Persistent grammatical difficulties in Specific Language Impairment

Deficits in knowledge or in knowledge implementation?

Duinmeijer, I.

Publication date
2017

Document Version
Final published version

License
Other

Iris Duinmeijer

Persistent grammatical difficulties in Specific Language Impairment

Deficits in knowledge or in knowledge implementation?

This study examines the grammatical abilities of children and adolescents with Specific Language Impairment (SLI). There were two research goals. Firstly, the persistence of grammatical problems over time was examined by comparing a younger group of children with SLI and an older group of adolescents with SLI. Secondly, this study explored whether difficulties in the grammatical domain in SLI purely reflect a grammatical deficit or may partly stem from problems in the implementation of grammatical knowledge due to problems in information processing. In the grammatical production tasks, the complexity of the linguistic context was therefore varied to examine whether this would cause a (larger) decrease in scores in the SLI groups. In addition, different measures of information processing ability were administered and the link between variability in performance and processing abilities was examined.

On the basis of the outcomes this book makes two major claims. Firstly, grammatical problems in SLI are persistent into adolescence. For some grammatical aspects such as grammatical gender, fossilization seemed to take place before children reach adolescence. Other aspects, such as verb inflection, had clearly been acquired by adolescence but differences between SLI and typical development (TD) in the amount of errors still remained. Secondly, grammatical performance in SLI was affected by the linguistic context in which grammatical knowledge had to be implemented. Such effects were small or absent in the TD groups. The effect of context was related to the verbal processing abilities of the groups. Grammatical problems in SLI therefore do not always reflect a deficit in grammar. Even when grammatical knowledge has been acquired, a child or adolescent with SLI is not always able to implement this knowledge in performance.

This dissertation is of relevance to researchers in the fields of language acquisition and language disorders, as well as to clinicians and teachers working with children and adolescents with language impairments.

: LOT Netherlands Graduate School of Linguistics

ISBN 978-94-6093-221-2

: LOT 440

Iris Duinmeijer

Persistent grammatical difficulties in Specific Language Impairment

Iris Duinmeijer

Persistent grammatical difficulties in Specific Language Impairment

Deficits in knowledge or in knowledge implementation?

University of Amsterdam

Amsterdam Center for Language and Communication

: LOT Netherlands Graduate School of Linguistics

Landelijke Onderzoekschool Taalwetenschap

Persistent grammatical difficulties in Specific Language Impairment

Deficits in knowledge or in knowledge implementation?

Iris Duinmeijer

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 photo by Eduard Lampe: Het Bungehuis

ISBN: 978-94-6093-221-2
NUR 616

Persistent grammatical difficulties in Specific Language Impairment

Deficits in knowledge or in knowledge implementation?

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie, in het openbaar te verdedigen in de Agnietenkapel op
op vrijdag 3 februari 2017, te 10.00 uur

door Iris Duinmeijer
geboren te Hoorn

Promotiecommissie:

Promotores:	prof. dr. A.E. Baker, Universiteit van Amsterdam
	prof. dr. F.P. Weerman, Universiteit van Amsterdam
Copromotor:	dr. J. de Jong, Universiteit van Amsterdam
Overige leden:	prof. dr. N.F. Botting, City University London, UK
	prof. dr. T. Marinis, University of Reading, UK
	dr. W.B.T. Blom, Universiteit Utrecht
	prof. dr. J.C. Schaeffer, Universiteit van Amsterdam
	prof. dr. A.C.J. Hulk, Universiteit van Amsterdam
	dr. J.E. Rispens, Universiteit van Amsterdam

Faculteit der Geesteswetenschappen

Table of contents

Acknowledgements

While I spent five years of my life on the 300 pages that follow, I am well aware that these six pages I wrote in a few evenings will probably be the pages that will be read the most. I will therefore not even try to keep it short, especially because I have a lot of gratitude to express.

First of all, I would like to thank my promotores, Fred Weerman and Anne Baker, and my co-promotor Jan de Jong for their belief in me as a candidate, their trust during the project and their tremendous guidance towards the end goal of this study. Each of you had a different supervision style and the three of you made a wonderful combination. Jan, you witnessed the emergence of my interest in language pathology during my bachelor and you have played an important role in my whole academic career. 'Tusen takk' for your well timed encouragements, the equality you always tried to create by calling me a colleague while I still felt like a newbie and for being a great example of a scientist with a strong clinical view. Fred, apart from sharing your impressive knowledge on Dutch grammar you often challenged me to formulate my ideas more concretely. Your raised eyebrow now and then, and the discussions that followed brought me many new insights. Thank you for your open attitude and your patience with my generative struggles. Anne, without your repetitious calls for conciseness this manuscript would have been twice as big. It is exceptional how you managed to be critical and encouraging at the same time. Thank you for all the little presents you bought your PhDs on trips abroad, for your counselling during stressful periods, for the correction of my Dunglish and for the fun we had during conferences.

Secondly, I would like to express my gratitude to my reading committee, consisting of Nicola Botting, Theo Marinis, Elma Blom, Jeannette Schaeffer, Aafke Hulk and Judith Rispens. You all have inspired me during my project and I am glad you found the time to read my manuscript and take part in the ceremony. Nicola, my three-month-stay at your lab in London has been of great value to me, both at a personal and a scientific level. Thank you (and all the other colleagues at City University) for your inspiration and hospitality. Theo, I profited much from the feedback you and Ianthi Tsimpli gave me on the gender results when I presented them in your lab in Reading. Thank you for sharing your thoughts and knowledge. Elma, I had the pleasant experience of collaborating with you for a very short period (in data collection and in teaching) before you went off to Utrecht. Thank you for the discussions we had on basic notions like competence and performance. I hope you have noticed that your comments have had an effect on my thinking. Jeannette, we joined forces in the data collection for this project, which led to a large dataset and the opportunity to combine our results. Thank you for the feedback you provided at several moments during the past five years. I hope we will be able to publish some joint articles in the future. Aafke, it was wonderful to have a gender-specialist amongst our research group. Thank you for your input on this topic during the past five years. Last but not least, Judith, thank you all the little corridor conversations that often took much longer than we intended. Your energy and enthusiasm have always inspired me and increased my wish to stay in science.

Performing this study would never have been possible without the altruistic help of all the children, adolescents and parents who participated in this study. I am so thankful that they agreed to spend hours performing my sometimes quite boring tasks. Many of your faces still reappear in my mind now and then and most of your names are still in the back of my head. Parents of the children and adolescents with SLI often responded to my expressions of gratitude by saying 'how important it is to have research on SLI to increase knowledge and awareness of these language impairments'. I truly hope that my study adds to this.

I am, furthermore, highly indebted to all the teachers and other staff members of the primary and secondary schools who participated: the '*Professor van Gilseschool'* in Haarlem, Beverwijk and Hoofddorp (in particular Veronique Key, Karin Wajon and José Keesmaat), '*de Weerklank'* in Leiden (in particular Yvonne van der Mee, Geke de Jong and Mirjam de Mooij), '*het Orion College'* in Amsterdam (in particular Audrey Franssen) *'het Rotsoord'* in Utrecht (in particular Leona Brandse and Jetty van Eys), '*de Ambulante Begeleidingsdienst'* in Houten (in particular Marieke de Vries-Hofman and Rian van Vollenhoven), *'de Linnaeusschool'* in Amsterdam (in particular Kees van Houtum, Marleen Heetveld and Elly) '*de Kaap'* in Amsterdam (in particular Yvonne van Putten), '*de Beatrixschool'* in Haarlem (in particular Tom Spits), '*de Petrus Dathenus School'* in Hilversum (in particular Tobias van Iwaarden), '*de Montessorischool'* in Hilversum (in particular Alex de Kruijff), '*het Futura College'* in Woerden (in particular Ciske Kokjes, Puck Delhee and Lia de Graaff), '*het Trivium College - locatie Mondriaan'* in Amersfoort (in particular Maartje van Kerckhoven) and *'College de Brink'* in Laren (in particular Claudio Castagna, Paulinka van Boven and Hermien).

A warm thank you also to all the wonderful student assistants that helped Jeannette and me in collecting our large dataset (or transcribing parts of it): Merel van Witteloostuijn, Leanne Matimba, Doatske de Haan, Sybren Spit, Kim Schoof, Irene Rademaker, Jorik Geutjes and Bart Siekman.

The Linguistics department of the UvA, the university context in which I performed my study, consists of many kind and brilliant people who helped me during my project, of whom I would like to mention some in particular. What should I have done without Dirk Jan Vet: our '*technische man die alles kan'* as I often described you to others. Thank you Dirk for all your help with the experiments and your always cheerful 'yes' when I knocked on your door with a new computer problem. Paul Boersma, I am grateful for the hours we spent searching for the best statistical methods to analyse my data. I do not believe I have ever met a person who can get so excited about numbers. Tuba Yarbay Duman, I would like to thank you for the collaboration in recruiting and selecting subjects, and the conversations we had. Olga Fisher, thank you for your support during the meetings to discuss my progress. Elly and Gerdien (and Brigitte, Suzanne, Marloes, Lilian and Emily) thank you for all the emails you sent, all the meetings you planned and for our pleasant conversations.

Performing my study would not have been as pleasant without the good company of my wonderful roommates and colleagues. Bibi, with you I spent more time than I ever spent with anybody else. Within one week after we started our projects, we had turned BH 3.41 into a little 'living room' by smuggling in a couch. I have always admired your luminous ideas and your tremendous energy (and the sometimes sudden lack of it). I witnessed you giving a perfect performance during your defence and I am glad to have you by my side as a paranymph. Tiffany, except for a Nespresso machine you brought much extra joy to our room. The small overlap in our research topics brought many interesting and stimulating discussions. In our 'new' room in the PCH we joined 'cosiness-forces' with Tessa, Merel and Imme. Thanks to the chats and lovely little coffee breaks with the four of you, finishing this manuscript was not hard at all.

A warm thank you also to all the other PhD's from the Linguistics department and the Dutch department, of whom many have become friends. Sophie and Margot, we graduated together in the research master Linguistics and as colleagues we frequently met for lunch or drinks. I truly appreciated the moments we shared our successes and struggles in science and in life and I hope our friendship will last.

Jing, Margreet, Konrad and Marlou – together with Bibi we started our projects at the same time and followed many courses together. Thank you for all the fun during the past five years and for being great examples of a scientific mother (Jing), an ambitious post-doc (Margreet), a successful grant-applicator (Konrad) and a great teacher and strong researcher (Marlou). Thanks also to all the other PhD's (former and current): Joke, Aude, Sterre, Karin, Vadim, Klaas, Mirjam, Jan Willem, Margarita, Katja C., Katja B. Sanne, Jelke, Maja, Heimir, Caroline, Patrick, Seid, Matthias, Camille, Caitlin, Brechje, Jeroen, Hernán, Rosa, Ulrika, Marloes and Jasmin. Having such nice colleagues to share lunch, drinks and celebrations with made the lonely life of performing a PhD so much more fun.

Within the field of research on language acquisition and language pathology, many people have provided feedback on presentations or helped me in other ways. I was very lucky to be part of the Grammar & Cognition research group. To all its members: a warm thank you for the interesting discussions on research and the informal atmosphere. Within the Netherlands, I furthermore would like to thank Angeliek van Hout for sharing her unpublished data on relative clause elicitation. Tessel Boerma, I want to thank you for the lovely discussions on SLI-topics and the fun we had during conferences. Rob and Margo Zwitserlood, we met at several conferences and I was always very happy to find out you were there. Thank you for your support and the good time we spent together in London during a workshop. Audrey Franssen, thank you for answering my questions on special education, and for co-authoring a book chapter on adolescents with SLI. Annette Scheper and Esther Parigger I would like to thank you for your encouragements to become a PhD and for inspiring me. I am not sure whether I would have been at this point without your example and encouragements.

I also would like to express my gratitude to numerous colleagues outside the Netherlands. In 2012, I visited Naama Friedmann in Tel Aviv for two weeks to work on a Dutch adaptation of the relative clause tasks she developed. Thank you Naama for your time and hospitality. I also would like to thank all members of the EUCLDIS group, in particular Richard Schwarz and Katrin Lindner, and all the members of the COST-action on Language Impairment in a Multilingual society. Those meetings provided me with useful information and many connections in the field. During my period in London in 2014, I made a short tour through England to visit language acquisition labs in other cities. Gina Conti-Ramsden, thank you for the opportunity to present my results at your lab in Manchester and for the lovely dinner we had with Ludovica Serratrice. The Manchester Language Study you and

Nikki performed has been of great value to this dissertation. Dorothy Bishop, it was an honour to present the results of a project that was largely based on one of your ideas in your lab in Oxford. Thank you for hosting me, and for being such a great example. Chloe Marshall and Shula Chiat thank you for meeting me at Exmouth market in London to exchange ideas over a cup of coffee or a dinner.

I also would like to acknowledge my colleagues of the Speech & Language Centre in Eindhoven, who were always interested and supportive during the years I combined my PhD with a clinical position. I sometimes miss our large team of clinical linguists (Annette, Naz, Ingrid, Esther, Annelies, Katja, Steffie and Wendy), the lunch breaks in the beautiful garden and the enriching multidisciplinary setting in which we could offer so much to children and their parents.

A special thank you goes to my dearest friends and family for reminding me there are other, more important things in life than just work. Helvi, Geer, Maaike and Suus, we've known eachother since we were little and our friendship is very comforting. I lost count of the weekends we spent together and the number of times we've uttered '*twee blauwe voor je ogen'* (which seems linguistically odd, but in our world makes perfectly sense). Thank you for all the fun, your interest and support. Maaike, your brave emigration to the other side of the world has not prevented us from staying connected. Thank you for always being there for me and for knowing me so well. Carola, thank you for the fact that we always seem to bond, no matter how long we haven't seen eachother. Eline, thank you for your friendship and for allowing me to be involved in some important moments in your life. Marlies, Linda, Fieke, Berber, Maya, Merle and Kris, thank you for sharing so much *'moeder-wel-en-wee'* in the past two years. Emma, Jessica, Maartje and Else, we met 12 years ago and your friendship has been valuable in many ways. Jessica, with you I started the Linguistics adventure. I hope we will be able to stay in touch. Maartje, you were a little professor already when you entered the Beta-Gamma bachelor. Thank you for your support and your strong opinions when I needed some advice. Else, I believe you are my most spiritually engaged friend. Your view on life often triggered me to reconsider whether my choices were truly mine. Emma, our friendship has many different facets and is highly valuable to me. Although we have different research fields, we shared many struggles during our projects. We have been running mates for 12 years and added joint ice skating lessons to the friendship list a few years ago. Somehow, it always feels like we are on the same track. I am therefore very glad to have you by my side as a paranymph. I would like to thank Gijs, Baucke, Caroline and Jelle for the marvellous editions of 'de Maaltijd'. My

relationship with Gerard also brought a whole new group of friends: the ‘Jan Arentz-gang en consorten’. Thank you for throwing so many good parties and being such an energetic group.
Mijn familie wil ik graag in het Nederlands bedanken, want je moeder in een andere taal dan je moedertaal bedanken voelt vreemd. Allereerst mijn lieve papa en mama: dank jullie wel voor het warme thuis dat jullie ons hebben geboden, en voor jullie grenzeloze betrokkenheid en liefde. Inhoudelijk hebben jullie dan misschien nog steeds niet zo’n goed beeld van wat ik tijdens die vijf jaar allemaal heb uitgespookt, maar jullie hebben altijd veel interesse getoond in wat het project met mij deed en naar welke plekken het me bracht. Ik weet zeker dat jullie trots zullen zijn, niet om de extra letters voor mijn naam maar om de weg die ik heb afgelegd. Dank ook aan mijn twee lieve broers, Ruben en Aram, die in de lengte en de breedte al lang niet meer mijn kleinere broertjes zijn. Ru, wat is het fijn dat je zo dichtbij woont en dat we elkaar vaak eventjes kort kunnen zien of spreken. De spontante avondjes met jou en Dorine zijn me heel dierbaar. Aram, wat is het bijzonder om jou in 20 jaar tijd van baby tot boom van een kerel te hebben zien opgroeien. De avond in Barcelona dat we voor het eerst met z’n drieën naar een club gingen omdat jij inmiddels ook (bijna) oud genoeg was koester ik. Dank ook aan Hanny en Rob, mijn lieve tante en oom die me in de laatste fase van mijn proefschrift wat extra schrijftijd hebben gegeven – en aan mijn lieve nichten Ranne en Yunka, voor jullie aanmoediging en de creatieve plannen om drukkosten voor het proefschrift te besparen. Naast mijn eigen familie is er ook een ander familie waar ik inmiddels officieel tot ben toegetreden en waar ik me helemaal thuis bij voel: Katy, Gert Jan, Conny, Jafeth, Esther, Peter, Dolf, Linda en Iris, dank jullie wel voor jullie interesse en steun tijdens de afgelopen jaren, en voor alle gekkigheid bij jullie thuis.

In het laatste jaar van mijn promotie ben ik zelf moeder geworden van een lief, temperamentvol, vrolijk meisje. Lieve Livia, dank je wel voor het extra geluk en de vreugde die je hebt meegebracht. Het duurt waarschijnlijk niet lang meer voor je dit zelf kunt lezen, want je lijkt een aangeboren interesse in boeken te hebben. Ik hoop dat je altijd zo gedreven en levenslustig zult blijven als je nu bent.

Last but certainly not least, I want to thank you, my dear Gerard, for your love! I once blamed the partner of a former colleague for not displaying any interest in what he was investigating in his PhD study. To me, a lack of interest seemed like a lack of support. You have shown me that very little interest in my topic could coincide with endless support. Thank you for being my opposite.

1 Introduction*

Wexler (1998) once claimed that children can informally be described as ‘little inflection machines’, since they seem to learn the inflectional rules of their language efficiently and very quickly in their first years of life. The description is a striking characterization of the way typically developing children acquire their first language, but for some children, the machine does not seem to work properly. They have profound problems in learning or applying the linguistic rules of their language, although there is no clear aetiology for these difficulties. This thesis focusses on this group of children, often referred to as children with Specific Language Impairment (SLI) (Leonard, 1998; 2009).

In the last decades, it has become clear that language impairments in children often persist into adolescence and adulthood. Longitudinal and follow-up studies indicate that the majority of identified cases of SLI in childhood still meet the criteria for language impairments in adolescence (Leonard, 2009). In childhood, grammatical problems are often seen as a core characteristic among the range of language problems that are displayed in SLI. However, relatively few studies discuss whether the nature of the problems remains the same as children grow older. The first aim of this study is to investigate the persistence of grammatical differences between individuals with SLI and typically developing (TD) peers.

Several theories exist as to the origin of grammatical differences between SLI and typically developing children. Grammatical errors in SLI are often interpreted as an indication of a lack of grammatical knowledge. Children with SLI are therefore hypothesized to have difficulties deducing the linguistic rules of their language. An alternative interpretation of grammatical mistakes could, however, be a problem in implementing knowledge in performance (Bishop, 1994). The second aim of this thesis is to investigate whether persistent grammatical difficulties in SLI originate

* Parts of this chapter were published in Duinmeijer (2013). Persistent problems in SLI: which grammatical problems remain when children grow older?, *Linguistics in Amsterdam, 6*, 28-48.

solely in the acquisition of grammatical knowledge or can also be attributed to problems in knowledge implementation.

In this first chapter, language impairment in individuals with SLI is defined and characterized in more detail (§1.1) and some theories are described that have been developed to account for SLI (§1.2). From this description, the general aims of this thesis will emerge; these will be discussed in more detail in §1.3. The structure of this book will be outlined in §1.4.

1.1 Definition, prevalence and characterization of SLI

As was mentioned above, children with SLI show problems in language development in the absence of a clear aetiology. The absence of a clear cause for the language problems is an important aspect of the diagnosis of SLI. Children are only considered as having SLI if their hearing is normal, non-verbal cognitive abilities are within the normal range (above 85) and there are no cognitive deficits, socio-emotional problems or behavioural disorders that can explain their language difficulties (Stark & Tallal, 1981; Leonard, 1998). The diagnosis is thus mainly based on exclusion criteria, the single inclusion criterion being a significant deviance in language acquisition in comparison to typical development.

Over the last decades and particularly in the past few years, the label SLI has been discussed since the group it is supposed to designate is unclear (Bishop, 2014; Ebbels, 2014; Reilly et al., 2014). There are several reasons for this. First of all, the same children might receive different labels, since other terms like 'developmental language disorder' or 'primary language impairment' are often used for children meeting the criteria for SLI.[1] Furthermore, the label denotes different children depending on where or by whom they are tested or diagnosed. Diagnostic criteria differ across contexts and countries. Inclusionary cut-off values for language scores range, for instance, from -1 to -2 standard deviations below age expectations (Tomblin et al., 1997).[2] Exclusion criteria are not uniformly applied either. The presence of some neurodevelopmental disorders - such as Autism Spectrum Disorder (ASD) or Attention Deficit Hyperactivity Disorder (ADHD) - does, for example, no longer leads to automatic exclusion because the link between these disorders and SLI is unclear (Bishop, 2014). Since the label SLI does not generate a clearly defined clinical group (Ebbels, 2014), a comparison of results between studies is complicated. Despite all the (valid) criticism of the label SLI, an

[1] Bishop (2014) provides an overview of all labels used to denote unexplained language difficulties, as well as a discussion on their validity.
[2] See §3.1 for more details on the criteria used in this thesis.

alternative label has not yet been agreed on and SLI is still the common term used in the literature. It will therefore be used in this thesis.

