative clauses can be regarded as complex grammatical structures. Characteristics such as (1) variety in types of relative clauses; (2) gender agreement between relative pronoun and the relativized noun; (3) different verb placement requirements in main and subordinate clauses; and (4) non-local dependencies belonging to center-embedded versions, all may contribute to a late acquisition and infrequent use of relative clauses in (Dutch) children with SLI.

The present study

In view of their complexity, as outlined above, targeting relative clauses in a intervention study is challenging. If we should find that school-aged children with SLI can learn to use and understand relative clauses through the 'Metataal' program, there will be all the more reason to adopt this approach for older children with persistent language difficulties. Our hypothesis is that metalinguistic therapy using the multimodal 'Metataal' approach will improve the use of relative clauses in older school-age children with SLI. To test our hypothesis, a concurrent within-subjects multiple-baseline design was chosen. In this quasi-experimental design, all participants undergo treatment in the same general time period. An advantage of this approach is that it moderates several threats to validity, history effects in particular (Carr, 2005; Harris & Jenson, 1985). Another advantage of a concurrent multiple-baseline design is that it saves time. The amount of time that can be spent on speech therapy with the older children in special schools is limited. The concurrent multiple-baseline design enabled us to obtain a complete data set within well-defined time constraints. To control as much as possible for extra-experimental variables (e.g. different speech therapists involved, therapy settings, school curriculum differences, and period in the school year) we chose to confine our study to just one school. This decision limited the number of children that could participate, and consequently ruled out the inclusion of a control group.

However, in this multiple-baseline design, the participants can be regarded to act as their own controls. During three subsequent monthly measurements prior to the treatment, the children are tested on criterion-referenced tasks, especially constructed for this study to test relative clause production and comprehension. If it is found that the children do not improve in their use of relative clauses during this baseline period, an increase in scores post-therapy should then be attributed to the 'Metataal' intervention. Furthermore, maintained improvement at a retention measurement administered 3 months later without providing any further intervention, would indicate that the children had effectively improved on relative clause use. Nevertheless, this quasi-experimental design does not meet the requirements of a randomized controlled trial, and should therefore be regarded as a Phase 2 early efficacy study (Fey & Finestack, 2009).

Method

Participants

A total of 13 monolingual Dutch speaking children with SLI were included. Informed written consent was obtained from the parents of all participating children. This study was reviewed and approved the Dutch Central Committee on Research Involving Human Subjects (CCMO; *Centrale Commissie Mensgebonden Onderzoek*). The descriptives of the participants can be found in Table 1. All children were enrolled in a special school for children with language impairments. They had non-verbal cognitive abilities within normal limits and no diagnosed comorbidity such as ADHD or autism spectrum disorders. The children had a mean age of 11;2 years (standard deviation 1;1 years, age range 9;3 – 12;8 years). Around this age the children supposedly have acquired sufficient reading skills, and their metalinguistic awareness is developed to such an extent that they are able to reflect on grammatical features of language. It must be noted that one child dropped out after three therapy sessions because of poor cognitive skills. Although at the initial selection, her non-verbal IQ was within the normal range, on retesting a nonverbal IQ of 76 was scored. During therapy, she was not able to make the abstraction from words to Lego® bricks. Consequently, 12 children (8 boys and 4 girls) completed all measurements and all sessions of the intervention program.

Table 1: *Means, Standard Deviations and Range of Age, and Quotient Scores of the Wechsler Non Verbal, PPVT-NL and CELF-4-NL Subtests.*

	Mean	*SD*	*Range*
Age (in months)	134.75	13.72	112-154
Wechsler Non Verbal-NL	99.64	10.31	85-116
Peabody Picture Vocabulary Test-NL	88.09	9.70	71-102
CELF Number Repetition Total	82.92	20.60	55-120
CELF Formulated Sentences	83.73	7.99	75-95
CELF Recalling Sentences	71.64	12.19	55-96
CELF Word Definitions	77.75	11.78	55-90
CELF Word classes	84.45	13.48	63-110
CELF Understanding Spoken Paragraphs	91.82	8.15	80-105

Procedure

Comprehensive manuals and protocols were constructed for the assessments and therapy program, which were followed strictly. The assessment of 6 children in the first round of the baseline measurements was carried out by the speech therapist who would also deliver the intervention. The remaining 7 children in the first round of testing and all children in the 4 subsequent rounds were tested by two research assistants. These assistants were experienced clinical linguists who were not involved in either the school or the intervention program. The assistants scored the tasks and scores were checked and analyzed afterwards by the experimenter. During the whole testing and treatment period, the speech therapist delivering the intervention remained blind to the test results. After the intervention program was completed, the children did not receive any further speech/language therapy. The retention assessment was carried out 12 weeks after the last therapy session, again by a research assistant. Furthermore, relative clauses were not part of the school curriculum for the duration of the intervention study.

Materials

Description and scoring of the relative clause tasks

Standardized language tests could not be used to measure the progress of relative clause use, because these norm-referenced tests do not contain enough items assessing relative clauses and therefore are not sensitive enough. Therefore, specific tasks were constructed especially to assess the relative clause types that were targeted in this intervention study. The six tasks that were developed were always administered in the same order. The three production tasks were presented first, followed by the two sentence repetition tasks and the comprehension task. All tasks contained two or three practice items depending on the difficulty of the task. For all tasks, two balanced A and B versions were constructed, which were alternated, in order to suppress a possible learning effect from multiple presentations of the tasks. Administration of the tasks lasted 30-45 minutes. The six tasks are described in detail below. Each examples of the tasks starts with the instruction of the investigator (IN), followed by the target response (TR) expected from the children, in italics.

Production task 1: Right-branching OS relative clause completion (10 items).

In this elicited production task, adapted from Novogrodsky and Friedmann (2006), the children had to complete sentences after a prompt was presented. Test items were adjusted to participant gender (example 7). Responses were scored as correct when a relative clause pattern was produced. The relative clause had to contain a relative pronoun and an inflected verb in clause final position.

(7) IN: Eén jongen (meisje) koopt een lolly en één jongen koopt een ijsje.
Welk(e) jongen (meisje) zou jij willen zijn?
Begin je antwoord met: Ik kies de jongen (het meisje) ….
TR: *die (dat) de lolly koopt.*
IN: One boy (girl) buys a lollipop and one boy (girl) buys an ice-cream.
Which boy (girl) would you like to be?
Start your answer with: I choose the boy (the girl)…..
TR: *who the lollipop buys.* ("I choose the boy who buys the lollipop").

Production task 2: Center-embedded SS relative clause production (10 items).
Because center-embedded relative clauses are difficult to elicit, this task contained test items with three written short sentences (example 8). From these three short sentences the children had to construct a sentence with an embedded SS relative clause. The sentences were first read aloud together by the child and the research assistant, who pointed at the words. The children were instructed to use all bold printed words (verbs) and not to use the coordinate complementizer '*en*' (and). Responses were scored as correct when a center-embedded relative clause pattern was produced. The relative clause had to contain a relative pronoun and an inflected verb in clause final position.

(8) IN: Er is een chauffeur. / Hij **draagt** een hoed. / Hij **zit** in de auto.
Begin je zin met: De chauffeur….
TR: *die een hoed draagt, zit in de auto.*
IN: There is a driver. / He **wears** a hat. / He **sits** in the car.
Start your sentence with: The driver…
TR: *who a hat wears, sits in the car.*
("The driver who wears a hat, sits in the car").

Production task 3: Recreating OS and SS relative clauses (16 items).
This task was based on the subtest 'Sentence Assembly' from the Dutch version of the CELF-4 (Semel, Wiig, & Secord, 2008). The different parts of the sentences were visually presented in frames in a scrambled and ungrammatical order (example 9). The test items were read aloud by the child and investigator. The children had to reconstruct the sentences and were instructed to start with the underlined word(s) which also had a capital letter. Responses were scored as

correct when the children produced a grammatical sentence containing a relative clause pattern.

(9) | Stefanie | dorst | drinkt water | die | heeft |

IN: Begin je zin met: Stefanie….
TR: *die dorst heeft, drinkt water.*
IN: Start your sentence with: Stefanie …
TR: *who thirst has, drinks water* ("Stefanie who is thirsty, drinks water").

Sentence repetition (task 4: 12 items, and task 5: 15 items).
Two sets of sentences containing both OS and SS type relative clauses had to be repeated by the children. Working memory limitations are often seen in children with SLI. In order to investigate the role of sentence length in a repetition task, two different sets of sentences with OS and SS type relative clauses were constructed. The first set contained twelve 7-word sentences and the second set contained fifteen 12-word sentences. The examples (10) and (11) show the two relative clause types presented in the 7-word condition. Responses were scored as correct when a relative clause pattern was produced. The relative clause had to contain a relative pronoun and an inflected verb in clause final position.

(10) Hij ziet een vrouw die cake eet. (7 words, OS type)
He sees a woman who cake eats. ("He sees a woman who eats cake".)

(11) De clown, die ballonnen opblaast, ziet Marloes. (7 words, SS type)
The clown, who balloons blows up, sees Marloes. ("The clown who blows up balloons, sees Marloes".)

Comprehension of relative clauses (task 6: 28 items).
In this picture selection task, sentences containing SS and OS relative clauses were presented to the participants. The children had to select the correct picture from a set of four pictures. The images were full color photographs of scenes with Playmobil® material. Figure 1 shows one of the test items. The accompanying test item of an embedded SS type relative clause is shown in example (12). Responses were scored as correct when the child selected the correct picture.

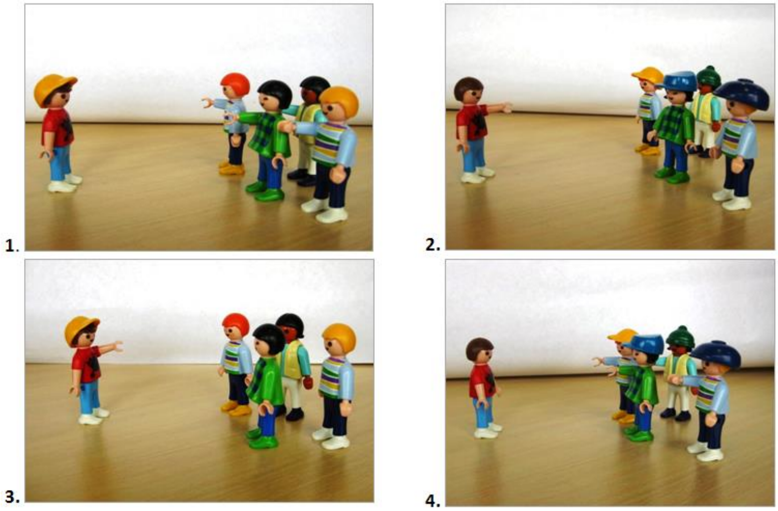

Figure 1. Example of one test item from the relative clause comprehension task.

(12) De jongen, die naar de kinderen wijst, draagt een pet.
The boy, who to the children points, wears a cap.
("The boy who points to the children wears a cap").

Intervention program

The children received a total of 10 individual therapy sessions lasting 30 minutes each, twice a week during 5 weeks. All therapy sessions were protocolled in a therapy manual which was followed strictly. First, the children were introduced to the concept that words could be represented by Lego® bricks in different shapes, sizes and colors. The children also practiced with exercises aimed at the identification of different clause types and conjunctions in written and spoken language. When the concepts and procedures of the intervention program were understood, the children started building simple declarative sentences consisting of a subject, verb and object. Subsequently, more elements such as prepositional phrases, plural nouns and subject-verb agreement were introduced. The next step was to build coordinated sentences by connecting two main clauses with an arched brick (a bridge). An example of a coordinated sentence built with Lego® bricks can be seen in Figure 2. Next, the children learned that this sentence could be shortened by deleting the superfluous subject in the second main clause. This operation resulted in a sentence containing coordination with reduction, as is illustrated in Figure 3. Obviously, the terms 'coordination' and 'subordination' were not used with the children. Instead, these sentences were called *'bruggetjeszinnen'* (bridge sentences). When the children performed well on building coordinated sentences, right-branching (OS) subject relative clauses were introduced. As can be seen in Figure 4, these relative clauses were built on the base plate on a lower level than the main clause. The bridge was used to connect the main clause with the relative clause. The last step in the intervention program was reserved for building center-embedded (SS) subject relative clauses. Two bridges were used to express subordination and center-embeddedness. The first bridge connected the main clause to the lower relative clause. A second bridge was needed to connect the relative clause to the second part of the main clause again. This extra bridge was called a 'comma bridge', and was placed between the two inflected verbs of the relative clause and the main clause. An example of a center-embedded (SS) type relative clause can be seen in Figure 5. The 'comma bridge' was depicted by a small corner brick on top of the arched brick. The children easily understand this concept, because in running speech it is natural to pause at that point in the sentence, and in written text a comma is often placed between two inflected verbs.

During intervention, the children did not have to memorize the functions of the different colors, bricks and bridges. They always had a crib sheet with pictures of all the bricks and examples of their functions at hand. Intervention was mostly oral, although minor reading and writing activities were included in the intervention. While building their sentences, the children could add, move and remove bricks. A new sentence could also be constructed on the base plate below the previous one, so the two sentences could be compared. Another option was to check possible differences between sentences built by the child and therapist. Conversations, short stories and pictures were used to elicit relative clauses during therapy. Each therapy session also had a game activity to consolidate the target sentences. The children did not practice with relative clauses outside of the therapy sessions and no homework was given.

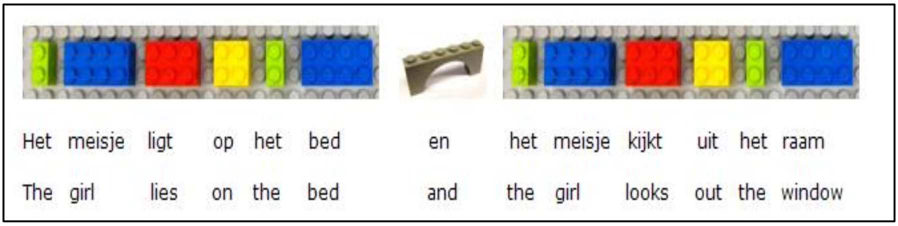

Figure 2. Coordinated sentences built with Lego® bricks.

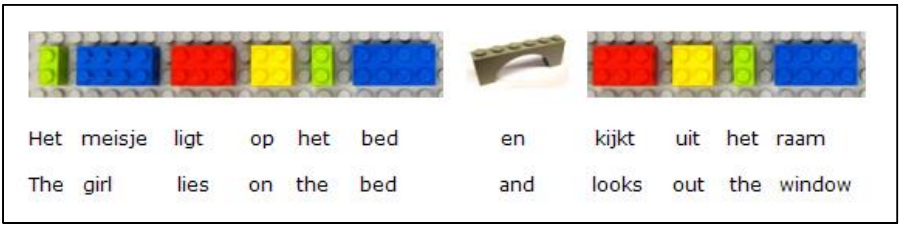

Figure 3. Coordination with reduction sentence built with Lego® bricks.

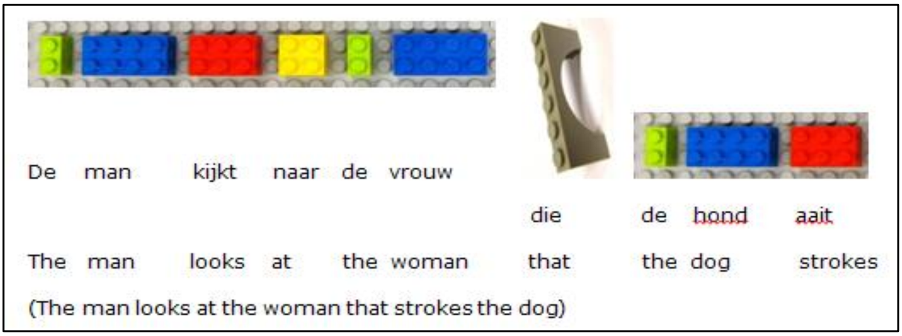

Figure 4. Sentence built with Lego® bricks containing a right-branching OS type relative clause.

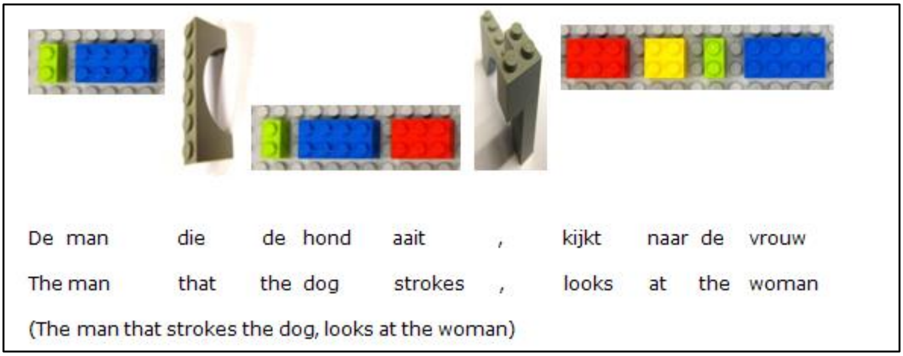

Figure 5. Sentence built with Lego® bricks containing a center-embedded SS type relative clause.

Results

The results (scores and significant differences between measurements) on the repeated baseline, post-therapy and retention measurements of the 6 relative clause tasks are presented in Table 2. Because the participant group was small and the data violated the assumptions for a GLM analyses on repeated measures, non-parametric tests were used. The differences between the five measurement points were analyzed with Friedman's non-parametric ANOVA, using the exact statistic. When Friedman's tests reached significance, subsequent post-hoc Wilcoxon signed rank tests were used. A Bonferroni correction was applied for all post-hoc tests, as a result of which, all effects are reported at a .0125 level of significance. The effect size *r* was computed by dividing the *Z*-score by the square root of the total observations (Field, 2009). An effect size *r* from .10 to .30 represents a small effect, from .30 to .50 a medium effect and beyond .50 a large effect.

Table 2: *Mean and Standard Deviations of the Raw Scores and Maximum Score (Max Score) on the Relative Clause Tasks at the 5 Measurement Points (T1-T5).*

RC tasks	Max Score	Pre-therapy			Post-therapy	Retention
		T1 *M* *(SD)*	T2 *M* *(SD)*	T3 *M* *(SD)*	T4 *M* *(SD)*	T5 *M* *(SD)*
			Relative clause production			
OS Completion	10	6.25 (2.05)	7.83 (2.17)	8.08 (1.51)	9.42** (0.90)	8.75 (1.14)
SS Composition	10	1.75 (3.25)	3.00 (3.86)	3.00 (4.00)	7.08** (4.23)	7.00 (4.20)
Recreation	16	6.50 (4.23)	8.00 (4.05)	9.25 (4.05)	11.42** (3.50)	12.33 (3.63)
			Relative clause repetition			
Repetition 7 word	12	11.42 (1.17)	10.92 (1.31)	11.25 (1.22)	11.00 (1.13)	11.50 (0.80)
Repetition 12 word	15	9.25 (3.72)	9.58 (4.85)	9.75 (3.96)	10.75 (3.30)	11.83 (3.46)
			Relative clause comprehension			
Comprehension	28	19.25 (2.30)	18.25 (2.90)	21.17 (2.37)	20.17 (1.75)	21.67 (2.19)

Note. ***p* < .01 between T3 (last baseline measurement pre-therapy) and T4 (measurement directly post-therapy)

Regarding the *OS Completion task*, analysis with Friedman's ANOVA showed that the scores of the children between baseline measurements, pre- and post-therapy tests and the final retention measurement did change significantly ($\chi^2(4) = 20.72$ and $p < .001$). Post-hoc Wilcoxon tests indicated that no significant difference was found between T1 (*Median (Mdn)* = 6.50) and T2 (*Mdn* = 8.50) ($Z = -1.86$, $p = .070$, and $r = -0.38$). This was also the case between T2 (*Mdn* = 8.50) and T3 (*Mdn* = 8.00) ($Z = -.23$, $p = .855$, and $r = -0.05$). However, post-therapy the difference was significant between T3 (*Mdn* = 8.00) and T4 (*Mdn* = 10.00) with $Z = -2.41$, $p = .008$, and $r = -0.49$. The difference in scores post-therapy at T4 (*Mdn* = 10.00) and the retention measurement at T5 (*Mdn* = 9.00) was not significant ($Z = -1.93$, $p = .047$ and $r = -0.39$). In summary, the children did not improve between the three baseline measurement points, but directly post-therapy a significant gain was found, which was maintained 12 weeks later.