Due to the disagreements on inclusion and exclusion criteria, studies estimating the prevalence of SLI come up with different numbers. A review study indicates that percentages range from 0.6% to 33.6% in children up to age 7 years of age (Law, Boyle, Harris, Harkness, & Nye, 1998). Tomblin and colleagues (1997) performed a large study (N=2084) among the population of American school-age children and found that 7.4% of children matched diagnostic criteria of SLI (in that study -1.25 SD on two out of five composite language scores). This percentage is often taken as a kind of international standard for the occurrence of SLI. Viewed from a statistical perspective, this means there is at least one child with SLI in every classroom. Nevertheless, language problems are not always recognized as such. In the population studied by Tomblin and colleagues, only 29% of the parents with a child that met the criteria for SLI reported to have been previously informed about their child having a speech or language problem (Tomblin et al., 1997). Clinical referral is therefore not automatic.

Longitudinal and follow-up studies indicate that, in the majority of cases, language problems are persistent into adolescence. Differences in language performance between SLI and TD persist. The studies of the identified cases of SLI in childhood show that around 70-80% meet the criteria for language impairments in adolescence (e.g., Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998; Johnson et al., 1999; Botting, Faragher, Simkin, Knox, & Conti-Ramsden, 2001; Conti-Ramsden, Botting, Simkin, & Knox, 2001a; Nippold & Schwartz, 2002; Clegg, Hollis, Mawhood, & Rutter, 2005; Law, Rush, Schoon, & Parsons, 2009; Tuller, Henry, Sizaret, & Barthez, 2012). Even when language problems seem to have disappeared in childhood (Bishop & Edmundson, 1987), they can re-appear at a later age in adolescence (Stothard et al., 1998). The fact that the number of children receiving a diagnosis decreases with age is partly due to the fact that language milestones are less clearly distinguishable in adolescence (Nippold, 1995). Furthermore, individual differences become larger in adolescents (Nippold, 1995; Reed, 2005) and standardized tests are scarce. It must be concluded that existing figures on the number of adolescents with language impairments rarely reflect the actual situation (Larson & McKinley, 1995).

In the Netherlands, the prevalence and persistence of SLI is also difficult to assess.[3] Children with an official diagnosis of SLI either go to special education schools or attend regular schools while receiving ambulatory support. The number

[3] In the Netherlands, the disorder is labelled 'Specifieke Taalontwikkelingsstoornis (S-TOS)'.

of pupils in special education is reported only once a year after the summer break, while figures of ambulatory care appear throughout the year. Statistics on the number of language impaired pupils are therefore estimates. In the transition from primary to secondary school many children with SLI make a move from special education to regular education with ambulatory support. For this age group, language impairments were, until 2014, grouped together with hearing impairments so that governmental reports on secondary special education only contain collapsed numbers (CBS, 2016).[4]

In 2013/2014, according to official figures the percentage of pupils in primary education who had a language impairment and were placed in special education was 0.37%. Another 0.28% received ambulatory care. Of all pupils in primary education, the percentage of children with an official diagnosis of a language impairment was thus around 0.65%, a figure at the lowest end of the range in the review study of Law and colleagues (1998). In secondary education, this percentage is reduced to around 0.39% (0.21% in special education and 0.18% with ambulatory care) (calculations on the basis of the numbers provided by the CBS, 2016). These low percentages suggest that many cases of SLI are not officially diagnosed. This may be because the inclusion criteria for a diagnosis in the Netherlands are quite strict compared to the criteria used by Tomblin and colleagues (1997) in the United States (see §3.1 for the Dutch inclusion criteria that were used in this thesis). That does not mean that pupils with language impairments who do not have an official diagnosis are not receiving any form of help. They most probably receive therapy from a speech and language therapist (for which an official diagnosis is not required) and they may be given extra instruction or individual help at school.[5]

The language impairments seen in individuals with SLI are heterogeneous, both in terms of the domains in which language problems occur and in terms of the severity of the problems. Problems occur in all domains of language (phonology, morphology, syntax, semantics and pragmatics). It regularly takes children with SLI longer to acquire the phoneme inventory of their mother tongue, and phonological

[4] In 2014, the system of special education has been drastically reorganized and statistics on the number of pupils with language impairments are not provided yet. We therefore base our estimation of the number of Dutch children and adolescents with a diagnosis SLI on the reports from 2013/2014 (CBS, 2016).

[5] In some individuals meeting the criteria for SLI, the language problem may be unnoticed or labelled otherwise (partly due to the strict exclusion criteria for SLI). A language problem might for instance show as a behavioural problem, since the discrepancy between cognitive abilities (non-verbal IQ) and language abilities often leads to frustration (Conti-Ramsden & Botting, 2004).

reduction processes can be perseverant over a long period (Fee, 1995; Roberts, Rescorla, Giroux, & Stevens, 1998). Furthermore, many individuals with SLI have problems in word learning, resulting in smaller vocabularies and word finding difficulties (e.g., Leonard, 1998; McGregor, Newman, Reilly, & Capone, 2002; Hick, Botting, & Conti-Ramsden, 2005). The ability to tell a story and to communicate effectively is often also reported to be impaired (Richardson & Klecan-Aker, 2000; Duinmeijer, de Jong, & Scheper, 2012). Most widely attested are, however, the problems in the grammatical domain, resulting in the omission or substitution of grammatical markers or errors in syntactic structures.

Many authors have tried to come up with classification systems to define subtypes within SLI (e.g., Conti-Ramsden, Crutchley, & Botting, 1997). The classifications are difficult to compare, due to differences in the classification criteria (linguistic characteristics, clinical observations or statistical methods). Some classifications are very general distinguishing only between an expressive disorder, in which children have adequate comprehension but problems in production, and a mixed receptive/expressive disorder, in which comprehension is also impaired (Rapin, 1996). Subgroups have also been defined on a more detailed level, using linguistic characteristics or clinical observations (Rapin & Allen, 1983; Bishop, 2004).

The largest group of children with SLI has problems in the domains of phonology and morphosyntax. In the existing classification systems of children with SLI, those children with such grammatical problems are often classified as ‘typical' or 'classical’ SLI (Bishop, 2004) and problems in the grammatical domain are often described as a core characteristic of SLI. The grammatical problems reported in SLI are quite extensive. Children are described as having difficulties in the comprehension and production of inflectional or derivational morphemes, clitics, free morphemes, complex syntax and non-canonical sentence structures (Dromi, Leonard, & Shteiman, 1993; de Jong, 1999; van der Lely & Battell, 2003; Hamann, 2006; Stavrakaki, 2006; Orgassa & Weerman, 2008; Schwartz, 2009; Contemori & Garraffa, 2010). Some grammatical aspects have been identified as 'key' to the language impairment, in the sense that all children with SLI seem to struggle with these aspects and the struggles do not simply reflect a delay but indicate a deviant developmental pattern. These key aspects are often called ‘clinical markers’, for they can be used to identify those children with language impairments. Which grammatical aspects can serve as clinical marker is, however, partly language-dependent. Verb inflections are, for instance, more problematic in languages with a sparse morphology than in languages with a rich morphology (Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992) (see Chapter 6 for a further discussion of this issue).

Many studies on the grammatical abilities of children with SLI note that there is not only considerable variability in scores between children, but also within children (Masterson & Kamhi, 1992; Bishop, 1994; Marchman, Wulfeck, & Weismer, 1999; Marshall & van der Lely, 2007; Song, Sundara, & Demuth, 2009; van Ewijk & Avrutin 2010; Weerman, Duinmeijer, & Orgassa, 2011; Keij, 2009). In 1991, Gopnik and Crago observed, for instance, that children with SLI omit grammatical features, but not consistently. They sometimes produce them correctly and sometimes omit them altogether. Bishop (1994) also noted variability in the correct use of English morphological markers in SLI (e.g., plural marking *s*-, or past tense verb inflections). However, the occurrence of errors is not random. Verb inflections are, for instance, produced correctly more often in certain phonological contexts than in others (Song et al., 2009). Performance on grammatical aspects in SLI thus seems to be (partly) dependent on the linguistic context (Marchman et al., 1999; Grela & Leonard, 2000; Roulet-Amiot & Jakubowicz, 2006; Marshall & van der Lely, 2007; Keij, 2009; Song et al., 2009; Weerman et al., 2011).

As has been set out above, grammatical problems are seen as a hallmark of younger children with SLI. If the language disorder persists, as it appears to do in the majority of cases, we do not yet know whether the predominant vulnerability in the grammatical domain remains the same. Most studies on older children with SLI discuss whether children still meet diagnostic criteria, based on general language tests. Very few studies have discussed whether adolescents with SLI still struggle with the specific grammatical aspects that are characteristic for SLI in childhood. A first, general aim of this thesis is therefore to investigate whether grammatical problems persist in individuals with SLI and which problems these are.

1.2 Theoretical accounts of SLI

Since the start of the research into SLI, explanations have been sought for these language problems. This can be done at different levels, as has been visualised in the causal model of developmental disorders by Bishop and Snowling (2004) shown in Figure 1.1 In this model a distinction is made between observed behaviour, cognitive processes, neurobiology and aetiology (genetic and environmental factors). Furthermore, the model illustrates the fact that various causes can underlie the same observed behaviour and vice versa, the same underlying cause can result in different observed phenotypes. It may therefore not be sufficient to search for explanations at one level only.

Although SLI is by definition not characterized by a clear aetiology (the first level in Figure 1.1), recent findings suggest a strong genetic basis for the disorder (e.g., Bishop, North, & Donlan, 1995; Bishop, 2008). Familial aggregation studies report higher rates of language impairments in parents or siblings of children with SLI (Leonard, 1998; Barry, Yasin, & Bishop, 2007; Tomblin, 2009). Familial aggregation does, however, not automatically mean that the effect is genetic, since family members often share environmental factors as well. A stronger indication of genetic influences is provided by twin studies in which one twin has SLI. Monozygotic twins, who are genetically identical, appear to show higher concordance rates of language abilities than dizygotic twins (Bishop et al., 1995; Bishop, 2008). Problems in language development thus seem to be inheritable. Although molecular genetic studies identified some candidate genes with strong effects on speech and language (e.g., Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998; Tomblin, 2009), it still remains unclear which combination of genes and environmental factors determines the aetiology of SLI.

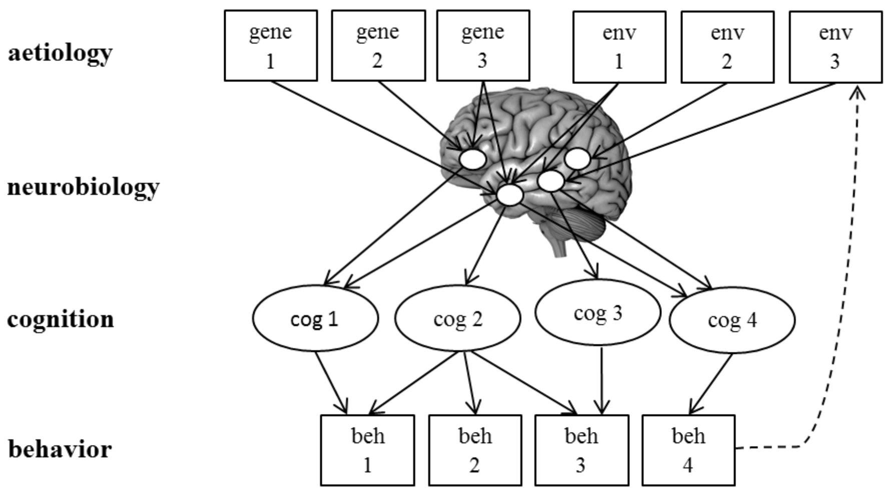


Figure 1.1. Causal model of Bishop and Snowling (2004, page 85) that shows the levels of causation for developmental disorders. The dashed line emphasizes that children’s behaviour (beh) can affect the environment (env) they experience

The same holds for studies focussing on the neurobiological level (the second level in Figure 1.1), in which structural and electrophysiological differences in language disordered children are measured. If the language impairment could always be linked to a localized neurobiological deviance – for example a focal lesion or a clear electrophysiological deviation - the label SLI would not be applied (Tropper & Schwartz, 2009). Rather, SLI is characterized by diffuse neurobiological anomalies that presumably originate in an abnormal (prenatal) maturation of the brain, driven by genetic and environmental factors. As a result, the brain may not be optimally configured for (language) learning processes (Tropper & Schwartz, 2009). Post-mortem studies and studies using neuroimaging techniques have shown several structural differences in the brains of individuals with SLI, such as a deviation from the standard cerebral asymmetry in the perisylvian region (Cohen, Campbell, & Yaghmai, 1989; Plante, 1991) or smaller brain volumes in subcortical regions (Jernigan, Hesselink, Sowell, & Tallal, 1991). Structural differences are, however, not found in all affected individuals, and non-impaired individuals sometimes show the same brain patterns. The reported structural differences may therefore be linked to problematic language abilities but cannot be regarded as the cause for language disorders (Tropper & Schwartz, 2009). Similarly, brain activation studies have shown several different activation patterns in SLI, such as in the amount of activation in certain areas or in the timing of activation (Leppänen, Lyytinen, Choudhury, & Benasich, 2005; Tropper & Schwartz, 2009). So far, language disorders cannot be linked to a clear difference in brain structure or brain activity.

Most theories on the origin of SLI are still based on behavioral findings and are descriptions of hypothesized underlying cognitive deficits (the two lowest levels of Figure 1.1). In the description of the (language) behaviour of children with SLI, two groups of accounts can be distinguished. The first group is often denoted as the *representational accounts*, because the origin of the problems is sought at the level of the representation of linguistic knowledge, i.e. in the language faculty itself. Representational accounts have in common that they explain why problems occur in particular linguistic aspects and not in others. The fact that children with SLI from Germanic language backgrounds seem to have disproportional problems with verb inflections, for instance, led to the development of several theories on the linguistic locus of these problems. Rice and Wexler (1996) accounted for these problems by proposing that children with SLI have a prolonged period in which tense marking is optional, while Clahsen, Bartke and Göllner (1997) interpreted inflection errors as a problem with features that mark agreement relations. Gopnik (1990) even assumed children with SLI are 'feature blind'. The locus of the linguistic deficit in SLI thus has varied across different representational accounts but they have shared the idea that a linguistic principle is either missing or matures late in SLI, which explains why specific problems show up in the grammatical domain.

The accounts mentioned above were developed during a time in which research on SLI aimed at pinpointing the specific linguistic problems characterizing the language problems in SLI (de Jong, to appear). It may therefore not be surprising that the accounts are quite specific, and cannot be generalized to problems seen in other linguistic domains. Even within the domain of grammar, several problems are attested in SLI that may not serve as clinical markers but nevertheless need to be explained in an account of SLI (de Jong, to appear). Furthermore, the theories mentioned above were based on clinical markers of SLI in Germanic languages, and cannot explain the profile of SLI in other languages, in which problems in verb morphology are not necessarily a core feature.

More recently developed representational theories try to explain a broader range of linguistic difficulties in SLI as well as cross-linguistic differences in the linguistic profile of SLI. The Representational Deficit for Dependent Relations (RDDR) Hypothesis (van der Lely & Battell, 2003), for instance, assumes prolonged optionality of the rule that drives movement in SLI; this is proposed because many of the characteristic problems in SLI involve movement. Similarly, the Computational Grammatical Complexity (CGC) Hypothesis (van der Lely, 2005) accounts for several problems in the phonological, the syntactic and the morphological domain by assuming that children with SLI have a problem in the representation of structures that require hierarchical organization (Marshall & van der Lely, 2007).

Over the past decades, however, it has been questioned whether the problem in SLI is purely linguistic. A growing body of research indicates that children with SLI have problems in different aspects of information processing, both verbal and non-verbal. Problems are found in memorizing and repeating different sorts of information (Aram, Ekelman, & Nation, 1984; Im-Bolter, Johnson, & Pascual-Leone, 2006; Marton, 2008; Windsor, Kohnert, Loxtercamp, & Kan, 2008; Schwartz 2009; Henry, Messer, & Nash, 2012; Jensen de Lopéz & Baker, 2015, Lukács, Ladányi, Fazekas, & Kemény, 2016), although not all aspects of information processing are impaired to the same extent. Problems in processing abilities seem to be persistent over time, especially in the auditory domain and in more complex processing tasks (Hick et al., 2005; Spaulding, Plante, & Vance, 2008,).

Another group of explanations is based on these more recent findings of problems in information processing and can be referred to as the *processing accounts*. In this type of account, the problems in SLI are thought to stem from an information processing problem, while the language faculty itself is intact. Some theories propose a quite specific problem in the processing of auditory or phonological information (e.g., Gathercole & Baddeley, 1990; Chiat, 2001; Gathercole, 2006) or in speech perception (Leonard, 1989; Joanisse & Seidenberg, 1998). Problems in speech perception would, for instance, explain why verb inflection errors in SLI are language dependent, since the saliency of verb inflection morphemes differs across languages (Leonard, 1989). Others assume a broader deficit in the processing of both linguistic and non-linguistic information (e.g., Kail, 1994; Windsor & Hwang, 1999; Weismer, Plante, Jones, & Tomblin, 2005).

Weismer and colleagues (2005) tested whether individuals with SLI showed differences in brain activation in neural systems involved in language processing and in general processing. The neuroimaging results pointed towards a general processing limitation in SLI because less activation was not only found in regions associated with linguistic processing but also in regions that reflect general processing. The processing accounts share the assumption of a processing deficit (either general or specific) that causes inadequate or incomplete processing of the input. This makes (linguistic) notions that require a great deal of input in order to be acquired particularly vulnerable for SLI. Grammatical aspects that are acquired relatively late in typical language development, either because they require a large amount of input or because the aspect becomes important in a later stage of language development, are therefore hypothesized to be problematic for children with SLI. If the delay in acquisition caused by the reduced intake extends the critical period (and early sensitive period for language acquisition), a child with SLI may not even acquire the feature or rule at all.

A quite prominent account of SLI that tries to account for both linguistic and non-linguistic problems in SLI but cannot strictly be labelled as a processing theory is the Procedural Deficit Hypothesis (PDH) proposed by Ullman and Pierpont (2005). In this account, SLI is regarded as an impairment in the 'procedural memory system' that underlies the learning and execution of motor and cognitive skills, while 'declarative memory', underlying the learning, representation and use of factual knowledge and episodic knowledge, is spared. The procedural system is hypothesized to be localized in the frontal cortex and the basal ganglia, in which anomalies in SLI have been attested (Ullman & Pierpont, 2005). Furthermore, grammatical skills are hypothesized to be dependent on procedural learning while lexical abilities are dependent on declarative memory, which is in accordance with the disproportionate difficulty with grammatical aspects in SLI (Ullman, 2001). Although the theory is promising in the fact that it accounts for both linguistic and non-linguistic symptoms and the specific vulnerability of the grammatical domain, evidence on the proposed duality (impaired procedural memory and intact declarative memory) in SLI is not conclusive (Poll, Miller, & van Hell, 2015). Whether procedural memory is impaired seems to depend on the aspects of procedural memory that are tested (Hsu & Bishop, 2014). Individuals with SLI also seem to have deficits in declarative memory (Lum, Gelgic, & Conti-Ramsden, 2010).

Few of the theories discussed above can explain the inconsistencies in the grammatical performance of children with SLI. Some theories account for variability in performance by assuming optionality of rules, or lack of accessibility to linguistic principles (Gopnik & Crago, 1991 in their Feature Deficit Hypothesis, and van der Lely & Battell, 2003 in their Representational Deficit for Dependent Relations (RDDR) Hypothesis). These theories, however, predict variability in performance in SLI to be random. This is not in line with the studies reporting that linguistic context influences performance (Marchman et al., 1999; Grela & Leonard, 2000; Roulet-Amiot & Jakubowicz, 2006; Marshall & van der Lely, 2007; Song et al., 2009; Weerman et al., 2011; Keij, 2009).

A theory that explains the inconsistencies in performance in SLI and also accounts for the fact that variability is not random is the Vulnerable Markers Hypothesis (VMH) (Bishop, 1994). The hypothesis was based on the observation that the errors in children's production of plural markers on nouns and past tense marking on verbs are consistent in the type of error (nearly all omissions of an inflection) but not consistent in the occurrence of the error (sometimes the inflection is produced and sometimes it is omitted). Rather than trying to account for inconsistencies in terms of incomplete grammatical knowledge only, Bishop proposed the idea that inconsistencies can also stem from a problem in implementing knowledge in performance.

In the VMH, a processing problem in SLI is assumed, due to a limited capacity system that has to handle several operations in parallel – for instance in the decoding of a message at different linguistic levels, or in processing and storing information at the same time. Such processing problems cause problems in the acquisition of grammatical knowledge. At the same time, however, the hypothesis assumes that the same processing problem hampers the implementation of grammatical knowledge, both problems leading to grammatical errors. The VMH is schematically illustrated in Figure 1.2.

According to the VMH, systematic differences are predicted between the contexts in which errors occur and in which the utterances are produced correctly. The hypothesis predicts that "errors will occur when the speech production system is stressed by the need to produce output that makes heavy demands on its processing capacity" (Bishop, 1994, page 528). This may, for instance, happen when the message that has to be conveyed is more complex or contains more complex sentences, or when the retrieval of lexical items requires more effort. Bishop argues

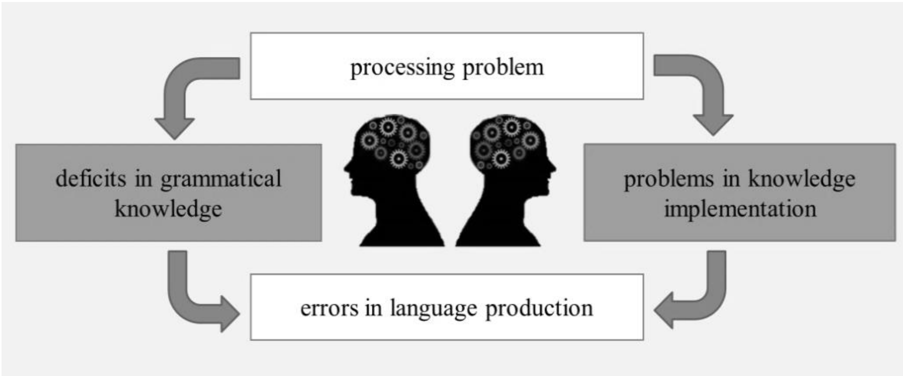


Figure 1.2. Schematic illustration of the Vulnerable Markers Hypothesis (VMH) (based on Bishop, 1994)

that experimental studies are needed to investigate under which conditions grammatical forms are affected.

Together with examining the persistence of grammatical problems in SLI, this thesis will further explore Bishop's idea that inconsistency of grammatical errors in SLI may be caused by a processing problem that not only affects the acquisition of grammatical knowledge but also the implementation of grammatical knowledge.

1.3 General research questions

In the previous sections, we already introduced the two important aims of this study. First of all, this study aims to gain insight in the persistence of the problems that are often regarded as characteristic for SLI in childhood. The first general research question is therefore: 'Do the grammatical differences between SLI and TD peers that are characteristic in childhood persist into adolescence?' This question is relevant from both a clinical and a scientific perspective. Clinically, knowledge about the persistence of grammatical problems can inform decisions on treatment and training of children and adolescents with SLI in clinical and educational settings. From a scientific point of view, insight into the persistence of grammatical problems may shed light on the question whether the grammatical problems in childhood are maturational, as some of the theories discussed above claim.

The second aim of this thesis is to test whether language problems in SLI can be explained by a processing deficit that affects both the acquisition of grammatical knowledge as well as its implementation in performance - an idea put forward by Bishop as the Vulnerable Markers Hypothesis (VMH) (1994). These two aims are not randomly combined but are intertwined in this study. We assume that the investigation of the assumptions and predictions of the VMH can be tested best in an older population of adolescent subjects in which grammatical knowledge has stabilized. Although adolescents might still acquire grammatical knowledge, their access to implicit learning mechanisms is assumed to be limited. Differences between knowledge (grammatical knowledge) and performance (knowledge implementation) might therefore show up more clearly in adolescence than in childhood, when grammatical acquisition is still very much in development.

The aim to test the assumptions and predictions of the VMH led to the formation of the second and the third research question. As was discussed above, the VMH predicts variability in the language output of individuals with SLI in linguistic contexts that involve a larger processing load (Bishop, 1994). The second question of this thesis is therefore whether any variability in the grammatical performance of children and adolescents with SLI is dependent on the linguistic context in which grammatical aspects are tested. Furthermore, the VMH hypothesizes that variability in performance in SLI can be explained by poor processing abilities in SLI. It predicts a correlation between processing capacities and the amount of variability. The third and final question of this dissertation is therefore whether variability in performance is correlated with (impaired) processing abilities. For this final question, it is important to verify an important assumption of the VMH, namely that individuals with SLI show difficulties in information processing that persist into adolescence.

To sum up, this thesis will address the following research questions:

1. Do the grammatical differences between SLI and TD peers that are characteristic in childhood persist into adolescence?

2. Is there variability in the grammatical performance of children and adolescents with SLI, and can this variability be explained by factors related to linguistic context?

3. Is any variability in linguistic performance related to (impaired) processing abilities?

In order to answer these questions, grammatical performance was tested in an experimental setting, with different tasks and in varying contexts. Three grammatical variables were chosen, which have previously been found to be vulnerable in Dutch SLI: subject verb agreement, grammatical gender, and relative clauses.

1.4 Structure of this book

This thesis is organised in seven further chapters. In Chapter 2, the literature on Dutch SLI and on the persistence of problems in SLI is reviewed in order to define the empirical background in which this study should be placed. The emphasis is on the studies discussing grammatical abilities and the choice for the three grammatical variables tested in this thesis (subject verb agreement, grammatical gender, and relative clauses) is elaborated on. Chapter 2 also elaborates on the operationalization of the concept 'complexity', and discusses which linguistic factors were incorporated in the tasks as potentially influencing processing load. The chapter finishes with the formulation of hypotheses related to the research questions of this study presented in the previous section.

Chapter 3 presents the methodology used to answer the research questions. Development over time was tested cross-sectionally by comparing four groups of children: a younger and an older group both with and without SLI. An adult group was added to check the validity of our tasks. The chapter provides information on the methods and criteria of selection, the general language abilities of the groups and the outcomes of the parental questionnaire that addressed the developmental history of the child. Besides a detailed description of the participants, this chapter also describes what type of tasks were used to examine processing capacities and grammatical abilities. The general procedures of administration, transcription, scoring, coding and analyses are briefly discussed.

Chapter 4 discusses the results on information processing. As discussed above, this plays a special role in this study since problems in information processing are hypothesized. The chapter starts with a brief discussion of models of information processing and a summary of what is known on processing abilities in SLI. Afterwards, the tasks selected for the measurement of processing ability are described, followed by the results. The chapter ends with a conclusion and discussion of the outcomes in the light of our hypotheses for the grammatical variables to be discussed in Chapters 5, 6 and 7.

Chapter 5 describes the first grammatical aspect tested, namely grammatical gender. The chapter starts with a discussion of the Dutch paradigm for grammatical gender and provides a summary of what is known on the acquisition of grammatical gender in TD children and children with SLI. This leads to a specification of the general hypotheses with respect to grammatical gender. The two tasks are described and the outcomes of the comprehension and production of grammatical gender in determiners and adjectives are discussed. Furthermore, the results are linked to the outcomes on the processing measures. The chapter concludes with an evaluation of the hypotheses.

Chapter 6 continues the description of grammatical abilities in children and adolescents with SLI with respect to subject verb agreement. Again, the chapter starts with an overview of the Dutch paradigm and then discusses what is known on the acquisition of this paradigm in TD learners and in individuals with SLI. On the basis of this discussion, specific hypotheses regarding subject verb agreement are formed. Three tasks designed to test knowledge of and performance on this grammatical aspect are described in detail. A discussion of the results follows and then in the final section, these results are related to the processing measures from Chapter 4 and evaluated with respect to the hypotheses.

Chapter 7 presents the final grammatical aspect tested in this thesis, that is relative clauses. The chapter starts with a discussion of their general characteristics, before zooming in on the characteristics of the Dutch relative clause. The literature on the acquisition of relative clauses in TD and in SLI is summarized, and specific hypotheses are formulated. The three tasks constructed to test knowledge of and performance on relative clauses are described in detail and the results are presented and linked to the processing measures. The chapter ends with a discussion of the results in the light of the research questions of this thesis and evaluates the hypotheses.

In the final chapter, the outcomes on processing abilities and the three grammatical variables are combined and conclusions regarding the research questions of this study are drawn. The chapter concludes with a discussion of the contribution of this study to the clinical and scientific field and suggestions and implications for future research are reviewed.

2 Background and hypotheses*

In order to formulate specific hypotheses in relation to the research questions presented in Chapter 1, background information on some topics is required. The first research question of this thesis is whether the grammatical differences between SLI and TD that are characteristic in childhood remain in adolescence. Since we will be answering the question for a Dutch population of children with SLI and since grammatical problems are language specific, it is important to describe their morphosyntactic profile in more detail. Furthermore, in order to be able to put the results of this study in context, the existing literature on the persistence of grammatical problems in SLI will be presented in more detail. These two topics will be covered in §2.1 and §2.2.

The second research aim of this thesis is to check whether there is variability in the grammatical performance of children and adolescents with SLI, dependent on the complexity of the linguistic context. In order to specify hypotheses related to this second research question, the concept complexity will be defined and the linguistic factors incorporated in the tasks will be presented (§2.3). The third research question does not require elaboration or definition at this point. The hypotheses are formulated in the final section of this chapter (§2.4).

2.1 Morphosyntactic profile of Dutch children with SLI

As was mentioned in Chapter 1, the grammatical profile of SLI is partly language-dependent. Since the literature on SLI has been dominated by studies on English subjects, descriptions of the morphosyntactic symptoms of SLI are often based on English typology. Dutch differs in a number of typological aspects from English that may significantly influence the Dutch grammatical profile of SLI. For instance, unlike present-day English, Dutch has a nominal gender system that is reflected in

* Parts of this chapter were published in Duinmeijer (2013). Persistent problems in SLI: which grammatical problems remain when children grow older?, *Linguistics in Amsterdam, 6*, 28-48.

several grammatical elements that accompany or refer to the noun. Similarly, conclusions on the dominant type of errors in SLI may be dependent on the type of language studied. In English, children with SLI are, for instance, often reported to omit inflectional markers like subject-verb agreement markers. However, the English verb inflection paradigm in present tense only distinguishes third person singular *–s* from bare stems. Omission and commission errors are hard to distinguish in such a system. In languages with more morphological distinctions in the verbal paradigm, such as Dutch, the dominant error type in SLI may therefore be different. In this section, we will briefly describe the morphosyntactic profile of Dutch children with SLI and we will discuss our choice for the three variables in this study.

An early study of the morphosyntactic profile of Dutch children with language impairments was conducted by Bol and Kuiken (1988). They developed a grammatical analysis tool for spontaneous speech (GRAMAT, a Dutch version of the LARSP, Crystal, Fletcher, & Garman, 1976) and described the grammatical development of children with SLI (N=14) between 4;2 and 8;2 years as opposed to TD children (N=12). The analysis tool provides a profile of the number of morphological markers, clause types and phrases structures typically used. From these analyses, it became clear that the children with language disorders had more unanalysable utterances, produced fewer utterances in general and used more simple sentence structures than their typically developing peers. At the word-level, the group with language disorders had significant problems with the use of personal and possessive pronouns. Furthermore, they used grammatical morphemes (diminutives and first person singular verb inflections) less frequently than their TD peers. Although that study provided a first indication of the morphosyntactic difficulties in Dutch children with language impairments, their analysis had disadvantages: grammatical errors were not considered and no information was provided on whether grammatical morphemes were realized in obligatory contexts.

De Jong (1999) studied the grammatical profile of children with SLI in more detail. He looked at inflectional morphology and argument structure in the elicited narratives of Dutch children with SLI (N=35, mean age 7;8). In accordance with the results of Bol and Kuiken (1988), Dutch inflectional morphology appeared to be a problematic area in children with SLI. Verb inflections were omitted (root infinitives) or lacked number agreement. Furthermore, children had problems with past tense marking. Besides problems in verb morphology, children with SLI seemed to favour verb argument structures with low levels of complexity. Studying the data of Bol and Kuiken in more detail, de Jong noticed that children with SLI less often express two arguments after a verb (ditransitive structures), and used more verbs that did not have any internal argument (intransitive structures). Children that made more verb inflection errors appeared to have lower levels of complexity in verb argument structure. De Jong thus concluded that inflectional errors characterize Dutch children with SLI, and that grammaticality should be considered in combination with complexity.

More recently, Zwitserlood (2014) looked at the most prominent morphosyntactic errors in the narratives of children with SLI and studied their developmental trajectories. A group of 30 children with SLI (age 6 at initial testing) were tested three times at 12-month intervals. Differences between SLI and TD peers mainly appeared in the grammatical correctness of the utterances, while fewer differences were found in utterance complexity. At all three moments of testing, children with SLI performed worse than both an age-matched and a language-matched group of TD children. Differences were found in verb morphology as well as in other morphosyntactic aspects (e.g. errors in pronouns, determiners, word order etc.). Verb-related errors showed a different developmental trajectory than errors in other morphosyntactic aspects in children with SLI. Error rates in verb morphology decreased till age 7, but then stagnated whereas other morphosyntactic errors decreased only after age 7.

Problems in verb morphology were confirmed by other recent studies on grammatical problems in Dutch SLI. Two perception studies indicated that children with SLI are less sensitive than their TD peers to errors in subject-verb agreement (Rispens & Been, 2007; Blom, Vasić, & de Jong, 2014). In a study of the narratives of 24 children with SLI (aged 7 and 9 years) and age-matched TD peers, clear group differences were found in the production of verb inflection for tense and agreement (Steenge, 2006; Verhoeven, Steenge, & van Balkom, 2011). Orgassa (2009) investigated subject-verb agreement in an experimental study with SLI between 6 and 8 years of age and also found clear group differences in the number of inflectional errors. Blom, de Jong, Orgassa, Baker and Weerman (2013a) suggest that present tense verb inflection errors can be used as a clinical marker with fairly high sensitivity and specificity (87% and 75% respectively).[1]

Children with SLI also seem to use strategies to avoid inflection (and movement) of lexical verbs, by using dummy auxiliaries and infinitives. The use of infinitives and dummy auxiliaries are stages in the typical acquisition of finiteness marking (van Kampen, 1997; Zuckerman, 2013). Constructions with a dummy

[1] The focus of the studies by Orgassa (2009) and Blom and colleagues (2013a) was on disentangling SLI and TD in bilingual populations. In bilinguals, verb inflection errors appeared to have much lower sensitivity and specificity. This indicates that verb inflection is a vulnerable domain in non-impaired second language learners as well.

auxiliary and an infinitive are regarded to be less complex options than constructions with an inflected lexical verb, both in linguistic theory and in terms of processing (de Jong, Blom, & Orgassa, 2013). Although not all studies report overuse of dummy auxiliaries in the production of children with SLI (Bastiaanse, Bol, van Mol, & Zuckerman, 2002), some studies report significantly more use of these constructions in children with SLI compared to younger TD children (de Jong, 1999; de Jong et al., 2013; Zwitserlood, van Weerdenburg, Verhoeven, & Wijnen, 2015b).

Yet another type of difficulty in the verbal domain is the relationship between auxiliaries and past participles reported in a study with children with SLI around 4 years of age (Wilsenach, 2006). In a sentence repetition task, children with SLI tended to omit auxiliaries, especially in more complex sentence structures or longer sentences. More noticeable were, however, errors in the past participle (omission of the prefix *ge*-). Problems in the verbal domain are thus not restricted to verb inflection errors, but verb inflection errors seem to be the most prominent.

The grammatical problems in Dutch children with SLI extend well beyond the domain of verb morphology, as the studies mentioned above (Bol & Kuiken, 1988; de Jong, 1999; Zwitserlood, 2014) already indicated. De Jong (1999) showed that the complexity of verb argument structures also seems to be an area in which differences between SLI and TD are found. Furthermore, Zwitserlood (2014) reported differences between SLI and TD in morphosyntactic rules that were not verb-related. Several grammatical aspects in the (pro)nominal domain have been shown to be particularly difficult to acquire. As has been reported in other languages, Dutch children also often omit articles (Wilsenach, 2006; van Ewijk & Avrutin, 2010) and struggle with personal and possessive pronouns (Bol & Kuiken, 1988; Zwitserlood, 2014). A number of recent studies have focussed on the acquisition of grammatical gender of nouns, as expressed in definite determiners and adjectives (Orgassa & Weerman, 2008; Orgassa, 2009; Weerman et al., 2011; Keij, Cornips, van Hout, Hulk, & van Emmerik, 2012). Children with SLI show high error rates in the assignment of gender to nouns and in the agreement between adjectives and nouns. The problems with gender agreement seem to be explained for a large part by erroneous gender assignment, as Weerman and colleagues (2011) argue. These problems do not seem to be solved before age 12 (Weerman et al., 2011; Keij et al., 2012). As well as verb inflection errors, errors in grammatical gender seem to be key in the morphosyntactic profile of Dutch children with SLI.

Although differences between SLI and TD are most prominent in the grammaticality of utterances, differences in syntactic complexity have also been reported (Bol & Kuiken, 1988; de Jong, 1999; Duinmeijer et al., 2012; Zwitserlood, 2014). In the cross-linguistic literature on SLI, problems in complex clauses like relative clauses or *wh*-questions are widely attested, but few studies have studied the acquisition of those structures in Dutch subjects with SLI. A number of studies do, however, report on reduced complexity of utterances in SLI or on the frequency of embedding (e.g., Bol & Kuiken, 1988; Duinmeijer et al., 2012). Whether differences appear does, however, seem to be dependent on the nature of the language production studied. Children with SLI were reported to use less embedded clauses in a retelling task in which the model story contained many subordinate clauses (Duinmeijer et al., 2012). When children generate a story themselves, differences in embedding seem to be absent (Duinmeijer et al, 2012; Parigger, 2012; Zwitserlood, 2014). Zwitserlood (2014) studied several measures of sentence complexity in a story generation task, and only found a difference between children with SLI and age-matched TD peers in the number of relative clauses used. The same author performed an intervention study targeting (subject) relative clauses and reported that problems in the comprehension and production of relatives seem to be present in children with SLI (Zwitserlood, Wijnen, van Weerdenburg, & Verhoeven, 2015a). However, the study did not include a control group, leaving the question whether Dutch children with SLI have problems in this domain unanswered. Furthermore, only subject relatives were targeted, while the literature on SLI reports predominant problems in object relatives.

From the Dutch literature on grammatical problems in SLI, it appears that verb inflection and gender marking are not only the most elaborately studied but also particularly vulnerable areas in Dutch SLI. Those two grammatical domains, subject-verb agreement and gender marking in determiners and adjectives, were therefore chosen as grammatical variables in this study. Relative clauses were added as a third variable, since this aspect has been reported to be vulnerable in many languages (e.g., Håkansson & Hansson, 2000; Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2004; Contemori & Garaffa, 2010) but little research exists on the Dutch relative clause.

As van Weerdenburg, Verhoeven and van Balkom (2006) indicated, the language profile of children with SLI might change over time. They examined a large group of children with SLI, comparing six-year-olds (N=147) with eight-year-olds (N=136) on a broad battery of (general) language tests and language-related cognitive skills. All the language measures they studied revealed significant differences from the norm in both age groups, but aspects like morphological skills (noun plurals and past tense) and recognizing syntactic patterns (picture recognition on the basis of a syntactic structure) appeared to have discriminative power only in the eight-year-olds. Over time, grammatical aspects may become increasingly or decreasingly important and different grammatical skills may be required at different developmental stages. In the following section, we will review the literature on persistent problems in SLI, with a special focus on the few studies discussing grammatical aspects.

2.2 Persistence of grammatical difficulties

Many studies on adolescents or adults with SLI report on the non-linguistic outcomes of the language disorder. While they are referring to the language impairment status of these children, they typically highlight persistent problems in the social, emotional, behavioural and psychiatric domain (e.g., Weiner, 1974; Baker & Cantwell, 1987; Beitchman et al., 1996a; Davison & Howlin, 1997; Beitchman et al., 2001; Clegg et al., 2005; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006; Conti-Ramsden & Durkin, 2008; Arkkila et al., 2009). Alternatively they focus on the influences of a language impairment on academic achievement and employment opportunities (e.g., de Ajuriaguerra et al., 1976; Hall & Tomblin, 1978; Aram et al., 1984; Stark et al., 1984; Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996b; Stothard et al., 1998; Rescorla, 2000; Snowling, Bishop, & Stothard, 2000; Conti-Ramsden, Knox, Botting, & Simkin, 2002; Nippold & Schwarz, 2002; Young et al., 2002; Law et al., 2009; Conti-Ramsden & Durkin, 2012; Conti-Ramsden, Durkin, & Walker, 2012; Conti-Ramsden, Mok, Pickles, Durkin & Conti-Ramsden, 2014). Far fewer studies discuss the linguistic outcomes of language disorders in adolescence or adulthood. Moreover, the majority of these few studies report only on measures of general language abilities, but do not indicate which specific grammatical problems remain (e.g., Stothard et al., 1998; Johnson et al., 1999; Conti-Ramsden et al. 2001a; Nippold & Schwartz, 2002; Clegg et al., 2005).