With respect to the *SS Composition task*, Friedman's ANOVA yielded a significant difference between measurements, $\chi^2(4) = 31.40$, $p < .001$. Post-hoc Wilcoxon tests indicated that no significant difference was found between T1 (*Mdn* = .00) and T2 (*Mdn* = .50) ($Z = -1.68$, $p = .125$, $r = -0.34$). The same result was found between T2 (*Mdn* = .50) and T3 (*Mdn* = .00) ($Z = .00$, $p = 1.00$ and $r = -0.00$). Post-therapy the difference was significant between T3 (*Mdn* = .00) and T4 (*Mdn* = 9.50) with $Z = -2.83$, $p = .002$ and $r = -0.58$. The difference in scores post-therapy (T4, *Mdn* = 9.50) and the retention measurement (T5, *Mdn* = 9.50) was not significant, with $Z = -.264$, $p = .984$ and $r = -0.05$. The children did not improve between baseline measurements, but post-therapy a significant gain was found that was maintained at retention measurement.

On the relative clause *Recreation task*, Friedman's ANOVA yielded significant differences over time, $\chi^2(4) = 34.40$, $p < .001$. Post-hoc Wilcoxon tests indicated that no significant difference was found between T1 (*Mdn* = 7.00) and T2 (*Mdn* = 7.50) ($Z = -1.58$, $p = .134$, and $r = -0.32$) and neither between T2 (*Mdn* = 7.50) and T3 (*Mdn* = 10.00) ($Z = -1.36$, $p = .203$, and $r = -0.27$). Post-therapy, the difference was significant between T3 (*Mdn* = 10.00) and T4 (*Mdn* = 11.50) with $Z = -2.67$, $p = .006$ and $r = -0.54$. The difference in scores directly post-therapy (T4, *Mdn* = 11.50) and the retention measurement (T5, *Mdn* = 12.00) was not significant ($Z = -1.72$, $p = .057$, and $r = -0.35$). On this third production task the children made significant gains post-therapy, which were maintained 12 weeks later.

The two sentence repetition tasks yielded different results. On the *Repetition task (7 words)*, analysis with Friedman's ANOVA showed that the children obtained significantly different scores over time, $\chi^2(4) = 13.08$, $p = .006$.

Post-hoc no significant difference was found between T1 (*Mdn* = 12.00) and T2 (*Mdn* = 11.50) (Z = -1.51, p = .250, and r = -0.31). Also no significant difference was found between T2 (*Mdn* = 11.50) and T3 (*Mdn* = 12.00) (Z = -2.00, p = .125, and r = -0.41). Post-therapy the difference was not significant between T3 (*Mdn* = 12.00) and T4 (*Mdn* = 11.00) (Z = -1.73, p = .250, r = -0.35). The difference between post-therapy scores (T4, *Mdn* = 11.00) and retention measurement (T5, *Mdn* = 12.00) was also not significant, with Z = -2.45, p = .031, and r = -0.50. On the *Repetition (12 word) task*, Friedman's ANOVA yielded significant differences, $\chi^2(4)$ = 17.51, p = .001. Post-hoc Wilcoxon tests indicated that no significant difference was found between T1 (*Mdn* = 11.00) and T2 (*Mdn* = 12.00) (Z = -.54, p = .652, and r = -0.11). The same result was found between T2 (*Mdn* = 12.00) and T3 (*Mdn* = 11.00) (Z = -.30, p = .848, and r = -0.06). Post-therapy the difference was not significant between T3 (*Mdn* = 11.00), and T4 (*Mdn* = 10.50) (Z = -2.16, p = .047, r = -0.44). The difference in scores directly post-therapy (T4, *Mdn* = 10.50) and the retention measurement (T5, *Mdn* = 13.00) was also not significant, with Z = -2.14, p = .043, and r = -0.44. Although the initial Friedman's ANOVA for both sentence repetition tasks was significant, no significant changes were found between the measurement points we were interested in.

With respect to the *Comprehension task*, Friedman's ANOVA showed a significant difference over time, $\chi^2(4)$ = 15.16, p = .002. Post-hoc Wilcoxon tests indicated that no significant difference was found between T1 (*Mdn* = 19.50) and T2 (*Mdn* = 19.00) (Z = -1.28, p = .236, and r = -0.26). The difference was also not significant between T2 (*Mdn* = 19.00) and T3 (*Mdn* = 21.00) (Z = -2.32, p = .018, and r = -0.47), and neither between T3 (*Mdn* = 21.00), and T4 (*Mdn* = 20.50) (Z = -1.01, p = .355, and r = -0.21). The difference between post-therapy scores (T4, *Mdn* = 20.50) and the retention measurement (T5, *Mdn* = 22.00) was also not significant, with Z = -1.86, p = .067, and r = -0.38. Although Friedman's ANOVA for the comprehension task was significant, no significant differences were found between the measurement points relevant for the present study.

Discussion

The purpose of this study was to examine the efficacy of the metalinguistic 'Metataal' approach for older school-age children with SLI. This intervention study made use of a repeated-baseline design and targeted the production of OS and SS type subject relative clauses. Relative clauses were chosen because older children with SLI still experience difficulties with these types of complex sentences. Furthermore, if the children would improve on such difficult and complex structures, the intervention approach would also be suitable to remediate other, less complex targets. Our prediction was that the children would not improve during the baseline measurements, but that a significant increase in scores would be observed post-therapy. We also predicted that the significant gains would be maintained at a retention measurement 12 weeks later.

In the present study, our predictions were partly confirmed. Post-therapy, improvement was not observed for all five administered relative clause tasks. The results indicate that the 5-week intervention program was indeed effective for the *production* of relative clauses. As predicted, no improvement was found on any of the tasks during the three baseline measurements, but immediately post-therapy, significant gains were found on all three relative clause production tasks. Moreover, this progress was maintained 12 weeks later. Contrary to our prediction, the children did not improve on the two relative clause repetition tasks and the comprehension task.

Although we did not have a control group in our study, it is quite plausible that the improvement in relative clause production is an effect of the 5-week intervention. As no significant improvement was seen during the three baseline measurements, the significant gains between the pre- and post-therapy measurements cannot be attributed to a test-retest effect due to a repeated administration of the tasks. This conclusion is strengthened further by the fact that also no significant differences were found between the post-therapy and retention measurements. Another argument for our conclusion that the intervention accounted for the improvement is the fact that the children specifically advanced on the production tasks only, and relative clause production was indeed the substance of the intervention program. Finally, if the improvement of the children would have been due to a test-retest effect, we would expect to see this improvement on *all* the different relative clause tasks, which clearly was not the case.

Concerning the sentence repetition tasks, at first measurement most children already performed at ceiling on the 7-word task. This task proved to be too easy. The children were able to repeat the target relative clauses as long as the sentences were short, although they did make omission and substitution errors. On the other hand, the 12-word sentence repetition task turned out to be very difficult. Although the children often were not able to fulfil the minimum requirements (i.e., realization of a relative pronoun and correct verb placement in the relative clause), they mostly did succeed in correctly conveying the meaning of the sentences. For instance, a phrase such as "the clown that has a balloon" was changed into "the clown with a balloon". Apparently, the children were able to process the sentences correctly, but failed to reconstruct the target sentences. This result fits in with the very poor results on the CELF-4-NL Recalling Sentences subtest (see Table 1) in this study. We can conclude that five hours of therapy targeted at relative clause production in children with SLI certainly did not improve their sentence repetition abilities. Although a sentence repetition task has considerable diagnostic value in identifying language disorders (Hesketh, Riches, & Vance, 2012; Seeff-Gabriel, Chiat, & Dodd, 2010), such a task apparently is not a suitable instrument to measure therapy gains in an intervention study.

The absence of significant gains on the sentence repetition tasks was also found for the relative clause comprehension task. This result is not that surprising, because the intervention program was mainly oriented towards remediation of relative clause production. During therapy, the semantics of relative clauses (i.e., giving extra information about a noun) did receive some attention, but most time was devoted to production training. This result contrasts with Camarata, Nelson, Gillum, and Camarata (2009), who found that children with SLI made significant gains in language comprehension when exposed to language intervention focused on production. Possibly, we did not find such result, because in our study children received only five hours of treatment in five weeks, whereas in the Camarata et al. (2009) study, intervention amounted to 24 hours in 12 weeks. However, we did observe changes in scores on the comprehension task over time. Half of the children obtained lower scores post-therapy than at the last baseline measurement. We interpret this dip in scores as a transitory phase of uncertainty resulting from the intervention. This pattern would fit a U-shaped learning curve. When children acquire new knowledge, they tend to apply that knowledge also in contexts that are an exception to the rule, as a result of which performance temporarily worsens. A well-known example is the past tense overgeneralization of regular verb inflection to irregular verbs. Such patterns are commonly seen in language acquisition, and reflect actual learning behavior. In our study, the observed uncertainty in the relative clause comprehension task had dissolved again at the retention measurement three months later.

A point that is not often discussed in metalinguistic intervention studies is that, although children with SLI have poorer metalinguistic skills than typically developing children (Kamhi & Koenig, 1985; Menyuk, 1993), these skills are in fact employed to remediate grammatical problems in a metalinguistic approach.

It is possible that metalinguistic intervention also enhances the metalinguistic awareness in children with SLI. Metalinguistic abilities were not tested in the present study, but improvement of metalinguistic skills could be a side effect of the intervention. Assessment of metalinguistic skills pre- and post-therapy could be included in future studies in order to answer this question.

In any case, this intervention study has shown that a relatively short training period of five weeks with only five hours of treatment resulted in a significant improvement of relative clause production in older school-age children with SLI. Furthermore, the medium to large effect sizes of the post-hoc comparisons between pre- and post-therapy and retention measurements render these results also clinically relevant. We have no concrete evidence that the children started using more relative clauses in their spontaneous speech. However, the speech therapist observed that during storytelling exercises and in conversation, the children seemed to use more relative clauses than before. The children also started noticing the relative clauses uttered by their teachers. Of course, this evidence remains anecdotal. In a future study the collection and analysis of language samples pre- and post-therapy would be a proper procedure to see whether relative clause use also improves in an unstructured setting.