Only a handful of studies on adolescents with SLI have investigated the persistence of specific grammatical aspects characteristic of SLI in childhood. Those studies, which mainly describe English subjects with SLI, will be reviewed here. We will start with the studies on problems in the morphological domain, followed by the studies discussing problems in the syntactic domain.

English children with SLI are characterized by problems in past tense marking and third-person singular marking, as was noted in §1.1. Both aspects of verb morphology are persistently problematic. Norbury, Bishop and Briscoe (2001), for instance, showed that older children with SLI performed significantly worse than chronological age-matched peers and language-age matched peers on production tasks for these aspects. Past tense marking proved to be the most difficult aspect. This was confirmed by another study measuring the same two aspects in a large cohort of 11-year-old English subjects with (a history of) SLI (Conti-Ramsden et al., 2001a). The SLI groups were mainly characterized by a higher amount of omission errors: in the majority of items past tense markings were left out (Norbury et al., 2001; Marchman, Saccuman, & Wulfeck, 2004). According to Marchman and colleagues, who tested 27 children with SLI in a wide age range and compared scores of younger and older children, children with SLI make substantial progression in past tense marking performance over time, but differences between SLI and TD persist (Marchman et al., 2004). Differences mainly persist in regular tense formation, as was indicated in a study by van der Lely and Ullman (2001). They compared a group of adolescents with grammatical SLI (N=12, age 9;3-12;10) to different control groups matched on morphological abilities and vocabulary (age range 5;5-8;9). The SLI group produced fewer regular forms, did not show an advantage of regular over irregular past tense formation unlike the TD groups, and showed effects of frequency in both regular and irregular past tense marking (while this frequency effect was only found in irregulars in TD) (van der Lely & Ullman, 2001). In another study by Norbury and colleagues (2001) the difference in regular past tense marking between SLI and vocabulary-age matched peers disappeared when novel verbs were tested. Differences between SLI and TD thus mainly seem to persist in regular past tense marking of existing verbs (Norbury et al., 2001; van der Lely & Ullman, 2001).

Miller, Leonard and Finneran (2008) showed that, even at the age of 16, English adolescents with SLI (N=48) are less sensitive to the omission or substitution of grammatical morphemes. Although scoring relatively high on grammaticality judgement tasks (90%), adolescents with SLI scored significantly worse than age-matched controls. They were less sensitive to the omission of tense morphemes (*-ed* and third person singular present *–s*), commission errors of the same tense morphemes (‘*it is impolite to stared at people at the mall’)*, and omission of non-tense morphemes (*-ing* and the possessive *–s*). Omission of *-ed* was the most difficult to perceive. The authors indicate that the adolescents with SLI do not seem to have a deficit in their grammatical representations because they have a high correct judgement rate on tasks tapping the perception of grammatical rules. However, they make some errors in their production and they still have problems perceiving specific grammatical violations. Similarly, Rice, Hoffman and Wexler (2009) found persistent differences between SLI and TD in the judgement of omissions of finiteness markers (*be/do*) in yes/no questions and *wh*-questions. They followed a sample of 20 children with SLI between 7;6 and 15;0 years of age and found differences with age-matched and language-matched peers at all 9 times of measurement. While controls reached ceiling levels from age 7 onwards, adolescents with SLI seemed to reach a lower asymptote. The authors indicate that judgements of finiteness marking may be a valuable clinical marker for both younger and older subjects with SLI.

English adolescents with SLI thus clearly have persistent problems with finite verb morphology. Spontaneous language data from Dutch older children with SLI (8-12 years, N=36) indicates that, in Dutch, grammatical errors mainly persist in subject-verb agreement. Persistent problems with past tense marking were expected on the basis of the English literature, but were not found in the spontaneous language samples of a group of adolescents with SLI (Hoek, 2010). Using experimental tasks for subject-verb agreement, Weerman and colleagues (2011) found that Dutch adolescents with SLI (N=9, age 12-13) performed better than the younger SLI group and obtained high accuracy rates. However their performance was not yet 100% (around 95%) while TD children reached ceiling around 6 years of age. The error rate appeared to be higher in main clauses, which in Dutch involve movement of the verb to second position of the sentence. Similar effects of syntactic context were reported by Blom and colleagues (2013a). Subject-verb agreement thus seems to remain a vulnerable area for SLI into adolescence in Dutch, but performance seems to be dependent on the linguistic context.

As was discussed in the previous section, morphological problems in Dutch children with SLI extend beyond the domain of verb morphology. Problems are attested in the assignment of grammatical gender to nouns and in gender agreement between adjectives and nouns (Orgassa, 2009). Keij (2009) tested knowledge of grammatical gender in articles in Dutch and compared younger (N=11, 6-10 years) and older age groups with SLI (N=9, 8-12 years) to controls. Both the younger and the older group of children with SLI had significant problems in gender assignment in comparison to their TD peers. Older children were better at assigning gender to nouns in a knowledge task (grammatical judgement), but they did not perform any better than the younger children on a production task. This observation was confirmed in a study by Weerman and colleagues (2011), who tested article assignment with a production task and obtained similar results. Adolescents with SLI (N=9, aged 12;3-13;3) were not significantly better in assigning correct articles to nouns than younger children with SLI (N=25, aged 6-8). In adjectival inflection (inflection according to the gender of the noun) adolescents with SLI seemed to have acquired the rules for inflection of the adjective, but because they persistently failed to assign the correct gender to nouns, the adjective was not inflected correctly. Gender assignment to nouns thus seems to be an area in which problems persist in SLI, although performance is dependent on the task used.

A number of studies have looked at the persistence of problems with complex syntactic structures, such as relative clauses or *wh*-questions. Problems with complex structures can express themselves in two ways. Children with SLI use them less than TD children and they make more errors in these structures. The studies reported here indicate that both types of problem are to be found in adolescents with SLI.

Marinellie (2004), for instance, analysed spontaneous language samples of older English children with SLI (N=15, average age 10;8) and age-matched TD peers on several complex constructions. Most of the structures used by the older TD children were also used by the older children with SLI, although to a lesser extent. Similar findings were reported by Hesketh (2006), who examined relative clauses in an elicitation task and a narrative task in a somewhat younger group of English children with SLI (N=66, aged 6-11). It was hard to elicit full relative clauses from the group of children with SLI (the authors reported that the construction was avoided) and in the narrative data they used more reduced relatives instead of full relative clauses (*'the man living upstairs goes to work'*).

This finding of less frequent use of subordination in SLI was recently confirmed for French (Tuller et al., 2012). Tuller and colleagues investigated the complexity of the spontaneous speech of French adolescents with SLI (N=18, aged 11-16) in comparison to different age groups of TD children/adolescents (6 years, 8 years and 11 years of age). The results are interpreted as indicating that French adolescents with SLI tend to avoid complexity in their spontaneous speech by using coordination/juxtaposition (rather than subordination), by omitting the complementizer and via direct speech. They used even fewer complex clauses than their 6-year-old language matched peers (matched on MLU in words).

In two small studies of Dutch, adolescents with SLI were reported as using complex sentence structures as frequently as TD peers. Studying narrative data, Burger and Rijpma (1998) found no difference between adolescents with SLI (N=10, average age 14;3) and their age-matched peers on MLU and syntactic complexity. The groups did, however, differ on the number of grammatical errors. A later study by van Groningen (2010) examined the correlation between the complexity and correctness of sentences in three age groups of Dutch children with SLI (8-9, 9-10, 10+, total N=24). She did find some support for the idea that complex syntactic structures are still difficult for older children with SLI. The mean number of errors per utterance correlated significantly with the complexity of the sentence (the more complex the sentence, the more errors were made). On the basis of the number of errors subordinate clauses appeared to be the most problematic structures. Dutch adolescents with SLI thus seem to use complex structures to the same extent as their TD peers in spontaneous speech, but find them more difficult since they make more errors. The contrasting findings between languages may be explained by methodological or cross-linguistic differences between studies. Perhaps, some contexts or methodologies elicit more complex constructions than others.

In sum, the review of the few studies that discuss grammatical problems in adolescents with SLI indicates that there are persistent problems in grammatical aspects. In the syntactic domain, persistent problems were found in complex syntactic structures, as indicated by fewer occurrences of complex structures and higher error rates within them. In the morphological domain, problems with finiteness marking of verbs have been shown to be persistent into adolescence and, in Dutch, problems with gender assignment are still attested in older children with SLI. However, performance in adolescents with SLI seems to be influenced by the linguistic context or type of task. In the next section, we will discuss which linguistic context factors were manipulated in our tasks to further examine these performance effects.

2.3 Linguistic context factors

As has been discussed in Chapter 1, variability in performance is often reported in SLI, not only between subjects but also within subjects. In this study, we will test a theoretical account of SLI that assumes that children with SLI have a processing problem that influences the acquisition of grammatical knowledge, but at the same time impedes the implementation of grammatical knowledge once it has been acquired (Bishop, 1994, see §1.2 for a more elaborate description). The hypothesis predicts that "errors will occur when the speech production system is stressed by the need to produce output that makes heavy demands on its processing capacity" (Bishop, 1994, page 528). In other words, variability in performance is predicted to be dependent on the processing costs of the message or the context in which the message is conveyed. In this section, we will describe how the concept 'complexity' is defined and operationalized here.

Why are some linguistic aspects or forms harder to process than others in comprehension, or preferred over other forms in production? In the processing literature, this is often linked to differences in processing cost between linguistic contexts. Linguistic phenomena that require more cognitive resources are assumed to have a higher complexity (or vice versa, phenomena with a higher level of complexity are assumed to require more cognitive resources). In the past decades, a range of different factors have been identified to affect linguistic processing, as reflected in accuracy rates or reaction times. Effects are either ascribed to differences in the amount of memory that is required or to differences in probability of the structures studied (Jaeger & Tily, 2011).[2]

Frequency is an example of a factor affecting linguistic processing due to input probability differences, although it is certainly not the only factor that constitutes probability. Effects of frequency have been shown in the processing of phonological, morphological and syntactic information, and in reading, spelling, listening and speech production tasks (Ellis, 2002). Elements with a lower frequency are generally processed more slowly and lead to higher error rates.

Proximity between elements in a dependency relation is another example of a factor for which influence has been attested within several linguistic phenomena. Agreement between determiner and noun is for instance reported to be cognitively less demanding if the elements are adjacent (Sagarra & Herschensohn, 2013) and the distance between verbs and its arguments has been found to affect sentence processing (Vasishth & Lewis, 2006). Whether the proximity of dependent elements must be calculated in terms of the number of words in between them, or in terms of the number of syntactic nodes involved (Hawkins, 1994) is a matter of debate (as is discussed by Jaeger & Tily, 2011), but effects of dependency length are commonly ascribed to differences in the amount of memory that is required.[3]

The grammatical variables tested in this study differ in nature and therefore required different operationalisations of the concept complexity. For each grammatical variable, a small number of linguistic factors were chosen that were assumed to affect processing on the basis of previous studies.

The assignment of grammatical gender has previously been found to be affected by the frequency of the noun and the distance between the determiner and the noun (Sabourin, Stowe, & de Haan, 2006; Keating, 2009). In more complex contexts – i.e. nouns with a low frequency and non-adjacent nouns and determiners – agreement accuracy between the gender of the noun and the form of the definite determiner appears to decrease. Noun frequency and dependency length were therefore used in the determiner assignment tasks to examine their effect on performance in SLI compared to TD. Similarly, agreement between the adjective

---

[2] Recently, a number of authors have tried to quantify linguistic complexity by applying information theoretical formulas to linguistic phenomena (del Prado Martín, Kostić, & Baayen, 2004; van Ewijk & Avrutin, 2010).

[3] An interesting line of research attempted to match psycholinguistic measures of complexity with typological patterns. Because speakers tend to minimize the use of cognitive resources, linguistic elements that reduce complexity are hypothesized to be distributed more widely across languages than elements that increase complexity (see Jaeger & Tily, 2011 and Hawkins, 1994, for a more detailed discussion of this topic).

and the noun was tested with nouns differing in frequency. However, the length of dependency relations could not be manipulated. This factor does not play a role in adjectival inflection - since adjectives are always adjacent to nouns. Instead, the number of adjectives that had to be inflected was varied between contexts (either 1 or 2). This should potentially influence processing costs (visually, because the items contained more pictures to distinguish, or conceptually, because children had to distinguish items on two scales (size and colour) instead of one (size or colour)).

Our second grammatical variable, subject-verb agreement, has been found to be affected by several linguistic factors. Frequency of the verb forms has, for instance, been found to influence processing accuracy of subject-verb agreement (Barker & Nicol, 2000), as has proximity between argument and verb (Vasishth & Lewis, 2006). Performance in SLI also seems to be dependent on whether the verbs are novel or existing (Weerman et al., 2011). Furthermore, previous studies on Dutch SLI revealed that children with SLI made more subject-verb agreement errors in main clauses than in embedded clauses, if the head of the sentence was already provided (Weerman et al., 2011; Blom et al., 2013a; Blom et al., 2014). In this study, we implemented slightly different factors. Based on previous findings on the effects of the phonological complexity of the verb stems on verb inflections (Marshall & van der Lely, 2007; Song et al., 2009; Blom et al., 2014), the phonological properties of verb stems in our production task were systematically varied. Inflection had to be added to verb stems that differed in sonority (plosive, fricative, liquidate) and in the number of consonants in the verb stem coda (1 or 2). Furthermore the syntactic context in which verbs were tested was varied: children had to produce inflections either in main clause contexts (where the head of the clause was already provided) or in subordinate clauses (where the main clause was provided, but the subordinate clause had to be constructed). For subject-verb agreement, we thus varied two phonological factors (sonority and coda complexity) and one syntactic factor.

For the third grammatical variable, relative clauses, effects of the role of the relative marker (or relativized head) in the subordinate clause on accuracy rates in comprehension and production have been widely attested in different populations, including SLI. In most languages, relatives appear to be easier to comprehend or produce when the relative marker takes a subject role than when it takes an object role. These differences are often ascribed to differences in processing costs, due to differences in the distance between the filler and the gap (Hawkins, 1999). As Hawkins (1999) describes, processing a relative clause requires identification of the gap (which is not an easy process because it is an empty element) and keeping the filler in working memory and simultaneously processing the other material that is encountered (see Chapter 7 for a more elaborate discussion of the characteristics of relative clauses). Especially in individuals with SLI, differences between subject relatives and object relatives prove to be persistently large (e.g., Friedmann & Novogrodsky, 2004; 2007; Jensen de Lopéz, Sundahl Olsen, & Chondrogianni, 2014). The type of relative clause (subject or object) was therefore varied in this study to see whether differences were larger in SLI than in TD and to test whether differences were related to processing abilities.

The comprehension of object relatives has furthermore been shown to be influenced by the characteristics of the arguments. Factors like animacy, number differences between subject and object, or the accessibility of the subject have been shown to influence the interpretation of object relatives (Mak, Vonk, & Schriefers, 2002; 2006; Kidd, Brandt, Lieven, & Tomasello, 2007; Adani, van der Lely, Forgiarini, & Guasti, 2010; Adani, Forgiarini, Guasti, & van der Lely, 2014). Object relatives with different characteristics were therefore compared. The characteristics of the object relatives were varied in the perception tasks instead of the production tasks for several reasons. One reason was that Dutch object relatives were hard to elicit. Furthermore, the previously found effects of context on object relatives that were described above were found in perception, while little evidence of these factors is provided in production (but see Mak et al., 2002 for a corpus study on the production of object relatives). A final reason was that relatives are acquired rather late in typical development. Because the literature often indicates that attainment of grammatical variables manifests itself earlier in comprehension or judgement tasks than in production tasks, the context effects on object relative accuracy would plausibly be easier to demonstrate in judgement/comprehension than in production.

Because some studies on SLI have reported differences between outcomes on comprehension and production tasks, for instance in the persistence of problems - as discussed in the previous sections (Miller et al., 2008; Keij, 2009), all three grammatical aspects were examined with multiple tasks (comprehension/judgement and production).[4] As well as a comparison of different contexts in which the grammatical variables were tested, we also looked at potential differences between perception and production.

[4] Gregg, referring to a model of second language acquisition, noted that ‘knowledge must stretch far beyond what a subject can produce, because it for instance must constrain what the subject does not produce’ (Gregg, 1990).

2.4 Hypotheses

The previous discussion of the relevant background literature and of the concept complexity and its operationalization in different context factors makes it now possible to specify hypotheses in relation to the research questions formulated in Chapter 1. Below, each of the research questions is linked to specific expectations of the outcomes.

1. Do the grammatical differences between SLI and TD peers that are characteristic in childhood persist into adolescence?

On the basis of previous studies on grammatical abilities in adolescents with SLI, grammatical difficulties are hypothesized to be persistent. Differences between SLI and TD in grammatical gender and subject-verb agreement have been attested in Dutch children with SLI. Grammatical gender has even been shown to be persistently difficult (Weerman et al., 2011). For relative clauses, no studies on Dutch children with SLI have previously been carried out but we expect this aspect to be (persistently) difficult in SLI based on the international literature and Dutch studies indicating difficulties in complex syntax (e.g., Duinmeijer et al., 2012; Zwitserlood, 2014). In statistical terms, we expect an effect of Group (SLI or TD) in the outcomes on the three grammatical variables but we do not expect an interaction between Group and Age (young or old). If an interaction between Group and Age exists, post-hoc analyses will be calculated to reveal whether this means the difference between SLI and TD disappeared in adolescence, or whether children with SLI showed significant improvement over time while still falling significantly short in comparison with TD peers (the latter outcome would support our hypothesis).

2. Is there variability in the grammatical performance of children and adolescents with SLI, and can this variability be explained by factors related to linguistic context?

As was described in Chapter 1 and earlier in this chapter, several studies have indicated variability in performance in SLI dependent on the context or task in which grammatical aspects are tested. In the literature on older subjects with SLI, some studies also note this variability in performance. The theoretical claim we aim to test here predicts patterns in variability, dependent on the amount of processing load involved. Some linguistic contexts are hypothesized to involve higher levels of processing load than others, due to increasing complexity of the information. In the previous sections, it was indicated how the concept 'information complexity' is operationalized. We also described the linguistic factors used in our tasks to test whether variability in performance in SLI is indeed explained by such linguistic factors that require a higher amount of processing load. The hypothesis is that these factors will affect performance in all subjects, but more so in subjects with SLI. In statistical terms, we expect to find differences between outcomes in different contexts, and we expect these differences to interact with Group (SLI or TD). Differences are expected to be larger in the SLI group than in the TD group, as will be tested with post-hoc analyses.

3. Is any variability in linguistic performance related to (impaired) processing abilities?

In order to answer the last question of this study, an assumption has first to be tested, namely that processing capacities are limited in SLI. As was discussed in §1.2, there is a growing body of evidence that processing abilities in SLI are different from TD peers – although not all aspects of information processing are problematic to the same extent and not all individuals with SLI seem to be limited in broader information processing abilities. The few studies discussing information processing in older children with SLI seem to indicate that processing limitations are persistent. We thus hypothesize to find differences in information processing between SLI and TD in the younger and the older groups. Only those aspects for which a persistent significant difference between SLI and TD is found will be taken into account for answering the third research question. The hypothesis is that there are correlations between variability in performance and processing abilities. Poorer processing abilities are expected to be associated with higher amounts of variability in performance (a larger impeding influence of linguistic context). Statistically, we thus expect to find a negative correlation between variability level and processing ability.

.

3 Methodology

In the first section of this chapter, the criteria for recruitment and selection of the subjects with SLI and the TD groups are described (§3.1). Subsequently, an overview of the group results on the variables general language ability, non-verbal IQ and socio-economic status (SES) are presented (§3.2). These variables were used as a check for typical development in the TD groups (general language ability) and as covariant variables in the analyses (non-verbal IQ). Furthermore, some background information on the medical history, achievement of language milestones and language problems in the family of subjects is provided in order to get a clearer picture of the characteristics of the groups (§3.3). The materials used for the investigation of the processing abilities and the grammatical abilities of the subjects are described in §3.4. Finally, the procedures of administration, transcription, scoring, coding and statistical analyses are presented in §3.5.

3.1 Recruitment and selection of subjects

To examine development in SLI, a cross-sectional design was chosen with two age groups: a younger group of children with SLI between 6 and 10 years of age and an older group of adolescents with SLI between 12 and 16 years of age (SLI-Y and SLI-O). Two groups of typically developing children (TD-Y) and adolescents (TD-O) were then matched to these children on the basis of age and gender.[1]

The younger SLI group was recruited in primary special education schools for children with a diagnosis of a language disorder, based on strict inclusion criteria that closely resemble the international criteria: a child had to score at least 1.5 SD below the mean on at least 2 out of 4 language domains (speech, auditory processing, grammatical development, lexical-semantic development) or 2 SD below the mean on the total score of a language test (Resing, Evers, Koomen, Pameijer, & Bleichrodt, 2005). Their performance IQ also had to be within the

[1] Data of the younger groups were collected in collaboration with Jeannette Schaeffer.

normal range (85-115), and they should have received at least 6 months of treatment from a speech and language therapist without satisfactory improvement.