The motor and tactile/ kinesthetic dimensions of the 'Metataal' approach can be regarded as a valuable addition to the metalinguistic approaches studied thus far. The children thoroughly enjoyed working with the material. Some children wanted to use the bricks in the classroom and asked why they had not started using them much earlier on. The three-dimensional characteristics of the 'Metataal' approach certainly added to their understanding of the grammatical features of the relative clauses. Furthermore, written language can be used together with the Lego® bricks, but this is not necessary. Therefore, the children with SLI that are also poor readers can work with the material as well. Hopefully, the positive results of this study will lead to further research with other grammatical targets, and in different therapeutic settings. The approach may also be effective in a group therapy setting or even in a classroom situation. Other groups of children with language impairment might also benefit from the intervention, such as bilingual children with SLI or children with co-morbidity. A question for further research is what really constitutes the active ingredient in the 'Metataal' program. It is possible that just the extensive exposure or practice with relative clauses was already enough to obtain positive gains. Future studies using a randomized control trial with participant groups assigned to a 'Metataal' condition compared to an extended exposure condition (usual care) would be most informative.

To conclude, this intervention study supports the evidence that grammatical skills in older school-age children with SLI can still be remediated with direct intervention using a metalinguistic approach. Therefore, the current tendency to diminish direct intervention for older children with SLI, as well as shifting treatment goals from grammatical targets and vocabulary to the training of compensation strategies should be reconsidered.

Acknowledgments
We want to thank the parents, children, and staff of Auris De Taalkring for participating in this study. We also want to thank Liset Bergevoet, Hanneke Creemers, Jonna Genuit, and Nadia Lemouesset for their help in constructing the 'Metataal' tests and therapy program.

References

Aram, D. M., Ekelman, B. L., & Nation, J. E. (1984). Pre-schoolers with language disorders: 10 years later. *Journal of Speech and Hearing Research, 27*, 232–244.

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: 1. Developmental and academic outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 804–814.

Birsh, J.R. (2005). *Multisensory teaching of basic language skills (2nd Ed.).* Baltimore MD: Paul H. Brookes Publishing Co.

Bishop, D. V. M. (1992). The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry, 33,* 3-66.

Bishop, D. V. M. (1994). Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics, 15,* 507-550.

Bishop, D. V. M. (2009). Specific language impairment. In C. L. Cooper, J. Field, U. Goswami, R. Jenkins, & B. J. Sahakian (Eds.) *Mental capital and wellbeing* (pp. 767-775). Malden MA: Wiley-Blackwell.

Bolderson, S., Dosanjh, C., Milligan, C., Pring, T., & Chiat, S. (2011). Colourful semantics: A clinical investigation. *Child Language Teaching and Therapy*, *27*, 344–353.

Buschmann, A., Jooss, B., Rupp, A., Feldhusen, F., Pietz, J., & Philippi, H. (2009). Parent based language intervention for 2-year-old children with specific expressive language delay: a randomised controlled trial. *Archives of Disease in Childhood*, *94*, 110–116.

Camarata, S., Nelson, K. E., Gillum, H., & Camarata, M. (2009). Incidental receptive language growth associated with expressive grammar intervention in SLI. *First Language*, *29*, 51–63.

Carr, J. E. (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions*, *20*, 219–224.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210.

Ebbels, S. H. (2007). Teaching grammar to school-aged children with Specific Language Impairment using Shape Coding. *Child Language Teaching and Therapy*, *23,* 67-93.

Ebbels, S. H., & Van der Lely, H. K. J. (2001). Meta-syntactic therapy using visual coding for children with severe persistent SLI. *International Journal of Language and Communication Disorders*, *36 Suppl*, 345–50.

Ebbels, S. H., Van der Lely, H. K. J., & Dockrell, J. E. (2007). Intervention for verb argument structure in children with persistent SLI: A randomized control trial. *Journal of Speech, Language, and Hearing Research*, *50*, 1330–1349.

Ellis Weismer, S., & Evans, J. (2002). The role of processing limitations in early identification of Specific Language Impairment. *Topics in Language Disorders, 22*, 15–29.

Fey, M. E., & Finestack, L. H. (2009). Research and development in children's language intervention: A 5-phase model. In R. G. Schwartz (Ed.) *Handbook of child language disorders* (pp. 513-532). New York: Psychology Press.

Fey, M. E., & Proctor-Williams, K. (2000). Recasting, elicited imitation and modelling in grammar intervention for children with specific language impairment. In D.V.M. Bishop & L. B. Leonard (Eds.) *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp.177-194). Hove: Psychology Press.

Field, A. (2009). *Discovering Statistics Using SPSS. 3rd Edition*. London: Sage.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working Memory and Language*. Hove: Lawrence Erlbaum Associates Ltd.

Gillon, G. T. (2000). The efficacy of phonological awareness intervention for children with spoken language impairment. *Language, Speech, and Hearing Services in Schools*, *31*, 126–141.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, *31*, 661–681.

Håkansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: a comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, *27*, 313–333.

Harris, F. N., & Jenson, W. R. (1985). Comparisons of multiple-baseline across persons designs and AB designs with replication: Issues and confusions. *Behavioral Assessment, 7*, 121-127.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Phonological and visual short-term memory in children with specific language impairment. *Journal of Cognitive Education and Psychology*, *11*, 45-56.

Hesketh, A., Riches, N., & Vance, M. (2012). Sentence repetition in children with communication impairment. Thematic panel presented at the 14th Meeting of the International Clinical Linguistics and Phonetics Association, Cork, Ireland, 2012, June.

Hestvik, A., Schwartz, R. G., & Tornyova, L. (2010). Relative clause gap-filling in children with specific language impairment. *Journal of Psycholinguistic Research*, *39*, 443–56.

Hirschman, M. (2000). Language repair via metalinguistic means. *International Journal of Language and Communication Disorders*, *35*, 251–68.

Johnson, C. J., Beitchman, J. H., Young, A. Escobar, M., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow- up of children with and without speech/language impairments: Speech/ language stability and outcomes. *Journal of Speech, Language, and Hearing Research, 42*, 744–60.

Kamhi, A. G., & Koenig, L. A. (1985). Metalinguistic awareness in normal and language-disordered children. *Language, Speech, and Hearing Services in Schools*, *16*, 199-210.

Lea, J. (1970). *The Colour Pattern Scheme - A Method of Remedial Language Teaching.* Oxted, Surrey: Moor House School.

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA, USA: MIT Press.

Levy, H., & Friedmann, N. (2009). Treatment of syntactic movement in syntactic SLI: A case study. *First Language, 29* 15-49.

Menyuk, P. (1993). Metalinguistic difficulties in children with specific language impairment: Implications for diagnosis and intervention. In Grimm, H. & Skowronek, H. (Eds). *Language acquisition problems and reading disorders: Aspects of diagnosis and intervention.* (pp.3-23). Berlin: De Gruyter.

Montessori, M. (1912). *The Montessori Method.* New York: Frederick A. Stokes Company.

Montessori, M. (1949). *The Absorbent Mind.* Madras: Theosophical Publishing House.

Montgomery, J. W. (2000). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 43*(2), 293–308.

Montgomery, J. W. (2003). Working memory and comprehension in children with specific language impairment: what we know so far. *Journal of Communication Disorders, 36*, 221–231.

Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *International Journal of Speech-Language Pathology*, *8*, 364–375.

Quail, M., Williams, C., & Leitão, S. (2009). Verbal working memory in Specific Language Impairment: The effect of providing visual support. *International Journal of Speech-Language Pathology*, *11*, 220–233.

Robertson, S. B., & Weismer, S. E. (1999). Effects of treatment on linguistic and social skills in toddlers with delayed language development. *Journal of Speech, Language, and Hearing Research, 42*, 1234–1248.

Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language and Communication Disorders, 45*, 691–702.

Schuele, C., & Nicholls, L. (2000). Relative clauses: evidence of continued linguistic vulnerability in children with specific language impairment. *Clinical Linguistics and Phonetics*, *14*, 563-585.

Semel, E., Wiig, E.H., & Secord, W.A. (2008). *CELF-4. Clinical Evaluation of Language Fundamentals. Nederlandse versie (Dutch version).* Amsterdam: Pearson.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*, 411–417.

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language*, *77*, 419–431.

Simkin, Z., & Conti-Ramsden, G. (2006). Evidence of reading difficulty in subgroups of children with specific language impairment. *Child Language Teaching and Therapy, 22*, 315-331.

Stothard, S. E., Snowling, M., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence, *Journal of Speech, Language, and Hearing Research, 41*, 407–18.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*, 1245-1260.

Van der Lely, H, Rosen, S, McClelland, A. (1998). Evidence for a grammar-specific deficit in children. *Current Biology*, 8, 1253–1258.

Van Geel, C. A. A. (1973). *Terreinverkenning van de therapie bij vertraagde spraak- en taalontwikkeling. Toepassing van 'Grammatica in vorm en kleur' (Exploration of therapy for speech and language delay. Application of 'Grammar in form and colour').* Hoensbroek: Studiecentrum Hoensbroeck, Opleiding Logopedie.

Ward, S. (1999). An investigation into the effectiveness of an early intervention method for delayed language development in young children. *International Journal of Language and Communication Disorders, 34*, 243-264.

6 General discussion and conclusions

In this dissertation, the results of a longitudinal study of two age-groups of Dutch-speaking children with SLI and an intervention study examining a metalinguistic approach for older school-age children with SLI are reported. The starting point for this study was that, to date, knowledge on the grammatical development of Dutch school-age children with SLI has remained limited. Furthermore, research on the effectiveness of interventions for older school-age children with SLI has also remained scarce. The general opinion in the Netherlands is that grammatical development in children with SLI beyond age 7 (Grade 3) is very limited and language intervention for these children is unlikely to make a difference. These assumptions are largely based on unsubstantiated notions of critical/sensitive periods for language acquisition in children with SLI. Therefore, the goal of the present study was to gain more insight into the grammatical development of children with SLI between the ages 6 and 10 years and to explore prospects for language intervention in school-age children with SLI, aged 10 and older.