As mentioned in §1.1, after primary school some of the children with a diagnosis of a language disorder remain in special education while others move into regular schools, where they still receive some support from specialists in special education. In order to find an older group of subjects with SLI that would be similar to the younger group in terms of severity of the language problems, the older group with SLI was recruited in both types of educational contexts (special education and regular schools with support). The older children had a diagnosis of a language disorder on the basis of the same inclusion criteria as the younger group. In the Dutch educational system, diagnoses of language impairments are valid for four years, so every child had received this diagnosis at some point in the previous four years.

For both SLI groups, additional exclusion criteria were applied to select those children for whom the language impairments seemed ‘specific’ to language or (as Bishop, 2014, puts it) whose language problems are ‘unexplained’. In order to be included, subjects with SLI had to have normal hearing and they could not have a diagnosis of Autism Spectrum Disorders (ASD) or Attention Deficit (Hyperactivity) Disorder (ADHD) so that language problems plausibly caused by another disorder were excluded. Children had to be monolingual speakers of Dutch in order to exclude effects of multilingualism. Furthermore, children with a profile of major speech problems were excluded. The files of the subjects with SLI were screened for these criteria. Only the parents of the children who met our criteria received a letter to ask for informed consent.[2]

In the younger SLI group, 38 informed consents were received, but six children were excluded after testing for several reasons. Three children appeared to have a comorbid disorder during the testing phase (ASD). In one child the communication problems seemed to be restricted to speech problems (cleft palate/Developmental Coordination Disorder). One child appeared to be growing up in a multilingual situation and one child was not motivated enough to participate. In the older SLI group, 35 adolescents were recruited. Four were excluded after testing

---

[2] This letter contained information on the goal of the project and the content and duration of the testing sessions. Parents were asked to sign and return the consent form if they agreed on participation of their child. They also were asked to answer the question whether any other languages than Dutch were spoken at home. The letter provided information about who to contact in case of questions or complaints. The design of this study and the correspondence to parents was officially approved by the Ethical Committee of the University of Amsterdam (reference number 1012-8).

on the basis of their multilingualism (1), comorbid disorders (1 ASD, 1 ADHD) or because they were not motivated to finish the tasks (1). This resulted in a younger group of 32 children with SLI and an older group of 31 adolescents with SLI (see Table 3.1).

The younger and older TD children were recruited via mainstream schools. Schools were chosen with similarly low ratios of multilingualism to the schools attended by the children with SLI, since a high level of multilingualism in a school can affect the variety of Dutch spoken by the monolingual children (Cornips, 2008). The older TD group was recruited from classes at educational levels similar to the levels of the older subjects with SLI. For the younger groups, matching for educational level was not possible since determination of the educational level of a child is only done just before the child moves on to secondary education at age 11/12.

Table 3.1. Demographic characteristics of the selected groups

Group	N	Mean age (SD)	Age Range	Gender
SLI-Y	32	8;4 (1;5)	6;4-11;0	25 M, 7 F
TD-Y	32	8;4 (1;5)	6;4-10;11	25 M, 7 F
SLI-O	31	14;5 (0;11)	12;0-15;10	20 M, 11 F
TD-O	30	14;5 (0;11)	12;0-15;9	19 M, 11 F
Adults	22	31;3 (13;1)	18;0-53;0	14 F, 8 M

Within the TD groups, the same exclusion criteria were applied, with the additional criterion that children should not have (a history of) language problems or impairment. Screening for exclusion criteria took place after receiving the informed consent forms from the parents and was based on the reports of teachers and parents. Only those children without any reported language problems were selected for testing. After the administration of the tests, the ‘typical language development status’ of the TD subjects was checked with a test for general language abilities (CELF-4-NL, Kort, Schittekatte, & Compaan, 2008, see §3.2 for a more elaborate description of this measure).

Six subjects in the older TD group were excluded on the basis of low general language scores (-1.5 SD or below). Furthermore, one younger TD subject and one older TD subject had to be excluded on the basis of color blindness (since adjectives for color were crucial in the adjectival inflection task) and ADHD, respectively.

From a group of 112 TD subjects that were recruited, tested and selected as proper control subjects, two groups were selected to match the SLI groups in terms of age and gender (N = 32 and N= 30, respectively, Table 3.1). Apart from the four groups of children and adolescents, a group of adults was tested on several tasks to check the validity of the experimental tasks and to determine levels of ultimate attainment.

Table 3.1 shows the mean age, age range and gender distribution in the different groups. The younger groups were approximately 5 years younger than the older groups. As expected in a sample of subjects with SLI, the gender distribution was not equal, so more boys than girls were selected in all groups.

3.2 Language ability, non-verbal IQ and SES of subjects

Within every group, background information on language abilities, non-verbal IQ and socio-economic status was gathered in order to control for these variables in group comparisons. As described in the previous section, general language ability was measured to be able to exclude TD children with scores below the average from our control group. This measure also gave an indication of the difference between the groups in terms of language abilities. Since non-verbal IQ and socio-economic status might be factors influencing language abilities, these variables were measured to see whether our groups were comparable in these areas. Any difference in terms of the cognitive abilities or social background between groups needs to be corrected for in the statistical analyses.

General language ability was measured using the Dutch version of the Clinical Evaluation of Language Fundamentals *(*CELF-4-NL, Kort et al., 2008). The CELF (Semel, Wiig, & Secord, 2006) is a widely used general measure of language ability and had been standardized for Dutch. The test consists of a considerable number of subtests but allows a quick evaluation of general language ability using only 4 subtests to obtain a Core Language Score (administration of these subtests takes around 30 minutes). The Core Language Score has been shown to distinguish between good and poor language abilities. Dependent on age, different subtests are used (see Table 3.2. for the different core tests per age group and a short description of each task).

Table 3.2. Description of the subtests of the CELF-4-NL and the ages at which they were used (Semel et al., 2006)

CELF Subtest	Description	5-8	9-12	13+
Word structure	*Complete sentences by using a targeted structure(s).* This task tests the ability to implement grammatical rules	Core		
Concepts and following directions	*Point to pictured objects in response to oral directions.* The task tests the ability to (a) understand instructions that increase in length and complexity and require logical thinking, (b) remember the characteristics and order of the objects that have been named, and (c) choose from a number of options those objects that have been named.	Core	Core	
Recalling sentences	*Imitate sentences presented by the experiments.* The task tests the ability to (a) listen to spoken sentences of increasing length and complexity and (b) repeat the sentences without changing the semantic, syntactic or morphological structure.	Core	Core	Core
Formulated sentences	*Formulate a sentence about a picture using a given word.* This task tests the ability formulate complete sentences that are semantically and grammatically correct and increase in length and complexity.	Core	Core	Core
Word classes 2 total score	*Choose which 2 words belong together (from 4 options) and explain why*. The composite score of this test shows how well a child can understand semantic links between words and how well the subject is able to describe the relationship.		Core	Core
Word definitions	*Define the meaning of a word*. This task tests the ability to (a) discover the most important meaning of words, (b) define words by referring to connections between word classes, and (c) describe the semantic links between words.			Core

Non-verbal IQ was tested by means of the Raven's Progressive Matrices (Raven, Raven, & Court, 2003), a standardized non-verbal IQ test for children and adults of any age. Participants are presented with black and white pictures showing a geometrical pattern with a missing piece. Underneath this picture are six or eight pieces in the shape and size of the missing piece but with different geometrical patterns. The participant needs to choose the matching missing piece (see Figure 3.1 for examples of the items). Instructions are given both verbally and non-verbally, that is by pointing. The test is standardized, has often been used with children, and is internationally accepted as an index of non-verbal cognitive abilities. The manual gives quartile scores (e.g., 25-50%), but percentile scores were computed by converting the quartile scores.[3]

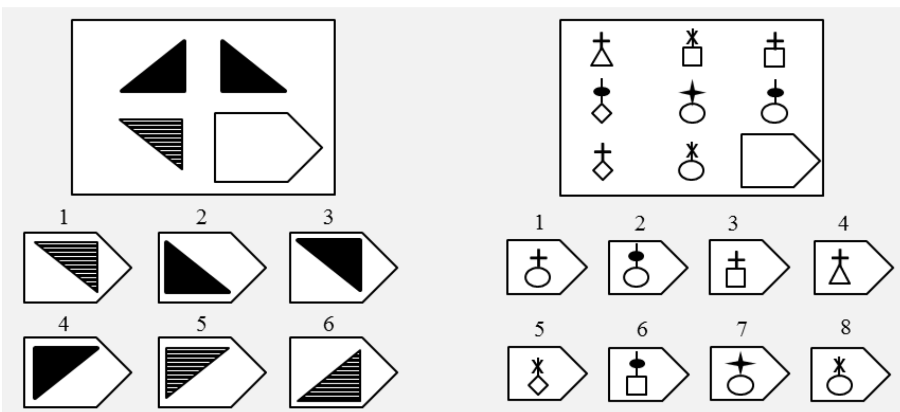


Figure 3.1. Examples of an easy item (left) and a difficult item (right) of Raven's Progressive Matrices

As a final background measure, socio-economic status (SES) was determined with the help of a parental questionnaire. The educational level of the mother (low, middle or high) that was reported in the questionnaire formed the basis for classification.[4]

---

[3] Percentile scores were computed on the basis of the norm scores of children and adolescents in the USA (Raven et al., 2003, page 92).

[4] *High* was used for university or college level (WO/HBO), *middle* was used for the higher vocational levels (MBO) and *low* was used for the lower vocational levels or for no education after high school (LBO/none).

In Table 3.3, the scores for general language ability, non-verbal IQ and SES of the different groups are shown. The background measures reveal important characteristics and differences between the groups.

Table 3.3. Language ability (CELF), non-verbal IQ (RAVEN) and SES of the participants

		SLI-Y N=32	TD-Y N=32	SLI-O N=31	TD-O N=30
Language ability (SS)	Mean	72.89	111.38	73.68	96.63
	(SD)	(8.34)	(9.30)	(11.45)	(12.95)
Non-verbal IQ (percentile[1])	Mean	49.34	69.81	25.45	38.77
	(SD)	(28.59)	(22.13)	(25.00)	(24.21)
Socio-economic status (SES)	Low	11	1	11	7
(educational level mother[3])	Middle	12	1	12	10
	High	2	26	5	10
	Missing	7	4	3	3

A first important observation is the fact that there was – as expected – a significant difference in general language ability between the groups with and without SLI. ANOVAs with the CELF Core Language Score as a dependent variable revealed that this difference was highly significant in the younger groups ($F(1,63) = 303.66$, $p = .000$, $\eta_p{}^2 = .830$), and smaller but still highly significant in the older groups ($F(1,61)= 53.90$, $p = .000$, $\eta_p{}^2 = .477$). Both SLI groups scored, as expected, around -2 SD below the mean on average (standard scores around 70). Compared to the norm, the younger TD group had a high average language score (+ .76 SD) and the older TD group had a low average language score (- .22 SD).

A more detailed view of the general language scores is provided by a scatterplot, where every dot represents the score of an individual (Figure 3.2). From this figure, it becomes apparent that there was no overlap in general language ability scores in the younger groups, but the older groups cannot be separated cleanly. Apparently, a number of adolescents with an SLI diagnosis performed within the normal range on this measure of language ability, and at the same time some TD adolescents had scores at the lower end (although TD subjects with scores lower than -1.5 SD were excluded). This latter result may be related to our recruitment process, since we chose to select TD adolescents at the same educational level and the same educational environment as the adolescents with SLI (as remarked above, this was not possible in primary school children). Although we do not claim to know

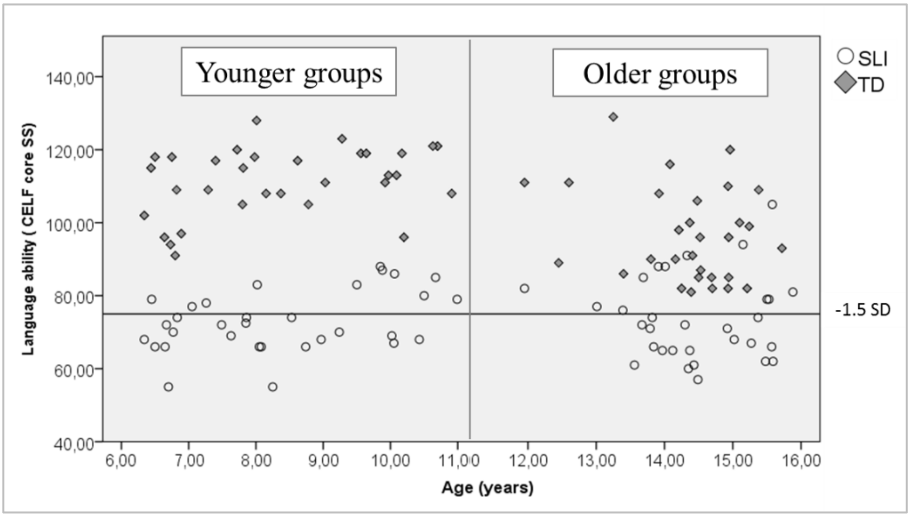


Figure 3.2. Scatter plot of general language ability (CELF core SS) in the different groups

the direction of causality, language abilities and educational level can be assumed to be related to some extent.

The fact that some individuals with SLI scored within the normal range on language ability might raise the question as to whether a CELF Core Score higher than -1.5 SD should not disqualify the children as being SLI for this study. As described above, all subjects with SLI had an official diagnosis on the basis of a much more elaborate test battery than the 4 subtasks that constituted the general language ability score we reported. It was therefore decided not to question their diagnosis. It is possible that their diagnosis was based on two scores of -1.5 SD below the mean in two language domains instead of a general language score of -2 SD below the mean (see the criteria for SLI in §3.1). The individual results on the different subtests of the CELF Core Language Score, show that almost all subjects with SLI obtained scores below -1.5 on at least two out of the four subtests (N=27 in the younger group and N=26 in the older group). Furthermore, the diagnosis might have been based on language domains that were not included in the tests constituting the Core Language Score of the CELF (the tests were chosen on the basis of their discriminatory power, not on whether they equally covered all language domains).

An average performance on the Core Language Score therefore does not seem to invalidate the language impairment status.[5]

A beneficial consequence of the recruitment of TD adolescents at similar types of educational level is the fact that in the older groups, non-verbal IQ-scores were more equal between the SLI group and the TD group. ANOVAs with IQ percentile scores as a dependent variable and Group as a between subjects factor show that there was a significant difference between SLI and TD in non-verbal IQ ($F(1,63) = 10.84$, $p = .001$, $\eta_p^2 = .081$). Post-hoc analyses reveal that the significant difference was strong in the younger groups ($F(1,63) = 10.25$, $p = .002$, $\eta_p^2 = .142$) and less strong in the older groups ($F(1,61) = 4.46$, $p = .039$, $\eta_p^2 = .070$). The older groups were thus better matched on non-verbal IQ but differences still existed. To account for the differences in IQ between SLI and TD, non-verbal IQ will therefore be entered as a covariate in the statistical analyses in the following chapters. The SLI and TD groups also differed considerably in terms of SES, as measured by the educational level of the mother. The difference is largest in the younger groups and quite small in the older groups (see Table 3.3). SES and non-verbal IQ are, however, strongly correlated ($r = .475$, $p < .001$). Because IQ will be corrected for in the statistical analyses, SES will not be taken into account in the further analyses of this thesis.

Ideally, the SLI and TD groups would be well-matched on non-verbal IQ and SES, and only differ in their language abilities. These differences in language abilities should also be evident in every possible language test. An important question is, however, whether cognitive abilities and language are fully separable. Many studies report a different in performance IQ between children with SLI and TD children (e.g., Krassowski & Plante, 1997; Gallinat & Spaulding, 2014). Furthermore, the relationship between IQ and language might change over time, as has been found in follow-up or longitudinal research on SLI (Botting, 2005; Clegg et al., 2005).

If we look at the correlation between the non-verbal IQ-scores and the language scores in the different age groups, we find weak correlations (Spearman's rho) in the younger groups ($r = -.079$ in the SLI-Y group and $r = .032$ in the TD-Y group, both not significant), but strong correlations in the older groups ($r = .603$ in the SLI-O group and $r = .496$ in the TD-O group). It seems as if the correlation between language abilities and non-verbal IQ becomes stronger over time, which

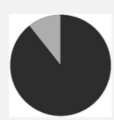
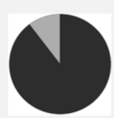
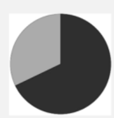
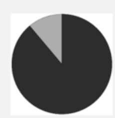
[5] The adolescent with SLI that scored above average on our measure of general language ability was participating in a higher educational level and had a fairly high non-verbal IQ (percentile score 95). His diagnosis was based on problems in spontaneous speech rather than on language tests.

might be regarded as evidence that language and cognitive development are intertwined.

3.3 Background information on subject groups

To create a better picture of the subject groups and the differences between them, all parents of the participants were asked to fill in a questionnaire covering the birth of the child, the medical history, milestones in the language development of their child, and the occurrence of (language) disorders in the family. This information was not used for selection or exclusion, but gave additional insight into group characteristics. Not all parents returned the questionnaires but the percentage of received questionnaires was more or less equal across different groups (SLI-Y 81%, TD-Y 91%, SLI-O 90%, TD-O 90%). Table 3.4 shows the general results.

Table 3.4. Background information on birth and health and hearing

	SLI-Y	TD-Y	SLI-O	TD-O
Delivery complications? ■ NO □ YES				
Health problems? ■ NO □ YES				
Frequent colds? ■ NO □ YES				
Frequent otitis media? ■ NO □ YES				

The table shows a higher reported rate of childbirth complications in the younger SLI group, but the problems that are reported vary in severity (from slow and long deliveries, to oxygen deprivation and urgent C-sections). Health problems in development are also more often reported in the younger SLI group (e.g., epilepsy, frequent vomiting because of relatively large tonsils, asthma and kidney infection) and SLI subjects seem to have a history of colds more often than reported in the other groups (some children had a cold many times during the last year). In both SLI groups, the number of children with past occurrences of recurrent otitis media is higher compared to the TD group. There are also some children with a history of ear infections in the TD groups but the differences are not that large.

The questions about the milestones in children's language development, and the occurrence of language problems in their families reveal – as expected – more differences between the groups (see Table 3.5). In the TD groups, no (or very few) language problems are reported in the family, while the reported occurrence of language problems in the families of children and adolescents with SLI is much higher.

Table 3.5. Background information on the occurrence of language problems in the family and language milestones in the child (varied babbling, first words and sentences of the subjects)

	SLI-Y	TD-Y	SLI-O	TD-O
Language problems in family? NO YES				
Dyslexia in family? NO YES				
Variegated babbling? NO YES				
Age of first words (mean, (SD))	20 months (0;10)	13 months (0;5)	21 months (0;8)	13 months (0;4)
Age of first 2-word sentences (mean, (SD))	30 months (0;13)	19 months (0;10)	32 months (0;9)	22 months (0;5)

Strikingly, the rate of reported dyslexia in the family is more or less equal across the SLI and the TD groups, but the reported rates are higher in the older populations. This might be explained by the fact that adolescents might have siblings or cousins that have recently been diagnosed with dyslexia (vice versa, the child groups might have siblings or cousins that will be diagnosed in the future).

Clear differences can be noted in the parental reports with respect to the early milestones in language development (variegated babbling, first words and first sentences). In the two TD groups, almost all children are reported to have gone through a stage of variegated babbling, while in the SLI group only about half of the children is reported to have shown this developmental stage.[6] If we look at the age at which subjects were reported to speak their first words and sentences, both SLI groups are commonly delayed in their onset of language production. The difference in onset of word production is significant between SLI and TD in the younger groups ($F(1,35) = 5.46$, $p = .025$, $\eta_p^2 = .135$) and the older groups ($F(1,32) = 9{,}30$, $p = .005$, $\eta_p^2 = .225$). The difference in the onset of two-word sentences is not significantly different between SLI and TD in the younger groups ($F(1,26) = 3.59$, $p = .069$, $\eta_p^2 = .121$), but it is in the older groups ($F(1,26) = 10.32$, $p = .003$, $\eta_p^2 = .284$). It should be mentioned though that not every parent answered these questions (often because they did not remember) and reports were mainly missing in the TD groups (on the question pertaining to the age of first words, 92% and 75% of the parents answered in the younger and the older SLI group respectively, and around 45% in the two TD groups. On the question about the age of first two-word sentences, the percentage of respondents who answered the question was 80% and 60% in the SLI groups, and 25% and 40% in the TD groups). If language development was delayed, parents were more likely to remember the age at which milestones were attained than when language development was normal (perhaps also because parents of language impaired children had been asked questions about the milestones several times in clinical settings while parents of TD children had not had that experience).

Although group comparisons revealed significant differences between groups in terms of their age of onset of first words and sentences, box-plots reveal that there is large variability in the SLI groups in terms of the onset of language production, and the range of onset age overlaps with the range in the TD groups (see Figure 3.3). This is consistent with the literature on language impairments. Late onset of language production is not always an indicator of language impairment at an older age, and an early onset does not always imply normal language development (Leonard, 2014).