This thesis focused on three main research questions. The first question on the development of grammatical skills of school-age children with SLI was addressed in Chapters 2 and 3. The second question on the relationship between grammatical development and speech disfluencies was addressed in Chapter 4. The third question on the trainability of grammatical skills in older school-age children with SLI was addressed in Chapter 5. In the following sections the answers to these questions will be finally discussed in perspective of a theoretical framework highlighting the debate on delay versus deviance, the role of the lexicon and speech fluency and the effectiveness of metalinguistic and multimodal intervention. Moreover, some clinical implications and suggestions for further research will be presented.

Delay versus deviance

The delay versus deviance issue was addressed in Chapter 2 by examining grammatical development longitudinally in Dutch school-age children with SLI between the ages 6 and 8 years. Grammatical complexity and correctness were analyzed in narratives elicited with a story generation task. In addition, the use of dummy auxiliaries was investigated. Overuse of these dummy verbs has been reported frequently in studies on Dutch SLI (de Jong, 1999; de Jong, Blom &

Orgassa., 2013; Orgassa, 2009) and can be regarded as a sign of delay or immaturity in grammatical development.

The results showed that the children with SLI were outperformed by the age-matched (CA) control group on mean length of utterance (MLU) at all three time points, but on the use of relative clauses only at age 8 years. On the other complexity measures, no group differences were found. The SLI group performed similarly to the LA group on standardized tests and MLU, thus demonstrating a two-year language delay. However, on correctness (i.e., verb-related errors and non-verb related errors), the SLI group was outperformed by both the age-matched (CA) group *and* the language-matched (LA) group at all three time points. This result best fits a delay-within-delay pattern (Rice, 2003), as the level of grammatical correctness in children with SLI does not even reach that of LA controls who are two years younger. Moreover, within the SLI group, grammatical correctness measures indicated a larger delay than the development of vocabulary and sentence comprehension.

In addition, the SLI group showed no development (i.e., decrease) in the use of dummy auxiliaries. This pattern did not resemble either of the control groups and might therefore be labelled as a deviance pattern. Grammatical development in the SLI group constituted a complex picture, consisting of diverging developmental patterns for the different measures of complexity and correctness. The magnitude of the delays varied for the separate measures and also changed over time. Therefore, it can be concluded that the children with SLI do not just show simple delays in their language development. Their developmental trajectories rather demonstrate atypical language learning (Thomas, 2006), which is perhaps the best characterization of language development in SLI. Nonetheless, a positive note is that the SLI group showed continuous progress on the grammatical complexity and correctness measures between the ages 6 and 8 years and no evidence for total stagnation was found.

Role of the lexicon

The relation between lexical and grammatical development in older school-age children with SLI was studied in Chapter 3. In typical language acquisition, lexical development is predictive of grammatical development (Bates & Goodman, 1997; Marchman & Bates, 1994; Tomasello, 2003). Although there is some evidence for a similar relationship between vocabulary and grammar in clinical groups (Marchman, Saccuman, & Wulfeck, 2004; McGregor et al., 2012), this relationship has so far not been studied longitudinally in school-age children with SLI. Grammatical development of children with SLI between the ages 8 and 10 years was investigated by analyzing grammatical complexity and correctness. Next, the relationship between lexical knowledge (i.e., scores on three vocabulary tests) and grammar was examined in the SLI group with correlation analyses.

The SLI group showed progress between the ages 8 and 10 years on all complexity measures. Percentages of dummy auxiliaries and verb-related errors decreased, but a stagnation was observed for non-verb related errors. As compared to the CA group, the 10-year-old children with SLI produced fewer complex sentences, among which complex-compound sentences (i.e., sentences containing two or more independent clauses combined with one or more dependent clauses). On the correctness measures, the SLI group was outperformed by both control groups at all three time points. The correlation analyses showed that individual grammatical correctness scores of the children with SLI were strongly associated with their vocabulary knowledge, most markedly so for non-verb related errors and the use of dummy auxiliaries.

The findings suggest a double deficit in children with SLI. On the one hand, processing difficulties appear to limit the use of complex syntax, and, on the other hand, the observation that non-verb related errors do not remit hints at incomplete linguistic knowledge. Furthermore, results of the correlation analyses were argued to be compatible with the declarative/procedural model of lexicon and grammar (Ullman, 2001, 2004) and with the procedural deficit hypothesis for SLI (Ullman & Pierpont, 2005). The declarative/procedural model posits that the mental lexicon and grammar are both subserved by two at least partially distinct memory systems in the brain. The procedural memory system is supposed to subserve rule-governed aspects of language (i.e., grammar), and the declarative memory system underlies explicit episodic and semantic (lexical) knowledge. According to the procedural deficit hypothesis (Ullman & Pierpont, 2005), the grammatical deficits in SLI can largely be explained by deficits in the procedural memory system. Children with SLI are claimed to have an impaired procedural memory system, leading to deficits in implicit sequence learning and grammar. However, the declarative memory system, which supports the acquisition of vocabulary, is supposed to remain largely intact in children with SLI. Therefore, it is suggested that they can compensate the procedural impairment by storing grammatical forms and learning grammatical rules explicitly in declarative memory. Given that the lexicon is supported by declarative memory, a well-developed vocabulary would suggest a high capacity for compensation of a procedural deficit by the declarative system. The correlation analyses in this study lend some support to this line of reasoning.

Role of speech fluency

It has been suggested in the literature that the analysis of speech disfluencies can offer a window on sentence formulation problems in older school-age children with SLI, even when their grammatical accuracy is fairly high (Finneran, Leonard, & Miller, 2009). Therefore, in Chapter 4, speech disruptions were investigated in the older school-age children with SLI, aged 8 to 10 years. The narratives were examined for frequencies, types and syntactic distribution of disfluencies. Results showed that the SLI group exhibited more stalls than the CA and LA groups, but that the groups did not differ on revisions. Frequencies and types of disfluencies in the SLI group resembled those of the LA group more than those of the CA group. The syntactic distribution of disfluencies followed the same pattern in all three groups: highest frequencies were found at utterance-initial position, followed by clause-initial, phrase-initial, and word-initial position. However, the SLI group had more disfluencies than CA peers at word-initial positions, suggesting that problems with lexical retrieval may contribute to their disfluency. Over time, only silent pauses decreased in the SLI group. The higher disfluency rates in the SLI group were regarded to reflect their compromised expressive language skills. It was concluded that, although some improvement in fluency was seen, sentence formulation remained challenging for older school-age children with SLI. The higher disfluency rates also hint at processing difficulties in the SLI group. Apparently, the children need more time to construct syntactic frames and retrieve lexical items.

Effectiveness of metalinguistic intervention

The effectiveness of intervention for older school-age children with SLI was addressed in Chapter 5. Based on the grammatical profiles obtained from the first two studies, the production of relative clauses was selected as intervention target. Twelve school-age children with SLI (age range 9;3 – 12;8 years) participated in a pilot-study using a quasi-experimental multiple-baseline design. In a metalinguistic intervention approach, called 'Metataal', grammatical rules were taught explicitly. In addition to using visual support, which is common practice for metalinguistic approaches, motor and tactile/kinesthetic support was provided. The present study used Lego® blocks in different shapes and colors to represent word categories and grammatical functions. The children effectively learned to build complex sentences with the blocks.

Five hours of individual therapy in five weeks produced significant progress on the relative clause production tasks, but not on the comprehension task. The gains on the production tests were also maintained three months later.

It was concluded that the motor and tactile/kinesthetic dimensions of a multimodal approach were a valuable addition to the existing metalinguistic intervention methods. The study provides evidence that metalinguistic intervention is effective in remediating grammatical difficulties in children with SLI, aged 9 and over. The current policy in the Netherlands to reduce intervention for children with SLI in the higher primary grades should therefore be reconsidered.

Synthesis of the findings

Concerning the delay versus deviance debate that was addressed in Chapter 2, it can be concluded that such a strict dichotomy does not do justice to the wide range of developmental patterns that were found in the first two studies. A delay-within-delay pattern partly suits the data, but labels such as 'atypical development' or 'atypical learning' may be even more appropriate for the observed irregular growth trajectories in the SLI groups. The children with SLI showed selective delays in complex syntax (i.e., relative clauses and complex-compound clauses), larger than expected delays on verb-related and non-verb related errors, with stagnation for the non-verb related errors. Moreover, the use of dummy verbs showed a stagnation up till age 8 years, and a subsequent decrease between the ages 8 and 10 years. The overall picture arising from the analysis of grammatical development in children with SLI between the ages 6 and 10 years supports the view that SLI is a dynamic condition with profiles changing over time (Conti-Ramsden & Botting, 1999; Van Weerdenburg, Verhoeven, & van Balkom, 2006).

An intriguing finding of this longitudinal study of grammatical development was that over time, more progress was found on verb-related errors than on non-verb-related errors. In the literature, the verb morphology problems of children with SLI are often emphasized (for overviews, see Leonard, 1998; Schwartz, 2009). Although the Dutch verb paradigms are only a little richer than is the case for English, this result indicates that studying SLI in many different languages is necessary to get a good grip of the symptoms of SLI across languages. Possibly, the fact that the Dutch inflectional system is slightly richer and a little more salient than the English system makes Dutch verb paradigms easier to learn. English has only one morpheme for present tense (*-s*) and one for past tense (*-ed*), whereas Dutch has two morphemes for present tense (*-t* and *-en*) and two for past tense (*-te/-de* and *-ten/-den*). In learning verb inflection, children are supposed to benefit from the richness of the inflectional system, as well as from the saliency of the grammatical morphemes. Such an explanation would combine the sparse morphology hypothesis (Leonard, 1987) and the surface hypothesis (Leonard, 1989).

The stagnation in non-verb related errors may be explained by the 'interface' characteristics of the elements involved. The correct use of pronouns, determiners and (pronominal) adverb '*er*' (there) is not exclusively guided by grammatical rules, but also by discourse/pragmatics-related constraints. The complexity of these grammatical elements may render them more difficult to learn or more difficult to process, both of which can be expected to lead to (more) errors. Another explanation may be found in the perceptual salience of the grammatical elements at stake. Dutch determiners, most pronouns, and (pronominal) adverb '*er*' (there) are monosyllabic words that are often pronounced with unstressed, schwa-like vowels. These phonologically weak morphemes representing abstract grammatical features are harder to identify and consequently more difficult for children with SLI (Leonard, 1989).