[6] Parents were simply asked whether their child went through a phase of variegated babbling, without further explanation.

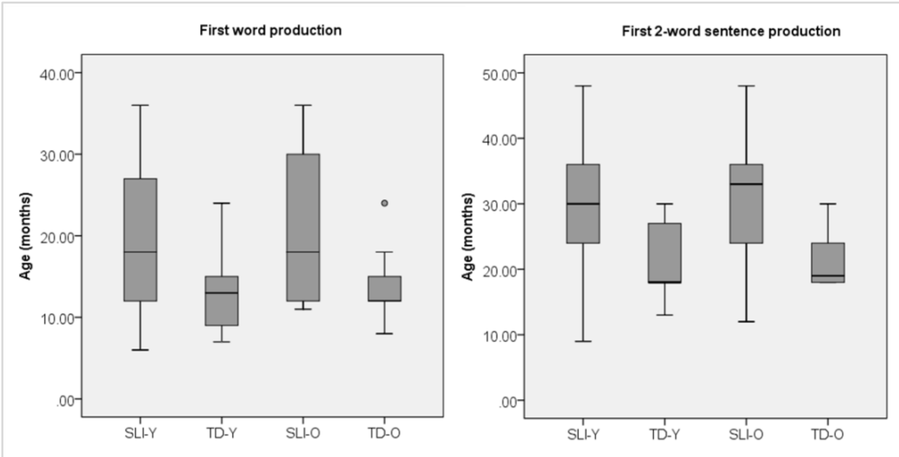


Figure 3.3. Age of first words and two-word sentences across the four groups

3.4 Materials

In the previous chapters it was already indicated that the three grammatical variables were tested using multiple tasks. Since some studies on SLI report differences between outcomes on comprehension and production tasks (e.g., Miller et al., 2008; Keij, 2009) we designed both a perception task (or two in case of the relative clauses) and a production task for each variable.

Perception can be tested in many ways. There is, for instance, a difference between online tasks and offline tasks. The suitability of online tasks like self-paced listening was initially examined, but these tasks proved not to be very sensitive to some of the grammatical variables studied in this thesis (see Marinis, Blom, & Unsworth, 2010 for a discussion of considerations in using online methods). In a self-paced listening study with TD children, errors in grammatical gender did not show a delay in reaction time at the expected post-critical element but the effect was shown at a later point in the sentence (Marinis, Blom, Chondrogianni, & Vasić, 2010; Blom & Vasić, 2011). This may result in undetectable difficulties in the SLI group, because processing speed is often reported to be generally reduced in SLI. There is only a limited number of elements after the critical element in which delays in response time can be measured. We therefore chose to use offline measurements.

Within offline tasks, knowledge of grammatical aspects can be tested in several ways, for instance by selection tasks or judgement tasks. In judgement tasks, participants have to decide whether something is correct, possible, plausible etc. Selection tasks, on the other hand, require the participant to choose from a number of options. The choice for judgement tasks instead of selection tasks was based on the following rationale: in selection tasks, there is no possibility of detecting unstable grammatical representations, i.e. the acceptance of multiple forms. When a child is forced, for instance, to say which form of the determiner is correct with a certain noun, it is no longer possible to observe that a child might accept both and does not have a clear representation of the grammatical gender of this noun. Since variability in the production of the definite determiner has been reported, we want to be able to capture this instability in the perception measure. We therefore chose to use judgement tasks (although the perception of relative clauses was additionally tested with a picture selection task).

A clear disadvantage of the judgement design is that you can only deduce something about the grammatical knowledge of participants when they reject items. If children, for instance, consistently reject object relative clauses, we can deduce that they have not yet acquired the grammatical knowledge of this structure. However, if they accept object sentences, it is not clear whether they have knowledge of the structure or accept the sentences for a different reason. For relative clauses, these drawbacks were circumvented by including simple incorrect sentences and by eliminating those children that failed to reject the simple incorrect sentences. For grammatical gender and subject verb agreement, such a control condition was not implemented (gender and subject verb items served as each other's fillers). Alternatively, the analyses were checked with a smaller sample in which the participants with very few rejections (<10) were eliminated. An exact description of the different judgement tasks can be found in the individual chapters on the three grammatical variables (§5.3.2, §6.3.2 and §7.3.2 respectively).

The production of grammatical elements can also be tested in several ways, for instance by recording spontaneous language samples, or by asking parents or teachers whether children use certain structures. These methods are, however, not appropriate for testing rather infrequent grammatical structures like relative clauses. By the time children have become adolescents, they might also have developed ways to avoid the structures they are struggling with (Nippold, Mansfield, Billow, & Tomblin, 2009). Furthermore, we wanted to manipulate linguistic context factors in order to test the hypothesis that performance in SLI is related to processing demands. As Bishop (1994), who studied spontaneous data, suggested: "experimental studies are needed" to investigate whether performance errors are due to problems in linguistic competence or in performance (Bishop, 1994, p. 508). An experimental design was therefore favoured. Sentence repetition was considered, but although this measure has proven to be useful for testing syntactic knowledge, its usefulness for morphological elements like grammatical gender or subject verb agreement is less clear. We therefore decided to test production with elicitation tasks. These will be described in more detail in the chapters on grammatical variables (§5.3.3, §6.3.3 and §7.3.3 respectively).

Both perception tasks and production tasks involve performance and rule implementation and variability in performance dependent on the linguistic context might theoretically occur in both types of task. For grammatical gender and subject verb agreement, we chose to manipulate the context factors in the production tasks because most of the 'performance effects' described in the previous section were noted in production tasks. For these variables, the influence of linguistic context factors was therefore only analysed in the production tasks (although some factors like noun frequency were automatically included in the judgement task because the same nouns were used in perception and production, but these effects were not analysed). For relative clauses, the context factors were manipulated in the production and the perception tasks.

In order to investigate the link between fluctuations in performance and processing difficulties, several processing measures were administered. Tasks were chosen for skills that showed differences between SLI and TD (Jensen de Lopéz & Baker, 2015), which led to the selection of a visuo-spatial inhibition task and several working memory tasks that differed in amount of verbal load involved. These measures will be described in detail in the following chapter on processing abilities (§4.3).

3.5 Procedure

All subjects were tested in quiet rooms during school time. Some children did not want to be tested during school time and were tested at home (4 SLI subjects with ambulatory care in the older group). Sessions lasted around 4 hours for the participants in the younger groups (experiments were carried out in collaboration with another project, for which additional testing was needed) and around 2.5 hours for the participants in the older groups. The participants in the younger groups were tested on different days, in 4 sessions of approximately 1 hour. The participants in the older groups were tested on a single day, with at least two short breaks between the individual tests. Tests were administered in a random order, but for all grammatical variables, production tasks preceded comprehension and judgement tasks in order to avoid priming and learning effects. Several tasks were administered on a laptop (HP Elitebook 2740p) with a touchscreen. Some tasks were run in E-prime (i.e. the judgement tasks and the visual working memory task) and answers were automatically scored. For other tasks, audio or video recordings were made in order to be able to score the tasks afterwards. All subjects received a small gift in exchange for their effort. The parents of participants received a short (non-diagnostic) summary of the outcomes of the project and schools were offered the opportunity to hear more about the results in a short presentation. The schools for special education of children with language impairments requested more information, and these schools were visited to present the results of the project in more detail.

The data from the processing tasks and the grammatical judgement and comprehension tasks were scored automatically or could be entered easily by the experimenter (scores were either 0 or 1). The data of the grammatical production tasks required transcription and coding, because the answers could have multiple forms. The coding schemes for every grammatical variable are presented in Chapters 5, 6 and 7 respectively, and are added in the appendices.

For most experimental tasks, the data was binary in nature: children judged or produced items either correctly (1) or incorrectly (0). The number of correct items was counted and an accuracy percentage was computed. This was especially necessary in the production task, since the total number of items that could be analyzed varied between children due to missing data. In accuracy percentages information on the number of items is, however, lost, which reduces the statistical power. For binary data, logistic regression is theoretically more appropriate than linear regression. However, because we manipulated linguistic factors in the context in which variables were tested, items did not have an equal probability of being correct/incorrect. Performing binary logistic regressions would therefore require a vast model (it has to compute the analyses on item and participant level) which was not feasible in the statistical program that we used.[7] Statistical analyses were therefore performed with linear models and were – where possible – checked with logistic analyses.

Each research question of this dissertation required a different statistical test. For the first question (whether differences between SLI and TD are persistent), two-way ANCOVAs were used to determine the effect of Group (SLI and TD) and Age (young or old) and the interaction between the two. The SLI and TD groups were well matched on age and gender, but differed in their level of non-verbal IQ, as described in §3.3. Non-verbal IQ was therefore entered as a covariate in all statistical analyses (scaled z-scores were computed to avoid ceiling/floor effects). In case of significant interactions (indicating that the effect of Group changes with Age), a post-hoc ANCOVA was computed to see whether differences between SLI-

[7] The statistical analyses were performed in SPSS.

O and TD-O were still present. Effect sizes were reported for all factors, using partial eta-squared ($\eta_p{}^2$).

The second question (whether there are effects of linguistic context on performance in production) was tested with the help of Repeated Measures. The different were entered as within-subjects factors, and Group was entered as between-subjects factor. If significant interactions between context factors and Group were found, post-hoc analyses were conducted to reveal whether the effect of context was larger in SLI or in TD. For all factors, effect sizes were reported using partial eta-squared ($\eta_p{}^2$).

The third question (whether variability in performance is related to processing abilities) was tested with correlational analyses. For every processing measure and every significant context effect, non-parametric one-tailed Spearman's rho (*r)* was computed to see whether the effects were related to processing ability. One-tailed tests were used because we expected to find negative correlations between processing ability and the effect of context (problems in processing were assumed to lead to larger fluctuations in performance).

As mentioned earlier, the precise details of the materials and analyses are described in the relevant chapters dealing with each grammatical variable.

4 Processing abilities

Human beings are usually capable of processing information quickly and efficiently. They are able to have a conversation in a noisy background, remember a range of instructions and integrate visual and auditory information to unravel the plot in a movie, just to mention a few examples. However, there are large individual differences in processing capacities, especially in more demanding processing tasks (Just & Carpenter, 1992).

Processing abilities play an important role in this study, as was already made clear in Chapter 1. This study aims to test the theoretical claim that impaired processing capacities in SLI not only influence the acquisition of grammatical knowledge, but also impact on the implementation of knowledge, causing variability in performance (see §1.4 for the research questions of this study and §2.5 for the hypotheses). As was briefly discussed in §1.2, there is growing evidence that children with SLI have problems in different aspects of information processing. Differences between SLI and TD are often found in processing, storing and repeating information of different types, both linguistic and non-linguistic (Aram et al., 1984; Im-Bolter et al., 2006; Marton, 2008; Windsor et al., 2008; Schwarz, 2009; Henry et al., 2012; Jensen de Lopèz & Baker, 2015; Lukács et al., 2016). Differences in information processing thus seem to be quite clearly present in SLI, although not all aspects of information processing are equally impaired and few studies have addressed whether those differences are persistent. In order to test the claim that (persistent) variability in grammatical performance is related to processing abilities, it is therefore important to check which aspects of information processing show persistent differences.

In this chapter the processing abilities of the children and adolescents with SLI will therefore be described. The chapter starts with a discussion of different models of information processing (§4.1), and a more elaborate review of reported problems in components of information processing in children and adolescents with language impairments (§4.2). In §4.3, the different processing tasks used in this study are described, and the results are discussed in §4.4. In the final section (§4.5) the assumption that processing capacities are persistently impaired in SLI will be evaluated. The chapter will conclude with the specification of processing measures that are suitable for analysing the relationship with grammatical abilities over time.

4.1 Models of information processing

Information can be perceived via multiple modalities or senses (visually, auditory, tactilely or via taste or smell). In this chapter, the focus will be on the visual and the auditory modality since these are the most relevant for language processing. In the processing of linguistic information, many different capacities play a role. Information has to be attended to, analysed, monitored, interpreted, linked to other information, stored for a an amount of time, passed on to long-term memory, and retrieved when necessary. The processing of information can therefore be unsuccessful at many different levels. Below, the different steps in information processing will be described in more detail.

When we are awake, a considerable amount of information is presented to our eyes and ears, but we do not pay attention to all of it. If we are, for instance, driving a car, we may just focus on the road and the sounds of other traffic, and may not pay attention to the further surroundings. Attention allocation is often a relatively unconscious process, since information that is more relevant or more familiar will automatically be favoured. We regularly detect a familiar word such as our own name in conversations we are not attending to (the cocktail party effect). We can, however, control our focus of attention to some extent. This ability to control attention allocation can be tested using focused or sustained attention tasks.

If information is (consciously or unconsciously) attended to, it has to be filtered. Executive functions (EF) like shifting and inhibition are thought to play a role in selecting which information is relevant for further processing, in determining how to process that information efficiently, in interpreting and inferring the meaning of the information and in passing the information on to long-term memory (Miyake et al., 2000). Taken together, executive functions form a supervisory control system that is involved in all stages of information processing. Several components of executive control have been proposed, but little consensus exists on which aspects should be distinguished as separate components (Vugs, Hendriks, Cuperus, & Verhoeven, 2014). Planning, initiation and monitoring are, for instance, components where authors do not agree. Likewise, the link between executive control and working memory is debated, as will also become clear from the discussion of different working memory models below. Although different components of executive control are often distinguished on a theoretical level, they are highly interrelated and are also connected to memory and attention. They can therefore not be tested easily in isolation. Nevertheless, several tasks have been developed to test the different components of executive control.

Incoming information that is attended to and selected will be stored for a few seconds in what is often called sensory memory (Alain, Woods, & Knight, 1998). If we, for instance, see an array of 10 letters, or hear a 5-syllable nonsense word, we will be able to reproduce some of that information immediately, without having processed the information thoroughly. Sensory memory is generally limited to around 5-7 elements, that can be stored for just a few seconds (Schraw & McCrudden, 2013).

Information that is held in sensory memory can either be passed on to short-term memory/working memory, or can be deleted from the system. If information is passed on, it is stored for a slightly longer period and processed in more detail. Strings of sounds or images are, for instance, decoded, interpreted (linked to other information and to long-term memory) and passed on to long-term memory. Short-term memory and working memory are not equivalent terms, since working memory implies that the stored information is manipulated, while short-term memory can just refer to short-term storage without any operations being applied. Whether or how much information is transferred to long-term memory is affected by a number of factors like repetition and organization of the information. When participants have to memorize words while they perform a subsidiary task (non-verbal, non-related), for instance, transfer to long-term memory appears to be reduced (Baddeley & Hitch, 1974).

In general, processing and storage of information are often thought to draw from a common pool of resources or activation, causing a trade-off between storage and processing capacities (Just & Carpenter, 1992). A heavier demand on processing abilities (e.g., more complex information) will, in this view, restrict the amount of information that can be stored. The view is based on correlations between processing speed and memory span, which are regarded as measures of processing and storage (Case, Kurland, & Goldberg, 1982). The idea of a trade-off between processing and storage has, however, been questioned by more recent studies. Rather than reducing memory span due to a trade-off between resources, slower processing speed may cause decreases in memory span, simply because the additional processing time increases the time in which items may be forgotten (e.g., Towse, Hitch, & Hutton, 1998). Nevertheless, individual differences in information processing might stem from differences in capacity (storage), or differences in the efficiency of mental processes (processing speed) (Just & Carpenter, 1992).

Over the last decades several models of information processing have been proposed of which the two most influential will be discussed here (Baddeley &

Hitch, 1974; Shah & Miyake, 1999; Baddeley, 2003; Gillam, Montgomery, & Gillam, 2009; Baddeley, 2012). Although the two models clearly represent a different view on the relationship between executive control functions and working memory, the two views are not incompatible considering the goals of this study. Furthermore, the systems that are described to play a role in information processing may also apply to production, although the two models are designed to explain perception.

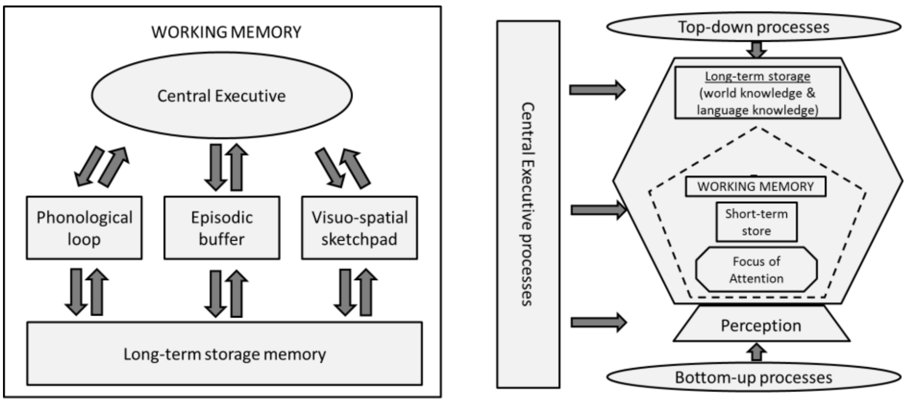


Figure 4.1. Model of working memory (Baddeley & Hitch, 1974, left scheme) and model of information processing (Gillam et al., 2009, right scheme)

In the well-known working memory model by Baddeley and Hitch (1974), working memory is thought of as a multiple-component system, involving a central executive, different short-term memory systems and long-term memory (see Figure 4.1, left). In this model, different types of incoming information enter the central executive, which controls the information flow to the different short-term memory components. Visual information is processed in the visuo-spatial sketchpad while phonological information is processed in the phonological loop. Both loops are linked by the episodic buffer, which binds information from different sources together into episodic chunks (Baddeley, 2003). In order to interpret and manipulate information that is held in the short-term memory components, previously stored information is retrieved from long-term memory. Baddeley and Hitch based their idea of a dissociation between visual and verbal short-term memory on outcomes of dual-processing tasks: if a person has to perform two tasks simultaneously in the same perceptual domain (either verbal or visual), performance is not as efficient in comparison to a single task. If, however, the two tasks address different domains, efficiency is comparable to single task performance. Working memory, according to

Baddeley and Hitch's model, thus involves attention and executive functions (the central executive), short-term memory and long-term memory. In other models, such as the model by Gillam and colleagues (2009, based on Cowan, 1999), working memory is represented as a smaller unit within the cycle of information processing and refers to the capacity to briefly store and manipulate information (see Figure 4.1, right). Executive functions are represented as a separate module, influencing the activation of working memory and focussed attention.

As mentioned above, these two views, although different, are not incompatible in the light of the goals of this study. When testing the hypothesis of a link between processing abilities and variability in language performance, it does not matter whether executive functions are viewed as part of working memory or as a separate module influencing working memory. What is important for this study is that information processing capacities influence language acquisition but at the same time affect language production. Traditionally, information processing models or working memory models depict the processes involved in the perception of information. Models of language production, such as the model by Levelt (1989), on the other hand, often do not include general information processing capacities. Figure 4.2 presents the two models of information processing from Figure 4.1 with an additional production component. Production is represented as a simple component, taking its input from long-term memory (linguistic knowledge, knowledge of the world, episodic memory) or short term memory/working memory, influenced by executive functions.

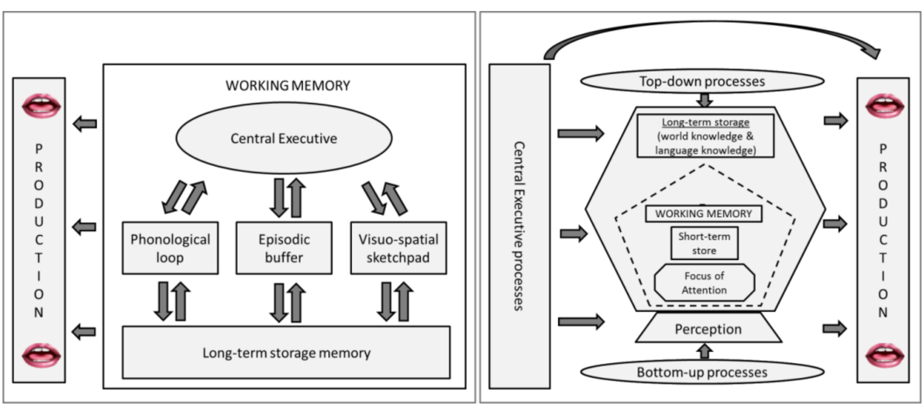


Figure 4.2. Model of working memory (Baddeley & Hitch, 1974, left scheme) and model of information processing (Gillam et al., 2009, right scheme) with an additional production component

It has become clear from the description of the different steps in information processing, that processing abilities cannot be tested with a single measure. In this study, five tasks were chosen that served as measures of processing ability, based on previously attested robust differences between SLI and TD. These measures involve executive control functions (inhibition) and working memory in different modalities (visual and verbal) and with differences in the amount of verbal load (as will be discussed in more detail in §4.3).

4.2 Processing abilities in children and adolescents with SLI

As briefly mentioned in Chapter 1, information processing is often less accurate in individuals with language impairments. This is perhaps not surprising, given the interrelatedness of information processing abilities and language abilities, as outlined in the previous section. A direction of causality is, however, not straightforward. Language problems might be due to problems in information processing, but vice versa, problems in language might also influence the potential for information processing.