In the four studies in this dissertation, several indications were found for processing limitations in the SLI group. Firstly, both SLI groups performed poorly on the Kaufman sequential memory tests compared to LA and CA control groups (see Chapter 2 and 3). Secondly, the children with SLI produced fewer complex sentences than their CA/LA peers, especially sentences with relative clauses and complex-compound clauses. This suggests that they have the competence to construct such sentences, but that it is difficult for them to do so. Conceivably, processing limitations hamper the production of these lengthy and complex sentences types. A third observation suggestive of a processing limitation is that the (older school-age) children with SLI had more speech disfluencies than their CA peers and to some extent also when compared to their LA peers. These higher rates of disfluencies are regarded to reflect difficulties in sentence formulation. The SLI group demonstrated higher word-initial disfluency rates, which can be related to impaired lexical retrieval. Furthermore, the children with SLI also had more (stutter-like) part-word repetitions than their CA peers, which are interpreted by Guo, Tomblin, and Samelson (2008) to reflect difficulties in accessing the phonological form of words. Finally, support for processing difficulties explaining the language problems in SLI also comes from the metalinguistic intervention study 'Metataal' (Chapter 5). The investigation of grammatical development demonstrated that the children with SLI performed poorly on relative clause production (Chapter 2). The same result has been found in many different studies for different languages (for an overview, see Jensen de López, Sundahl Olsen, & Chondrogianni, 2014). However, the 'Metataal' intervention study (Chapter 5) suggests that children with SLI can learn to produce relative clauses, when the processing load is reduced by offering visual support. Working memory limitations have often been reported for SLI (Montgomery, Magimairaj, & Finney, 2010). According to Coco, Garraffa, and Branigan (2012) these working memory limitations affect relative clause production in children with SLI. To complete this line of reasoning, working memory performance in children with SLI has been shown to benefit from visual support (Quail, Williams, & Leitão, 2009). In conclusion, the findings from the different studies in this thesis offer some support for processing limitations in children with SLI.

To conclude, the answers to the research questions in this thesis end in three important take home messages. First, the longitudinal examination of grammatical development in older school-age children between the ages 6 and 10 years does not reveal patterns of (total) stagnation. On the contrary; the children demonstrate continuous development with regard to grammatical complexity and correctness. Secondly, speech disfluencies reflect sentence formulation difficulties in older school-age children with SLI. Supposedly, difficulties in constructing syntactic frames and in lexical retrieval both contribute to the observed higher disfluency rates in children with SLI. Thirdly, school-age children older than 10 years can (still) benefit from a metalinguistic intervention approach aimed at improving their grammatical skills.

Clinical implications

A number of findings of this dissertation are relevant for clinical practice. In the 'Metataal' intervention study, evidence was presented that an intervention based on a metalinguistic approach can improve production of relative clauses in older school-age children with SLI. Five hours of therapy in a 5-week intervention program yielded significant improvement on tests measuring relative clause production. Relative clauses were chosen because it could be considered a challenging target. If the program would be effective for these targets, it probably will also work for other grammatical constructions.

Currently, the bulk of speech therapy in special education for children with SLI is allocated to pre-schoolers and school-age children up to 7 years of age. In the Netherlands, resources for children in the higher grades are rather limited. However, a metalinguistic and multimodal approach promises to be effective in resolving the (remaining) grammatical difficulties of children in the higher grades of primary school. This asks for a reconsideration of the provision of speech therapy for the children with SLI in these grades.

Conceivably, the provision of speech therapy could follow a U-shaped pattern, starting with intensive therapy for the pre-schoolers with SLI, lasting up to age 6 or 7 years. In the following years, when reading, spelling and arithmetic become dominant in the curriculum, speech therapy might be less intensive, to be intensified again in the final primary school years, between the ages 10 and 12 years. At that age, most children have reached an appropriate reading level and have obtained sufficient metalinguistic skills to benefit from a metalinguistic therapy approach. Metalinguistic therapy could then be provided in different ways: individually, in small groups, or perhaps in classroom settings. Therapy in a class-room setting would probably be most cost-effective. Metalinguistic training could be supported further by the development of special adaptive apps, enabling students to practice at their own level, providing feedback, and choosing the next practice levels. Perhaps, in the near future apps with speech recognition can be developed that can also be used to evaluate the answers of the students. Of course, identifying the most efficient and effective form(s) of intervention for these age groups demands further research.

Importantly, the results of this study suggest that vocabulary expansion in children with SLI should constitute a substantial part of the curriculum in schools for children with SLI. The findings reported in Chapter 3 showed that lexical knowledge is strongly correlated with grammatical correctness and the use of dummy verbs. When vocabulary scores increase, grammatical error rates and the use of dummy verbs decrease accordingly. Obviously, correlation does not tell us anything about the direction of causation (if any), but a plausible interpretation may be that a strong vocabulary may help overcome grammatical weakness. Vocabulary skills are generally considered to be a relative strength in children with SLI, as compared with their morphosyntactic deficits and are therefore not always regarded as problematic and needing extra attention. However, vocabulary expansion benefits language comprehension and the ability to express oneself. A causal relationship between vocabulary and grammar should be examined further.

Additionally, the grammar profiles of Dutch school-age children with SLI between the ages 6 and 10 years obtained in this study may be helpful for practitioners to select appropriate targets for speech and language therapy. It can be difficult to select appropriate intervention targets for certain age levels. For instance, dummy auxiliaries, which can be regarded as a prominent characteristic of Dutch SLI, did not decrease between the ages 6 and 8 years. However, a decrease was found between the ages 8 and 10 years. Possibly, dummy verbs are used as an economy 'strategy' until verb inflection paradigms are acquired for the most part, after which the use of dummy verbs gradually becomes superfluous. Thus, it may not be necessary to treat the overuse of dummy verbs, as their use will gradually fade out.

Finally, this dissertation shows that the analysis of narratives offers the clinician valuable information on the language skills of the children that cannot be obtained with standardized tests. A narrative analysis can inform practitioners on grammatical complexity and grammatical correctness, the quality of vocabulary, speech fluency, and on the expression of coherence and cohesion. This information can be used for diagnostics, the selection of appropriate therapy goals and the evaluation of intervention. Furthermore, a narrative task is usually not tiresome for children, as they generally like storytelling. Although transcription and coding of a narrative sample can be laborious, the valuable information that can be derived from a narrative analysis is worth the effort. It is therefore advisable to invest in a further development of tools that can facilitate narrative analysis (lexical, grammatical, discourse structural) for speech therapists (and researchers).

Suggestions for further research and future directions

This study demonstrated that analyzing the language profiles / development of 6-10 year old children with SLI contributes to a better understanding – both theoretically and clinically – of the impairment. It is expected that investigation of grammatical development in even older children, adolescents, and adults will provide more valuable information on issues such as compensation strategies, fossilization, trade-offs between semantic complexity and grammatical correctness, and fluency breakdowns, adding to our knowledge of language impairments in these age groups.

There is an urgent need for intervention studies addressing the language difficulties of children with SLI. The government and health insurers ask for evidence-based practice, and in the light of recent, extensive austerity measures, the demand for evidence of effectiveness of interventions becomes stronger and stronger. Moreover, practitioners themselves also want to know whether their approaches are effective. In this light, studies showing that certain approaches do not work deserve publication just as much as those that show their effectiveness. Furthermore, to this day, studies addressing the efficacy of intervention approaches aimed at language *comprehension* hardly exist. This is a gap that will have to be filled in the near future. The results of such intervention studies also have to be made available to the field and have to be translated into methods that can be (effectively) implemented in actual practice. The rapid development in new technologies (tablets, apps, and virtual learning environments) can boost innovation of therapeutic techniques and methods, such as telepractice and adaptive software supporting intervention.

Well-designed intervention studies can also be valuable from a theoretical perspective. Such studies can be very informative on compensation strategies and atypical learning processes in children with SLI. For instance, as regards the relationship between vocabulary and grammar reported in this thesis, intervention studies appear to be an excellent tool to further investigate a direction or possible causal relationship between the two.

The above recommendations with regard to research and practice and their interaction, are possible only if research institutions and institutions providing care and education to language-impaired individuals work closely together. In this respect, some promising initiatives have recently been taken in the Netherlands. However, the collaboration of universities and institutes for special education with respect to the research agendas, fund-raising, data collection and making these data available for research purposes should be expanded further. Because of advances in computer technology, the collection, storage, and sharing of data has become much easier, especially the aggregation of behavioral, neurocognitive, and genetic data. Such a large and accessible database can also reduce the recurring burden on children, schools, and health care institutions having to provide researchers with new data on a regular basis.

Furthermore, another positive development is that more and more practitioners graduate at universities and become scientist-practitioners. When these professionals are employed and facilitated in schools and institutions, they can coach and supervise practitioners, and thereby increase the amount of practice-based research. The special schools and health care institutions already gather large amounts of data on their students and clients. With minor adjustments to data collection and storage, this information can be used for practice-based research. Such a practice-based research would entail some important benefits, such as raising the level of expertise at the schools and institutions, easing the dissemination of evidence-based knowledge, and stimulating an investigative attitude in staff and personnel.

Last but not least, a point that certainly needs more attention is to raise awareness of SLI (or TOS, *TaalOntwikkelingsStoornis* in Dutch) in society by publicity campaigns. More awareness of the impact of SLI on individuals and society requires information campaigns, using all available media. An ambassador, such as a famous actor, epitomizing the SLI awareness movement would be very helpful. However, it is also very important to raise awareness in clients themselves by means of psycho-education and self-help groups. All too often, children and adolescents with SLI do not know what having a language impairment entails and accordingly, they are not assertive enough in dealing with their language difficulties. By empowering individuals with SLI, they can also play a role in informing society and raising awareness. Testimonials of individuals with SLI probably are the most effective and impressive way to explain to the general public what having a language impairment means.

More awareness of SLI in society could lead to a better understanding of language impairments and thus aiding acceptance and participation. Children with persistent SLI are at high risk of poor literacy and educational failure. In adolescence and adulthood, high rates of unemployment, social isolation and psychiatric disorder have been found. Therefore it is necessary to invest more in education, health care, and research for these individuals with SLI, if only for reasons of cost-effectiveness.

References

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

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42,* 1195–1204.

Coco, M. I., Garraffa, M., & Branigan, H. P. (2012). Subject Relative Production in SLI Children during Syntactic Priming and Sentence Repetition. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012), Sapporo, Japan, August, 2012* (pp. 228–233).

De Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure*. Groningen: Grodil. (Doctoral dissertation, University of Groningen). Retrieved from: http://irs.ub.rug.nl/ppn/297976826.

De Jong, J. Blom, E. Orgassa, A. (2013). Dummy auxiliaries in children with SLI - a study on Dutch, in monolinguals and bilinguals. In E. Blom, I. van de Craats, & J. Verhagen (Eds), *Dummy auxiliaries in first and second language acquisition* (pp. 251-278). Berlin: De Gruyter Mouton.

Jensen de López, K., Sundahl Olsen, L., & Chondrogianni, V. (2014). Annoying Danish relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of Child Language*, *41*, 51-83.

Leonard, L. B. (1987). Is specific language impairment a useful construct? In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Disorders of first language development* (pp. 1–39). Cambridge, NY: Cambridge University Press.

Leonard, L.B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10,* 179-202.

Leonard, L. B. (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339–366.

Marchman, V. A., Saccuman, C., & Wulfeck, B. (2004). Productive use of the English past tense in children with focal brain injury and specific language impairment. *Brain and Language*, *88*, 202-214.

McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders*, *42*, 35–47.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: an update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*, *19*, 78–94.

Orgassa, A. (2009). *Specific language impairment in a bilingual context. The acquisition of Dutch by Turkish-Dutch learners*. Utrecht: LOT dissertation series, No.220. (Doctoral dissertation, University of Amsterdam). Retrieved from: http://www.lotpublications.nl/publish/articles/003604/bookpart.pdf.

Quail, M., Williams, C., & Leitão, S. (2009). Verbal working memory in Specific Language Impairment: The effect of providing visual support. *International Journal of Speech-Language Pathology*, *11*, 220–233.

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy and J. Schaeffer (Eds.) *Language competence across populations. Toward a definition of specific language impairment (*pp. 63-95). Mahwah: Lawrence Erlbaum Associates.

Schwartz, R.G. (2009). Specific language impairment. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 3-43). New York, NY: Psychological Press.

Thomas, M. S. C. (2006). Language acquisition in developmental disorders. In M. Kail, M. Hickmann, & M. Fayol (Eds.), *Proceedings of the International Conference on First and Second Language Acquisition, Paris 2006*. (pp. 1–32). Amsterdam: John Benjamins Pub Co.

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

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*, 37–69.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92,* 231-270.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*, 399–433.

Van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry, 47*, 176-189.

Samenvatting in het Nederlands

In deze dissertatie wordt de grammaticale ontwikkeling van kinderen met een primaire taalontwikkelingsstoornis (TOS) tussen 6 en 10 jaar in kaart gebracht en worden therapiemogelijkheden voor leerlingen met een TOS in de bovenbouw van het basisonderwijs onderzocht.

Een TOS wordt vastgesteld op basis van een aantal exclusie- en inclusiecriteria. Gebruikelijke exclusiecriteria zijn dat de taalstoornis kan niet worden verklaard door gehoorproblemen, aantoonbare neurologische afwijkingen, ernstige lichamelijke of emotionele problemen, cognitieve beperkingen of ongunstige omgevingsinvloeden. Inclusiecriteria zijn onder andere dat de taalproblemen ernstig en hardnekkig moeten zijn. Een TOS heeft grote gevolgen voor het kind en zijn omgeving en grijpt in op de sociaal-emotionele ontwikkeling, de leerontwikkeling en op communicatie en participatie.

Tot nog toe hebben oudere basisschoolkinderen in het onderzoek naar Nederlandstalige kinderen met een TOS vrij weinig aandacht gekregen. Ook is er nog weinig bekend over therapiemogelijkheden voor kinderen met een TOS van tien jaar en ouder. De vier deelstudies uit het huidige onderzoek leveren een bijdrage aan onze kennis over kinderen met een TOS in de basisschoolleeftijd.

De grammaticale ontwikkeling van twee leeftijdsgroepen van kinderen met een TOS (30 kinderen per groep) is geanalyseerd door de verhalen van kinderen met een TOS, uitgelokt met de TAK Verteltaken, drie jaar lang te volgen. Deze verhalen zijn vergeleken met die van twee controlegroepen van kinderen met een normale taalontwikkeling (ook 30 kinderen per groep). De controlegroepen waren kinderen met dezelfde kalenderleeftijd (KL) en kinderen met een vergelijkbare taalleeftijd (TL). De kinderen in de laatste groep waren gemiddeld twee jaar jonger. Bij de groep kinderen met een TOS van 8-10 jaar zijn ook de onvloeiendheden in de spraak geanalyseerd. De verschillende types onvloeiendheden en hun positie in de zin kunnen inzicht bieden in de formuleringsmoeilijkheden van kinderen met een TOS. Daarnaast is op basis van het verkregen grammaticaal ontwikkelingsprofiel een pilotstudie verricht naar het effect van een therapieprogramma bij kinderen met een TOS van 10 jaar en ouder. Deze pilot is uitgevoerd op een cluster 2 school (onderwijs aan kinderen met auditieve en communicatieve beperkingen).

In hoofdstuk 1 worden definities van TOS besproken en de belangrijkste theoretische verklaringen voor TOS beschreven. Daarnaast wordt een overzicht van het wetenschappelijk onderzoek naar grammaticale problemen van Nederlandstalige kinderen met een TOS gepresenteerd.

Hoofdstuk 2 bevat de resultaten van een longitudinaal onderzoek naar de grammaticale ontwikkeling van kinderen met een TOS op de leeftijden 6, 7 en 8 jaar. Voor dit onderzoek is gekeken naar grammaticale complexiteit en grammaticale correctheid (fouten) in de verhalen uitgelokt met de TAK-R Verteltaken. De uitkomsten zijn gerelateerd aan het debat of er bij TOS sprake is van een achterstand in de grammaticale ontwikkeling of dat de taalontwikkeling ook volgens afwijkende patronen verloopt.

Uit de analyses van de verhalen van de kinderen blijkt dat de kinderen met een TOS tussen 6 en 8 jaar een continue groei laten zien in grammaticale complexiteit (gemiddelde uitingslengte en het gebruik van samengestelde zinnen). Ook neemt het aantal grammaticale fouten (werkwoord gerelateerde en niet-werkwoord gerelateerde fouten) af naarmate de kinderen ouder worden. Over het algemeen lijkt de taalontwikkeling van kinderen met een TOS op die van de twee jaar jongere TL controlegroep. Dit geldt vooral voor de complexiteitsmaten en de scores op reguliere taaltesten voor woordenschat en taalbegrip. De 8-jarige kinderen met een TOS gebruikten minder relatieve bijzinnen dan de 8-jarige KL controlegroep. Als we echter kijken naar grammaticale fouten, dan scoren de kinderen met een TOS niet alleen ver onder het niveau van de KL controlegroep maar ook veel lager dan de TL controlegroep. Bij de kinderen met een TOS is de grammaticale correctheid veel lager dan verwacht op grond van hun scores op woordenschat en taalbegrip.

Deze bevindingen kunnen het best gekarakteriseerd worden als een onevenwichtig profiel (Leonard, 1998) of een achterstand-binnen-achterstand (*delay-within-delay*) patroon (Rice, 2003). In ieder geval duiden de bevindingen op een asynchrone ontwikkeling van verschillende deelvaardigheden binnen de taalontwikkeling van kinderen met een TOS.

Een kenmerk van Nederlandse kinderen met een TOS is het overmatige gebruik van dummy werkwoorden (hulpwerkwoorden die geen betekenis dragen in de zin, zoals ‘gaan’ en ‘doen’ gecombineerd met een infinitief). Een overmatig gebruik van deze dummy werkwoorden kan worden gezien als een indicatie van een vertraagde grammaticale ontwikkeling. Jonge kinderen met een normale taalontwikkeling gebruiken ook dummy werkwoorden, maar bij deze groep dooft het gebruik van dummy werkwoorden geleidelijk uit. Een mogelijke verklaring voor het gebruiken van dummy werkwoorden is dat kinderen moeite hebben met de vervoeging en verplaatsing van het lexicale werkwoord. Ze vermijden daarom deze operaties door een dummy werkwoord in te voegen. In de huidige studie werd bij de kinderen met een TOS tussen de leeftijden 6 en 8 jaar geen ontwikkeling gevonden in het gebruik van dummy werkwoorden. Dit ontwikkelingspatroon werd niet gezien bij de twee controlegroepen, en zou daarom als afwijkend beschouwd kunnen worden.

Hoofdstuk 3 beschrijft de resultaten van een longitudinaal onderzoek naar de grammaticale ontwikkeling van 30 kinderen met een TOS op de leeftijden 8, 9 en 10 jaar. Bij deze studie werden eveneens TL en KL controlegroepen (30 kinderen per groep) gebruikt. Bij de kinderen met een TOS zijn de ontwikkeling van de grammaticale complexiteit en correctheid vervolgens gerelateerd aan de woordenschatontwikkeling. Het doel was om na te gaan of er een relatie tussen grammatica en woordenschat bij de kinderen met een TOS bestaat. Bij jonge kinderen met een normale taalontwikkeling blijken deze aspecten namelijk sterk aan elkaar gerelateerd. Daarnaast is ook bij deze oudere kinderen met een TOS gekeken naar de ontwikkeling van het gebruik van dummy werkwoorden.

De grammaticale vaardigheden van de kinderen met een TOS lieten een gestage ontwikkeling zien tussen 8 en 10 jaar met betrekking tot grammaticale complexiteit en de werkwoord gerelateerde fouten. Ook het gebruik van dummy werkwoorden nam af. Deze afname werd niet gevonden voor de niet-werkwoord gerelateerde fouten. Op de complexiteitsmaten presteerden de kinderen met een TOS beter dan de TL controlegroep maar zwakker dan de KL controlegroep. De 10-jarige kinderen met een TOS maakten vooral minder (meervoudige) samengestelde zinnen dan de KL groep. De grammaticale correctheid bleef ook bij deze oudere leeftijdsgroep van kinderen met een TOS ver achter bij het niveau van beide controlegroepen.