As discussed in the previous section, attention and executive functioning play an important role at all levels of information processing. There are several indications that children with SLI have problems in focused or sustained attention and executive functioning. Comorbidity between language problems and ADHD is, for instance, relatively high, much higher than would be expected on the basis of the prevalence of these individual disorders. Children with SLI often meet the criteria for ADHD (Tannock & Schachar, 1996; Botting, 2006; Duinmeijer, 2007) and language problems are often reported in populations of children with ADHD (Love & Thompson, 1988; Blankenstijn & Scheper, 2004; Parigger & Baker, 2005; Jónsdóttir, 2006; Parigger, 2012). The symptoms of ADHD, as indicated in the Diagnostic Statistic Manual of Mental Disorders (DSM), do not only involve poor attention skills as the name of the disorder implies, but also include problems in following and remembering instructions and ordering and structuring tasks and activities. ADHD is thus (partly) defined by problems in EF (Pennington & Ozonoff, 1996; Tannock & Schachar, 1996). The high rate of overlap between SLI and ADHD might therefore be seen as an indication of high rates of attention problems and problems in executive functioning in SLI.

Apart from high comorbidity rates between SLI and ADHD, studies investigating attention and executive functions in subjects with SLI who do not meet the criteria for ADHD also report difficulties in these domains (Kapa & Plante, 2015). Both in the auditory and the visual domain, problems in focused and sustained attention are reported (Snowling et al., 2006; Finneran, Francis, & Leonard, 2009; Spaulding, Plante, & Vance 2008; Gillam & Hoffman, 2004; Noterdaeme, Amorosa, Mildenberger, Sitter, & Minnow, 2001; Dispaldro et al., 2013; Gallinat & Spaulding, 2014). Furthermore, children with SLI have more problems in updating old information when new information comes in, in inhibiting automatic actions or responses (e.g., Bishop & Norbury, 2005; Im-Bolter et al., 2006; Marton, Kelmenson, & Pinkhasova, 2007; Windsor et al., 2008; Henry et al., 2012; Lukács et al., 2016) and in planning (e.g., Marton, 2008; Henry et al., 2012). Parents and teachers also report more difficulty in EF in children with SLI than in TD peers (Wittke, Spaulding, & Schechtman, 2013). Problems in EF is also found in self-reports by adolescents with SLI (Hughes, Turkstra, & Wulfeck, 2009) and atypical scores in attention in children with SLI persist into adolescence (Young et al., 2002; Weismer et al., 2005; Hughes et al., 2009). Problems appear to be more persistent in the auditory domain, and in more complex tasks (Hick et al., 2005; Spaulding et al., 2008). Executive functions in individuals with SLI have received more attention in recent years, but many studies focussed on single functions and did not link the outcomes to language measures.[1] How impairments in executive functions are related to language acquisition in SLI therefore remains undetermined (Kapa & Plante, 2015).

The majority of studies on information processing in SLI have focused on memory components. Children with SLI have problems in the recall of different types of information, ranging from linguistic information like sentences or words to nonlinguistic information like non-speech sounds or visual patterns (Aram et al., 1984; Conti-Ramsden, 2003; Gray, 2003; 2006; Horohov & Oetting, 2004; Schwarz, 2009; Henry et al., 2012). Problems in the visual domain are not found consistently: many studies do report differences (e.g., Hoffman & Gillam, 2004; Bavin, Wilson, Maruff, & Sleeman, 2005; Hick et al., 2005; Menezes, Takiuchi, & Befi-Lopes, 2007; Marton, 2008; Nickisch & Von Kries, 2009; Henry et al., 2012) but Archibald and Gathercole (2006) did not find any effects. Problems in the auditory domain are, on the other hand, consistently found (Schwartz, 2009). Children with SLI have problems repeating strings of auditory information, even when the tasks are simple. However, they seem to have more significant problems in complex tasks – in which verbal storage is for instance combined with verbal or visuo-spatial processing (Weismer, Evans, & Hesketh, 1999; Marton & Schwartz, 2003; Bavin et al., 2005; Hick et al., 2005; Archibald & Gathercole, 2006). Again, studies report that

[1] Although some studies did, for example Dispaldro and colleagues (2013), who found that attention engagement accounted for differences in grammatical comprehension.

problems arise especially in high processing load conditions (Weismer, 1996; Fazio, 1998; Hoffman & Gillam, 2004; Archibald & Gathercole, 2007). The problems with memory components seem to persist into adolescence and adulthood (Aram et al., 1984; Young et al., 2002; Clegg et al., 2005; Weismer et al., 2005; Reed, 2005; Poll, Betz, & Miller, 2010).

Apart from problems in (aspects of) executive functioning and memory components, many studies in SLI report problems in processing speed (Schwarz, 2009). Slower reaction times in SLI have been found both in verbal and visuospatial memory tasks and executive function tasks (Tallal & Piercy, 1973; Miller, Kail, Leonard, & Tomblin, 2001; Schul, Stiles, Wulfeck, & Towsend, 2004; Im-Bolter et al., 2006; Henry et al., 2012). Furthermore, children and adolescents with SLI appear to learn at a slower pace (Tomblin, Mainela-Arnold, & Zhang, 2007; Kemény & Lukács, 2010; Lum, Conti-Ramsden, Morgan, & Ullman, 2014). Although processing speed cannot be entirely separated from processing capacity, it might be a separate, additional source of impairment. On top of a problem in storage capacity or processing efficiency, children with SLI might just process information at a slower pace. Bishop (1994) seems to keep the three possibilities open, by saying performance errors might be due to 'slowed processing, in a limited capacity system that is handling several operations at the same time' (p. 507). Her view of the processing problem in SLI seems to be that it is a combination of problems in processing speed (slowed processing), problems in memory capacity or efficiency (limited capacities) and problems in attention and executive functioning (handling several operations at the same time). In addition, different problems may appear in different children.

Although intuitively the high rate of processing problems in SLI would seem to be a causal factor underlying the language problems in this population, it is hard to prove the directionality of a causal relationship since the influence of processing on language is reciprocal. Problems in information processing may impede language development, but at the same time language problems may hinder efficient and fast information processing. The older a child with SLI becomes, the more the two domains may have influenced each other in a negative way. Adolescents with SLI are, therefore, expected to show persistent problems in processing (Stothard et al., 1998). In order to find out whether differences in processing capacities are due to language problems or vice versa, Henry and colleagues (2012) performed regression analyses on a large data set including children with SLI, children with low language functioning who did not meet the criteria for SLI and those with typical development. Correcting for differences in verbal abilities between individuals, they checked whether group differences in processing abilities would disappear or remain. If group differences remained, this would indicate that poor processing abilities in SLI were unlikely to be the result of their language problem. Their results showed significant difficulties in processing tasks in SLI, even after correction for verbal abilities. Furthermore, the children with low language functioning showed similar difficulties in the processing domain as the group of children with SLI. They therefore concluded that problems in processing abilities are part of the cognitive profile that leads to language impairments.

Because of the robust findings of problems in SLI in auditory working memory tasks, and the fact that these measures seem to have a good sensitivity and specificity in distinguishing children with SLI from typically developing peers, poor performance on a nonword repetition or sentence repetition tasks have often been suggested as good clinical markers for SLI (e.g., Conti-Ramsden, Botting, & Faragher, 2001; Gray, 2003; 2006; Chiat & Roy, 2007; Poll et al., 2010,). Subjects with SLI do not form the only population showing problems on these tasks (which means specificity drops when other impaired populations are included since not only SLI will score in the clinical range). The tasks might, however, provide a valuable tool to distinguish SLI from typical development in adolescence. As discussed in §1.1, diagnosis of language impairments in adolescence is more difficult since language milestones are less clear in this period and individual differences become larger (Nippold, 1995). Impaired processing capacity could therefore be a promising clinical marker of language impairment in adolescence (Poll et al., 2010).

4.3 Selected processing tasks

In order to test processing abilities in this study, several tasks were chosen that had discriminated between SLI and TD in previous studies (Jensen de Lopéz & Baker, 2015). Jensen de Lopèz and Baker (2015) performed a review of the literature on information processing problems in SLI, and concluded that findings are most robust for inhibition and working memory. An inhibition task and four working memory tasks were therefore selected as processing variables in this study. The working memory tasks varied in terms of the modality in which processing was tested (visual or auditory) and the amount of verbal load in the auditory domain (see Figure 4.3). This choice was made in order to be able to investigate the effect of the type of verbal load involved in processing. In recalling spatial positions of pictures, little verbal load is assumed since this is a visual task. The auditory tasks naturally involved language processing, but to different extents. Recall of digits is, for instance, regarded as a working memory task with relatively low verbal load since digits are highly automatized. In the repetition of non-words, the amount of verbal load is thought to be higher. This task relies on phonological processing , in which linguistic factors like phonological skills and vocabulary size have been shown to play a role (Munson, Kurtz, & Windsor, 2005; Rispens, Baker, & Duinmeijer, 2015). Finally, sentence recall was regarded as a task with high verbal load, since all language levels (phonology, morphosyntax, lexico-semantics) are addressed. In the next sections, the different information processing tasks will be described in detail.

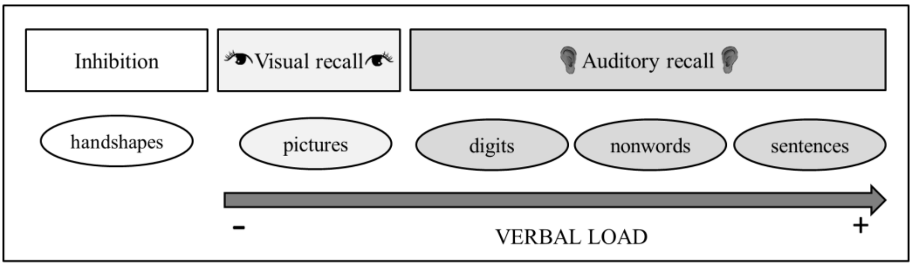


Figure 4.3. Amount of verbal load in the different information processing tasks

4.3.1 Inhibition (executive functioning)

Inhibition was tested using the motor part of an inhibition test battery (The VIMI), designed by Henry and colleagues (2012). This task is based on Luria's Hand Game (Luria, Pribram, & Homskaya, 1964) and includes two conditions that are each tested twice. Initially, the experimenter presents the child with sequences of gestures that are either a pointed finger or a fist. The child has to copy the hand shape after every gesture (copy condition). Subsequently, the rule changes and the child is asked to inhibit the copying response and present the alternative hand shape (inhibition phase). To exemplify: when the experimenter shows a pointed finger, the child also has to appoint her finger in the copy condition, but has to make a fist in the inhibition condition (see Figure 4.4). The same procedure is repeated with a flat horizontal and a flat vertical hand shape (note that the second copy condition involves both switching and inhibition to some extent).

Each condition consisted of 20 trials, so in total 80 trials were administered. The test was videotaped and the total time of performing each condition was recorded. Correct and incorrect responses were coded offline. The slightest hesitation or incorrect hand shape was scored as an error. Although the number of errors in each condition are usually combined and taken together as a measure of inhibition, this study calculated separate scores for the copy and inhibition conditions. The difference between the copying and inhibition accuracy was computed and taken as a measure of inhibition. The rationale behind this is that a total score of copy and inhibition trials does not distinguish between general low performance and problems in inhibition. A child might, for instance, produce errors in the copying phase due to motor problems or problems in coordination. The difference between the inhibition and the copy condition, on the other hand, takes general performance problems into account and is therefore considered a better measure of inhibitory control.

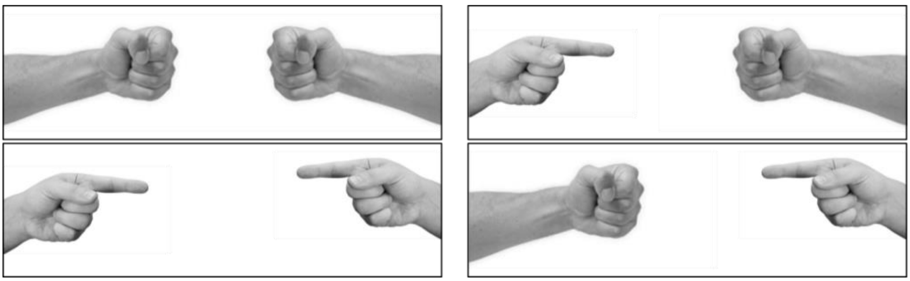


Figure 4.4. Example of copying (left) and inhibition (right) in the motor inhibition task (Henry et al., 2012)

4.3.2 Visual recall: odd-one-out

To test information processing in the visual domain, visuo-spatial recall was tested using the odd-one-out task (Henry, 2001, based on Hitch & McAuley, 1991). In this task, children were presented with sequences of pictures with three nonsense figures. In each picture, one of the figures was different from the other two. Children first had to discover the odd one out in every picture by pointing to it. After the sequence was finished, the children were asked to indicate in empty boxes where the odd ones had been (see Figure 4.5 for an illustration of a two-sequence item). Sequences started with lists of one item and continued with longer lists up to six items. Per list length, four trials were presented and the test was stopped if two out of four trials (50%) were incorrect. If the child pointed to the wrong picture in the selection of the odd one out (perhaps due to problems in pattern recognition or attention), the incorrect response was taken as input for the memory part of the test. The results are therefore not dependent on pattern recognition but are thought to reflect visuospatial working memory. The odd-one-out task has been used in previous research by various authors to measure visual-spatial working memory in SLI, although there are differences between studies in the number of items per list length and the break-off rule used (e.g., Henry et al., 2012; Lukács et al., 2016). Furthermore, some subjects remembered positions by rehearsing left-right in their head, which means the task is not necessarily entirely visual. A recent study by Botting & Cowan on the use of verbal strategies in a visual recall task indicated that TD children especially seem to use those strategies (Botting, 2014). They showed that if visual recall was tested with an interfering articulation task (impeding the use of verbal strategies to recall the visual information), the scores dropped in both the SLI and the TD group, but relatively more in the TD group.

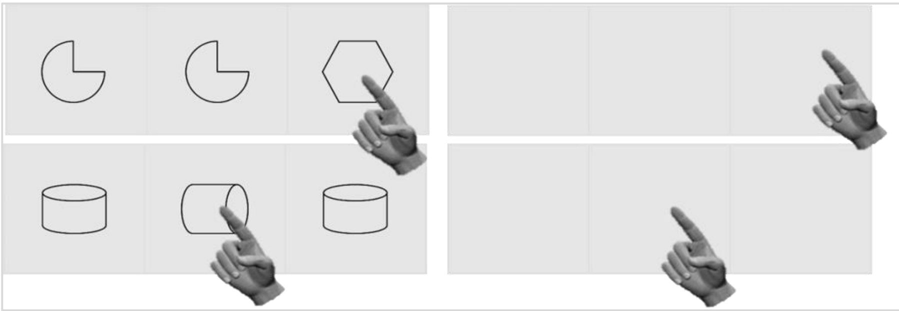


Figure 4.5. Example of picture recall in the odd-one-out task (Henry, 2001)

The task was administered on a laptop with a touch screen, and responses were coded automatically. The total number of correct items and the total number of administered items were scored and a memory level was computed by counting the number of pictures that could be recalled correctly in a row (a corrected score was also computed on the basis of the guidelines used by Henry et al., 2012). However, since variance becomes smaller in the conversion of the scores to a memory level, the total number of items correct was ultimately taken as the outcome measure since this is more sensitive.

4.3.3 Auditory memory: digit recall

In order to measure auditory memory in a task with low verbal load, the digit span task of the Dutch version of the Wechsler Intelligence Scale for Children (WISC-III-NL) was used (Kort et al., 2005). In this task, children had to repeat increasing sequences of numbers (both forward and backward) that were orally presented to them by the experimenter (see Figure 4.6. for an example). Sequences ranged from three to eight digits and for each length two trials were administered. If the child failed on both trials of a sequence length, the test was broken off. The WISC-III is a widely used standardized intelligence test, and the digit span subtask has been used frequently to measure auditory working memory in various populations. To detect a possible difference between short-term memory and working memory abilities, the number of digits that could correctly be repeated forwards and backwards was calculated. Furthermore, the total number of items correct (forward plus backward) constituted a raw score (RS), which was converted into a standard score (SS) on the basis of norms provided by the manual (Kort et al., 2005).[2]

Forward:		Backward:	
Experimenter:	2-6-9-5	Experimenter:	3-8-9
Participant:	2-6-9-5	Participant:	9-8-3

Figure 4.6. Example of forward and backward digit recall in the digit span task (Kort et al., 2005)

4.3.4 Auditory memory: nonword repetition

To measure auditory memory in a more verbally loaded task, a Dutch nonword repetition task was used (Rispens & Baker, 2012). In this task, children were orally presented with 40 nonwords that varied in length (from two to five syllables) and phonotactic probability (high/low). The items followed Dutch phonological rules, and all syllables consisted of consonant-vowel pairs, while avoiding consonant clusters. Children were asked to listen carefully to each word and repeat the nonwords as faithfully as possible. The task has been used in many studies on Dutch populations of children with language disorders and TD children (e.g., Rispens & Baker, 2012; Duinmeijer et al., 2012; Rispens et al., 2015). The responses of the subjects were recorded and later transcribed by the researcher. The percentage of correctly repeated nonwords was calculated (total %) and divided into percentages correct per word length (two, three, four or five syllables) and phonotactic probability (high or low) (see Figure 4.7 for examples).

	Low phonotactic probability	High phonotactic probability
2 syllables	*weugof*	*raanom*
3 syllables	*veujoetup*	*loowaamas*
4 syllables	*fuiseuwoesut*	*saaviebeemer*
5 syllables	*fuugiwuinoefep*	*baamerienooves*

Figure 4.7. Examples of nonwords in the Dutch nonword repetition task (Rispens & Baker, 2012)

[2] The standard score was used in the analyses of correlations between processing and grammatical measures.

The items were scored on the word level, meaning that one incorrect sound in the repeated nonword resulted in a 0-score. An alternative way of scoring this task is by analysing recall at the phoneme level (Rispens & Baker, 2012). Although this level of measurement is thought to be more sensitive to individual differences (Graf Estes, Evans, & Else-Quest, 2007), scoring at the word level has been found to be sufficient to reveal group differences (Conti-Ramsden et al., 2001).

4.3.5 Auditory memory: sentence repetition

To measure auditory memory in a task with high verbal load, the sentence recall task of the Clinical Evaluation of Language Fundamentals, which was part of the general language ability measurement in this study (§3.2), was used *(*CELF-4-NL, Kort et al., 2008).[3] Children were presented with sentences of increasing length and complexity (with different starting points for different ages), and had to recall the sentence as exactly as possible. Every change in the morphological, syntactic or semantic structure of the sentence was counted as an error. Sentences without any errors were given 3 points; sentences containing one error were credited 2 points; two or three errors resulted in a score of 1 point and no points were credited if the sentence contained more than three errors. The test was scored online and the experimenter stopped the task if children obtained five 0-scores in a row. The sum of scores constituted a raw score, which was converted into a standard score on the basis of the Dutch norms (Kort et al., 2005).

4.4 Results

In the following paragraphs, the results of the five different processing measures are shown (§4.4.1-§4.4.5). For all five processing tasks, statistical tests were performed using analyses of variance with Group and Age as independent variables, and with non-verbal IQ as a covariate (ANCOVAs). Because the likelihood of finding a significant result based on coincidence increases in performing multiple comparisons, the *p*-level was adjusted to .01 using a conservative Bonferroni correction (.05/number of comparisons). After the discussion of the separate results, an overview of the effect sizes of different processing measures will be provided.

The results of the inhibition task are shown in Figure 4.8. For each group, the range of accuracy percentages is represented by the boxes and the whiskers (the bold line in the middle of the box represents the median). The left side of the graph shows

---

[3] Recall that the general language measure was used to exclude typically developing children with low language functioning. It was, however, not used as an inclusionary task for the SLI-group and can therefore be regarded as an independent processing measure.

accuracy percentages for the copy trials and the inhibition trials. Statistical analyses were, however, performed on the difference between the two, because the difference score is considered the best measure of inhibitory control (see §4.3.1). The difference scores across the four groups are shown on the right side of the graph. For inhibition, no significant effects of Group (p = .725) or Age (p = .091) were found.