Bij de kinderen met een TOS bleek grammaticale correctheid sterk te correleren met scores op de TAK passieve en actieve woordenschattests en de TAK woorddefinitietest. Een grote woordenschat ging gepaard met lage percentages (van vooral) niet-werkwoord gerelateerde fouten en dummy werkwoorden. De resultaten van de correlatieanalyses bieden ondersteuning aan het declaratief /procedureel model voor lexicon en grammatica (Ullman, 2001). Volgens dit model zijn er min of meer gescheiden geheugensystemen verantwoordelijk voor leren van woorden (het declaratief geheugen) en het leren van grammaticale regels (het procedureel geheugen). De hypothese dat kinderen met een TOS een aangedaan procedureel geheugen hebben (PDH: *procedural deficit* *hypothesis*, Ullman & Pierpont, 2005) met daarnaast een min of meer intact declaratief geheugen zou verklaren dat kinderen met TOS vooral grammaticale problemen hebben en een relatief beter ontwikkeld lexicon. Volgens de PDH kunnen kinderen met een TOS hun relatief sterkere declaratieve geheugen gebruiken om te compenseren voor hun zwakke procedureel geheugen. De kinderen zouden grammaticale regels expliciet kunnen leren door deze regels op te slaan in het declaratieve geheugen.

Hoofdstuk 4 beschrijft de ontwikkeling van onvloeiendheden geanalyseerd in de TAK Verteltaken van 30 kinderen met een TOS tussen 8 en 10 jaar. De resultaten van de TOS groep werden vergeleken met TL en KL controlegroepen van 30 kinderen elk. De onvloeiendheden werden geanalyseerd voor frequentie, type, en de positie van onvloeiendheden binnen de zin. Bij jonge kinderen met een normale taalontwikkeling zien we een relatie tussen vloeiendheid en de mate van taalvaardigheid. Rond 2 á 2 ½ jaar maken de meeste kinderen een periode van verhoogde onvloeiendheid door. Zodra het proces van zinsformulering meer geautomatiseerd verloopt, neemt deze onvloeiendheid weer af. Kinderen met een TOS hebben grammaticale problemen en vaak ook problemen met woordvinding. Uit diverse internationale studies is gebleken dat kinderen met een TOS meer onvloeiendheden produceren. Dit bleek zelfs het geval te zijn als zinnen grotendeels grammaticaal correct geformuleerd worden (Finneran, Leonard & Miller, 2009). Een analyse van de verschillende types onvloeiendheden en de posities van onvloeiendheden binnen de zin kan inzichtelijk maken in hoeverre grammaticale en lexicale problemen beide bijdragen aan de onvloeiendheden van kinderen met een TOS.

Onvloeiendheden in de spraak kunnen worden onderverdeeld in twee functionele categorieën. Enerzijds zien we onvloeiendheden die de spreker meer tijd geven voor het formuleringsproces (stille en gevulde pauzes, herhalingen van klanken, lettergrepen, woorden en zinsdelen) en anderzijds zien we revisies, waarbij de spreker reeds geproduceerde spraak verbetert.

Uit de analyses bleek dat er nauwelijks verschillen waren tussen de twee controlegroepen met een normale taalontwikkeling. De kinderen met een TOS produceerden echter wel meer onvloeiendheden dan beide controlegroepen. Het verschil met de TL groep was kleiner dan het verschil met de KL groep. Deze resultaten wijzen erop dat onvloeiendheden in de spraak meer gerelateerd zijn aan het niveau van taalvaardigheid dan aan de leeftijd. Het aantal revisies verschilde niet tussen de groepen. Dit suggereert dat de kinderen met een TOS niet afwijken van de controlegroepen in het kunnen monitoren van hun spraak.

Daarnaast vertoonden alle drie de groepen hetzelfde patroon van grammaticale distributie van de onvloeiendheden binnen de zin. De hoogste percentage onvloeiendheden werd gevonden aan het begin van de uiting, gevolgd door respectievelijk aan het begin van een bijzin, aan het begin van een zinsdeel en in woord-initiële positie. De kinderen met een TOS produceerden echter meer onvloeiendheden in woord-initiële positie dan de controlegroepen. Deze bevinding suggereert dat problemen met woordvinding bijdragen aan de onvloeiendheid bij kinderen met een TOS.

De patronen van de onvloeiendheden veranderden in de drie groepen maar beperkt tussen 8 en 10 jaar. Bij de kinderen met een TOS groep verminderde alleen de frequentie van stille pauzes. In alle groepen verminderden de aantallen stille pauzes langer dan 2 seconden. Bij de KL groep verminderde ook het aantal onvloeiendheden aan het begin van een zinsdeel. Dit werd geïnterpreteerd als een verdere perfectie van het proces van zinsfomulering, omdat planning van uitingen voornamelijk plaatsvindt aan het begin van zinnen en zinsdelen.

Hoofdstuk 5 doet verslag van een pilotstudie naar het effect van een metalinguïstische en multimodale taaltherapie, genaamd 'Metataal', bij oudere kinderen met een TOS. Tot op heden worden de meeste effectstudies uitgevoerd bij jonge kinderen met een TOS. De laatste jaren verschijnen er echter meer internationale studies naar metalinguïstische interventie voor oudere kinderen met een TOS. Bij deze benadering leren kinderen grammaticale regels expliciet en het gebruik visuele ondersteuning vormt een belangrijk onderdeel van de aanpak. In de 'Metataal' pilot werden relatieve bijzinnen aangeleerd aan kinderen met een TOS van 10 jaar en ouder. Naast het bieden van visuele ondersteuning werden ook de motorische en tactiele/kinesthetische modaliteiten ingeschakeld. De kinderen leerden letterlijk zinnen te bouwen met behulp van Lego blokjes. De blokjes stonden voor verschillende woordsoorten en grammaticale functies. Met een beperkte set van blokjes konden alle woord- en zinsstructuren gebouwd worden.

Het gekozen onderzoeksdesign was een quasi-experimenteel design met een '*repeated baseline*'. Daarbij werden 3 nulmetingen uitgevoerd voor de start van de therapie, daarna volgde de behandeling die werd afgesloten met een nameting direct na afloop van de therapie. Drie maanden later werd een retentiemeting uitgevoerd om te kijken of het geleerde ook beklijfd was. De tests bestonden uit productietaken, nazegtaken en een begripstaak. Twaalf kinderen werden vijf weken lang tweemaal per week behandeld met een speciaal ontwikkeld behandelprogramma dat strikt gevolgd werd. Gedurende de hele duur van het project bleef de behandelaar blind voor de resultaten van de metingen.

Tijdens de baselinemetingen werd geen vooruitgang gevonden op de tests. Direct na de therapiefase waren de kinderen significant vooruitgegaan op de productietaken, maar niet op de begripstaak. Deze vooruitgang op de productietaken was nog steeds aanwezig bij de retentiemeting. Het programma was misschien te kort of niet intensief genoeg om ook een verbetering in het begrip van relatieve bijzinnen teweeg te brengen.

Deze pilotstudie ondersteunt het bewijs uit andere interventiestudies dat grammaticale problemen van oudere kinderen met TOS effectief behandeld kunnen worden met een metalinguïstische aanpak. De multimodale aanpak van 'Metataal' kan gezien worden als een waardevolle aanvulling op al bestaande therapievormen. Het huidige beleid om directe interventie voor oudere kinderen met TOS te verminderen zou dan ook heroverwogen moeten worden. Verder onderzoek is nodig om te kijken welke manier van aanbieden (individueel, groep, klassikaal) en welke frequentie het meest optimale en kosteneffectieve resultaat oplevert.

In hoofdstuk 6 worden de resultaten van de vier deelstudies samengevat, onderlinge relaties tussen de studies worden besproken en de bevindingen worden gerelateerd aan theorieën over taalontwikkelingsstoornissen. Daarnaast worden een aantal klinische implicaties en voorstellen voor verder onderzoek en verder beleid gepresenteerd.

De conclusies van de vier deelstudies in deze dissertatie zijn allereerst dat de grammaticale ontwikkeling van kinderen met TOS tussen de leeftijden 6 en 10 jaar nog voortduurt. Zowel voor grammaticale complexiteit en correctheid zijn er geen patronen van totale stagnatie gevonden. Daarnaast zien we bij oudere kinderen met een TOS een sterke associatie tussen woordenschat en grammaticale vaardigheden. Bij kinderen met een TOS in deze leeftijdsgroep kan de analyse van onvloeiendheden in de spraak ook meer inzicht bieden in hoeverre problemen met grammatica en woordvinding bijdragen aan formulerings-moeilijkheden. Ten slotte werd gevonden dat een relatief korte multimodale en metalinguïstische interventie de grammaticale vaardigheid van oudere kinderen met een TOS nog verder kan verbeteren.

Curriculum Vitae

Rob Zwitserlood was born on 29 May 1959 in Schaesberg, The Netherlands. After obtaining his VWO diploma in 1977 from the Bernardinuscollege in Heerlen, he studied English Language and Literature at the University of Utrecht for two years. From 1982 to 1986 he studied speech and language therapy at the Opleiding voor Logopedie en Akoepedie te Utrecht. After graduation, he worked as a speech and language therapist at various schools for special education (LOM, MLK, ZMOK). In 1987, he started to work at the special education schools for children with severe speech and language disorders. In 1999, he obtained his diploma (Preventief) Ambulante Begeleiding at Fontys Hogescholen, and started working as an educational coach (ambulant begeleider) at ACC Dienstverlening, later the Koninklijke Auris Groep. Between 2001 and 2005 he was also employed at the Seminarium voor Orthopedagogiek (Post HBO Special Education) as a teacher Language Development and Disorders. In 2007 he obtained his master degree in Linguistics (Speech and language: Processing and disorders) at the University of Utrecht. In 2009 he started at the Utrecht Institute of Linguistics OTS as an external PhD student, funded by the Koninklijke Auris Groep, This thesis is the result of the research he carried out there between 2009 and 2014. Rob Zwitserlood is currently working as a staff member/researcher at the Koninklijke Auris Groep.