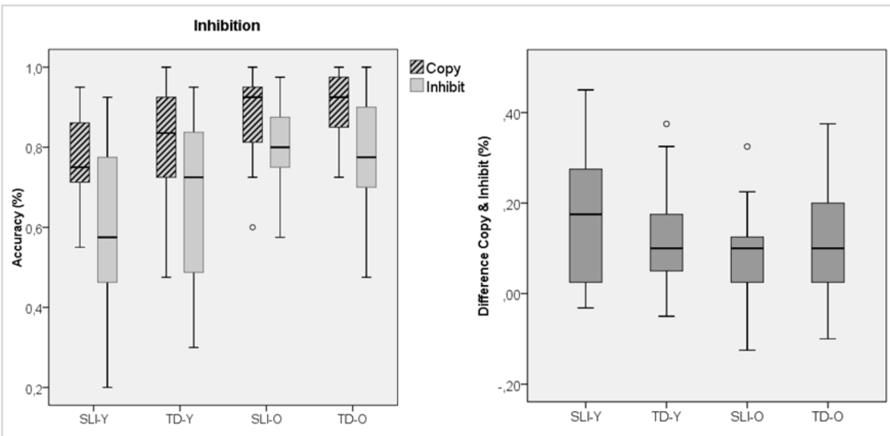


Figure 4.8. Inhibition (copy accuracy, inhibit accuracy and a difference score between the two) across the four groups

Inhibition, as tested with the motor inhibition task of the VIMI and computed on the basis of the difference between copy and inhibit trials, did not show significant differences between SLI and TD.

Figure 4.9 shows the results of the visual recall task and the digit recall task. Differences between the SLI and the TD groups are clearly present in both tasks, although the differences seem to decrease with age. In general, the ability to recall pictures increases with age, as expected. Such an effect is not visible in the digit recall results because these results are standard scores in which age is accounted for (the standard scores have an average of 10, and a standard deviation of 3). The younger TD group is scoring well above the normal range, while the older TD group is scoring in the normal range (at the lower end). The older TD group is therefore more representative of the TD population (as well as being a better match to the SLI group in terms of non-verbal IQ and SES).[4]

[4] SES has been found to influence EF significantly (Hackman, Gallop, Evans, & Farah, 2015).

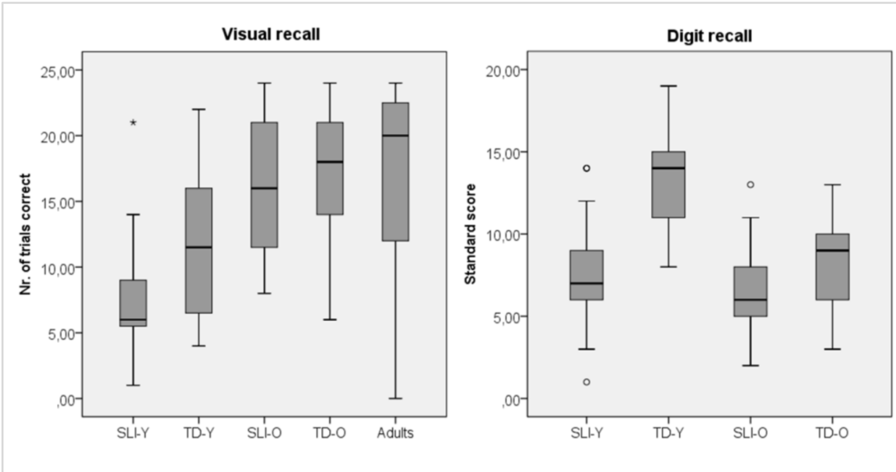


Figure 4.9. Visual recall and digit recall across the four groups

For visual recall, the effect of Group is not significant ($F(1,120) = 5.69$, $p = .02$, $\eta_p^2 = .045$) while the effect of Age is significant ($F(1,120) = 69.06$, $p < .001$, $\eta_p^2 = .365$).[5] Post-hoc analyses reveal that there is a significant difference between SLI and TD in the younger groups ($F(1,61) = 9.66$, $p = .003$, $\eta_p^2 = .137$), but not in the older groups ($p = .801$).

On digit recall, a significant effect of Group is found ($F(1,121) = 40.27$, $p < .001$, $\eta_p^2 = .251$), as well as a significant effect of Age ($F(1,121) = 12.86$, $p < .001$, $\eta_p^2 = .097$) and a significant interaction between Group and Age ($F(1,121) = 16.68$, $p < .001$, $\eta_p^2 = .122$). The difference between SLI and TD in recalling digits is large in the younger groups ($F(1,61) = 48.14$, $p < .001$, $\eta_p^2 = .441$) but not significant in the older groups ($p = .092$), as is shown by post-hoc analyses.[6] To summarize, differences in visual recall and digit recall do not seem to be persistent in SLI.

In Figure 4.10, the outcomes of the nonword repetition task and the sentence repetition task are presented. The nonword repetition data represents the percentage of accurately repeated words, while the sentence repetition scores are standard scores (corrected for age, with an average of 10 and a standard deviation of 3). Again, the younger TD group performs well above the population mean, while the

[5] recall that we adjusted the *p*-level to .01 to correct for the increased risk of finding significant results while there is actually no effect

[6] Differences between TD and SLI are similar in forward and backward recall, the only difference being that the difference in backward recall is significantly different in the older groups as well ($F(1,58) = 4.24$, $p = .044$, $\eta_p^2 = .067$).

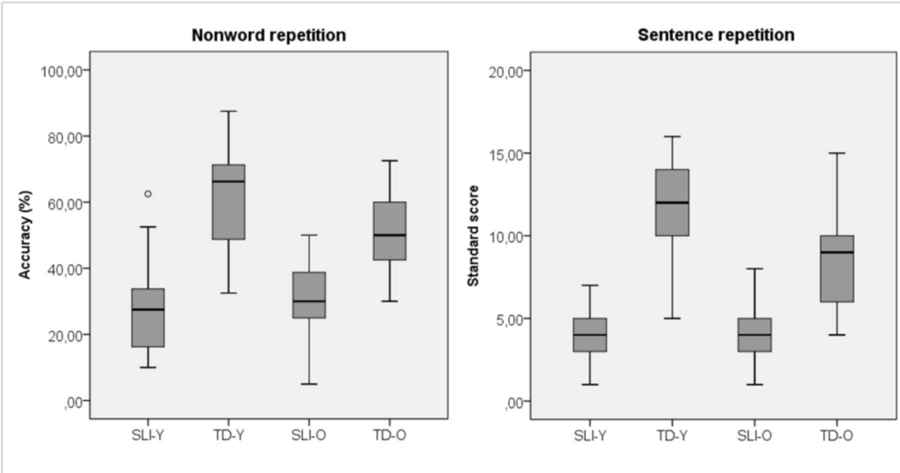


Figure 4.10. Nonword repetition and sentence repetition across the four groups

scores of the older TD group are just below the average. Both tasks show large differences between the SLI and the TD groups.

Statistical analyses show large effects of Group (F(1,120) = 112.14, $p < .001$, $\eta_p^2 = .483$) and no effect of Age ($p = .366$) for the nonword repetition task. The interaction between Group and Age is significant (F(1, 120) = 8.85, $p = .004$, $\eta_p^2 = .069$) and post-hoc analyses reveal that the difference between SLI and TD is largest in the younger groups (F(1,61) = 79.32, $p < .001$, $\eta_p^2 = .565$) but still large in the older groups (F(1,58) = 35.59, $p < .001$, $\eta_p^2 = .380$).[7]

Similar results are obtained for sentence repetition. The analyses reveal a large effect of Group (F(1,119) = 206.88, $p < .001$, $\eta_p^2 = .635$), a smaller effect of Age (F(1,119) = 9.47 $p = .003$, $\eta_p^2 = .074$) and a significant interaction between Group and Age (F(1,119) = 16.41, $p < .001$, $\eta_p^2 = .121$). The difference between SLI and TD is largest in the younger groups (F(1,60) = 170.58, $p < .001$, $\eta_p^2 = 740$) but still very large in the older groups (F(1,58) = 57.54, $p < .001$, $\eta_p^2 = .498$).

On the basis of these results the assumption that processing capacities are limited in SLI can now be evaluated. In Table 4.1, the results of the different

[7] The effects of phonotactic probability significantly influenced accuracy, but more so in the TD groups than in the SLI groups, as is expected on the basis of the literature. Syllable length affected nonword repetition accuracy in all groups, but not significantly different in SLI than in TD. These effects were not shown in detail because they are not relevant for the research questions of this study.

measures are combined and the effect sizes of the differences between SLI and TD are shown for the younger and the older groups (the outcomes of the post-hoc analyses).

Table 4.1. Effect sizes (η_p^2) of the difference between SLI and TD in the younger (Y) and the older groups (O) for the five different processing measures

	Inhibition η_p^2	Visual recall η_p^2	Digit recall η_p^2	NWR η_p^2	SR η_p^2
SLI-Y vs. TD-Y	.014	.137*	.441*	.565*	.740*
SLI-O vs. TD-O	.020	.001	.048	.380*	.498*

Table 4.1 shows that the effect sizes increases as processing measures become more verbally loaded. As was discussed in §4.3, working memory was tested in different modalities (visual or auditory) and within the auditory modality, the amount of verbal load was varied. The order of effect sizes of the different working memory tasks corresponds to the ordering of the tasks in terms of the hypothesized amount of verbal load (visual recall > digit recall > nonword repetition > sentence repetition).

For all five processing measures, the effect sizes are larger in the younger group than the older groups. The difference between SLI and TD thus seems to decrease with age. Some caution is required in drawing this conclusion since, as noted earlier, the older groups were also better matched than the younger groups, in terms of IQ, SES and educational level. Although we tried to correct for the large differences on these scales in the younger groups by entering IQ as a covariate, we cannot be sure that the differences were entirely accounted for (see Chapter 8 for a more elaborate discussion of this issue).

4.5 Conclusion/discussion

The results on the different processing measures show that the assumption that processing abilities are limited in SLI needs some refinement. Whether processing limitations are found seems to be linked to the amount of verbal load in a processing task. Furthermore, differences between SLI and TD in processing abilities seem to decrease with age and persistent differences are only found in the tasks with a high verbal load (nonword repetition and sentence repetition).

The finding that effect sizes increase with an increasing amount of verbal load in processing tasks (which is also reported by Meir, to appear) may provide a new perspective on inconsistencies in the literature regarding non-verbal processing abilities in SLI. As was discussed in §4.2, working memory differences between SLI and TD are quite robustly found in the verbal domain, while problems in the visual domain are not found consistently. Many studies do find differences in visual working memory in SLI (e.g., Henry et al., 2012) while others report absence of any effects (e.g., Archibald & Gathercole, 2006). This may be due to the fact that effects are smaller in the non-verbal domain and thus the significance level is much dependent on aspects like sample size. Henry and colleagues had a fairly large sample size of SLI (N=41) while Archibald & Gathercole tested visual memory in a small sample of children with SLI (N=15). The current study found a significant effect of SLI in visual memory (despite correction for non-verbal IQ), but the small effect size (.124) suggests that the effect is not very robust and would not be replicated in studies with smaller sample sizes. Similar lines of reasoning apply to contrastive findings on non-verbal executive functions like motor inhibition. Despite the fairly large sample size in this study, inhibition appeared here not be different in SLI compared to TD – although we used the same task as Henry and colleagues (2012) who did find an effect. Here, another factor may play a role. As was discussed in §4.3.1, the usual way of scoring the motor inhibition task of Henry and colleagues (2012) is to combine the accuracy rates in each condition (both copy and inhibition). We, however, chose to compute the difference between the two conditions, because the total scores do not distinguish between general low performance and problems in inhibition. The difference between the inhibition and the copy condition was therefore considered a better measure of inhibitory control. Significant differences between SLI and TD found previously with the same inhibition task may therefore be due a bias in the scoring procedure. They may reflect problems in motor control in SLI rather than problems in inhibition.

The fact that the discriminative power of processing tasks seems to decrease with age might be another source of contradictory results between studies. Similar tasks might produce different results, depending on the age of the subjects. Changes over time in the discriminative power of processing tasks have also been found by Jensen de López and Søndergaard Knudsen (2014). They reported that differences in non-verbal working memory disappeared in the older group (14-year-olds) in comparison to younger subjects, while differences in verbal measures remained. In this study, only nonword repetition and sentence repetition distinguished SLI subject from TD subjects in adolescence. Effect sizes are, however, fairly large. Because language milestones in adolescence are less clearly defined and language assessment tools for adolescents are not available in some countries, these processing tasks might be valuable in the diagnosis of adolescents with language impairments, as has been suggested by Poll and colleagues (2010).

As Kapa and Plante (2015) stated on the basis of their review of the studies reporting processing limitations in SLI, the relation between general processing abilities and language acquisition remains undetermined. This study aims to test the claim that processing limitations can account for (persistent) variability in grammatical performance in SLI. Processing measures that show persistent limitations in SLI are hypothesized to be correlated with variability in grammatical performance. Because persistent differences only appeared in the nonword repetition and the sentence repetition tasks, these are the two processing measures for which (negative) correlations are expected with the grammatical results in the following chapters. Correlations for the other measures will, however, be computed in order to be able to link the results to previously reported relations or the absence thereof between grammar and processing measures. These analyses will not provide information on the direction of causality in the co-occurrence of processing limitations and language problems in SLI. They will, however, provide some more insight into the way processing abilities are linked to variability in language performance.

5 Grammatical Gender

As discussed in §2.5, this thesis aims to answer the question whether differences in grammatical abilities that are often reported in children with SLI remain into adolescence. It also investigates whether performance in certain aspects of grammar is influenced by the linguistic context in which they are tested, and whether fluctuations in performance can be explained by processing abilities. These processing abilities of children and adolescents with SLI were analysed and evaluated in the previous chapter. It was clear from discussion in Chapter 2 that several grammatical variables are vulnerable in SLI including grammatical gender, verb agreement and relative clauses. In this chapter and the following two chapters, the grammatical abilities of children and adolescents are evaluated, each chapter dealing with a separate grammatical variable. The grammatical variable to be discussed in this chapter is grammatical gender.

In many languages, nouns are classified into categories on the basis of their gender. Gender systems can have a semantic basis, or can be (more) related to form, that is when phonological/morphological properties of words determine the gender classification (Corbett, 1991). Some languages, like Dutch, have neither a clear semantic basis nor a clear formal basis for their gender classification. In these languages, agreement for gender in other linguistic elements is the clearest indication of the gender of a noun. Hockett (1958) therefore commented that "genders are classes of nouns reflected in the behaviour of associated words" (p. 231).

This chapter discusses grammatical gender in Dutch children and adolescents with SLI and their TD peers. The first section describes the Dutch grammatical gender system (§5.1), followed by a review of the literature on Dutch gender acquisition in SLI and the specific hypotheses for this grammatical variable (§5.2). The tasks used to test comprehension and production of grammatical gender will be presented in §5.3, the results discussed in §5.4, ending with a discussion and conclusion (§5.5).

5.1 The Dutch grammatical gender system

Dutch has a two-way grammatical gender system that distinguishes *common* and *neuter* nouns.[1] Each noun has a grammatical gender, which is reflected in several grammatical elements that accompany the noun or make reference to it. Gender is, for instance, reflected in definite determiners, but also in demonstratives, attributive adjectives and the pronominal system. In this thesis, the focus is on the expression of grammatical gender in definite determiners and attributive adjectives, since these elements have been studied before in Dutch SLI.

The Dutch gender system lacks transparency, since the gender of a root noun cannot automatically be inferred from its phonological, morphological or semantic characteristics (Booij, 2002). The system does contain morphological cues in the nouns and some lexical patterns in the gender distribution of nouns. Diminutives, for example, always have neuter gender (*de poes* 'the cat' – *het poesje* 'the cat.DIM'). Infinitival verb forms used as nouns and nominalized verb forms with the prefix *ge-* are also neuter (*het spelen* ‘the playing’, *het gespeel* 'the playing'). Whereas neuter is also assigned to nouns with the suffix *-isme* (*het Marxisme* 'the Marxism'), nouns ending in the suffix *–heid* or *–ine* always have common gender *(de blijheid* 'the happiness', *de discipline* 'the discipline'). On the lexical level, some semantic classes have predictable gender, such as sports or metals (always neuter) and plants and seasons (always common) (Haeseryn, Romijn, Geerts, de Rooij, & van den Toorn, 1997; Polišenská, 2010). However, in the majority of cases, the gender of a Dutch noun is unpredictable.

The best cue for deducing the gender of a Dutch noun in context seems to be the accompanying singular definite determiner, since it is highly frequent and appears close to the noun. Common nouns are combined with the definite determiner *de*, and neuter nouns are combined with the definite determiner *het* (see Table 5.1 for some examples).[2] Gender is neutralized in plural contexts, since the same form is used with both common and neuter plural nouns (the same form as the singular determiner for common gender). In indefinite contexts, the same indefinite determiner *een* is used for all singular nouns and in plural contexts there is usually no determiner at all.

[1] In older forms of Dutch, the gender system was threefold, distinguishing between feminine, masculine and neuter. In standard Dutch, feminine and masculine have merged into common, but some Dutch dialects still preserve the feminine/masculine distinction (Audring, 2006).

[2] Hence most grammars of Dutch use the notions *de*-nouns and *het*-nouns.

Table 5.1. The Dutch determiner system

	Definite		Indefinite	
	Common (C)	Neuter (N)	Common (C)	Neuter (N)
Singular	***de*** *kat* 'the cat'	***het*** *paard* 'the horse'	***een*** *kat* 'a cat'	***een*** *paard* 'a horse'
Plural	***de*** *katten* 'the cats'	***de*** *paarden* 'the horses'	***Ø*** *katten* 'cats'	***Ø*** *paarden* 'horses'

As is obvious from the above, the common form *de* is used in many contexts. There are also far more common gender nouns than neuter nouns (2:1 in terms of types and 3:1 in terms of tokens in adult usage, van Berkum, 1996), again leading to a greater frequency of the form *de*. In combination with the fact that the common form appears in the neutralized context, the greater frequency often forms the basis for the claim that common gender (*de)* is the default in Dutch.[3] Error patterns of children, second language learners and language impaired populations invariably show overgeneralization of the common definite determiner to neuter nouns, while the opposite overgeneralization rarely occurs (Bol & Kuiken, 1988; Blom, Polišenská, & Weerman, 2008b; Orgassa, 2009). In terms of acquisition of this grammatical variable, common gender therefore seems to be the default.

In order to assign the correct definite determiner to a noun, learners have to know the grammatical gender of the noun and choose the correct form of the definite determiner. If they have not yet acquired a gender feature, they can also rely on the holistic form of determiner and noun taken together. For example they can store the fact that *paard* 'horse' is accompanied by *het*, without having specified the noun *paard* 'horse' as neuter. When a learner produces the correct definite determiner, it is therefore not clear whether this is due to the correct grammatical gender feature being present on the noun or to correct storage of determiner and noun combinations (chunking).

Another domain in which grammatical gender is expressed in Dutch is the inflection of adjectives. Here, however, the gender distinction is less clearly visible

[3] Looking at relative frequency or morphological dominancy is not the only way to deduce the default grammatical gender. Some authors assume neuter to be the 'linguistic default', in contrast to common as the 'learner default' in Dutch. See for more information Kester, 1996; Roodenburg & Hulk, 2008 and Tsimpli & Hulk, 2013.

than in the determiner system. A gender distinction is made only in the use of attributive adjectives, which are part of the noun phrase and are placed between determiner and noun. Attributive adjectives are always inflected with a schwa (ə), except in one particular case: if the noun is neuter and used in an indefinite, singular context, the schwa is not used and a bare form is used instead (see Table 5.2). In contrast to the determiner system where the gender distinction appears in the definite singular context, the gender distinction in adjectival inflection is thus only visible in the indefinite singular context.

Table 5.2. Attributive adjectival inflection in Dutch

	Definite		Indefinite	
	Common (C)	Neuter (N)	Common (C)	Neuter (N)
Singular	*de rod-**e** kat* 'the red cat'	*het rod-**e** paard* 'the red horse'	*een rod-**e** kat* 'a red cat'	*een rood-**Ø** paard* 'a red horse'
Plural	*de rod-**e** katten* 'the red cats'	*de rod-**e** paarden* 'the red horses'	*rod-**e** katten* 'red cats'	*rod-**e** paarden* 'red horses'

Again, the common form of the attributive adjective, that is with the schwa-ending (ə), is often considered the default, since it is appears in most contexts of the paradigm and it is highly frequent in the attributive context. In language acquisition and in populations with language problems or Dutch as an L2, it is this common form of the indefinite adjective that is most often overgeneralized to the neuter indefinite context (Weerman, Punt, & Bisschop, 2006; Blom et al., 2008b). This behaviour cannot be explained by the fact that adjectives are almost always inflected. In predicative contexts, no distinction between common and neuter gender is made and adjectives always have the bare form (*het paard is rood-Ø* 'the horse is red'*)*. Learners thus quite frequently encounter bare forms of adjectives (although in a different sentence structure) and they do not seem to make inflection errors in predicative contexts (Weerman et al., 2006).

In order to produce the correct form of the adjective, it is less feasible that learners use the 'chunking' option discussed above for definite determiners. Storing the adjective and noun (e.g., *rood paard* 'red horse') together will not provide a learner with the correct form since the inflection is dependent on whether the context is definite or indefinite (*een rood paard* 'a red horse' *vs. het rode paard* 'the red horse'). Storing determiner, adjective and noun together as a chunk (*een rood paard,* 'a red horse'*)* is possible, but subjects would have to store the right inflection for the whole range of possible adjectives that can be used with the noun (*een rood paard* 'a red horse', *een lief paard* 'a gentle horse', *een wild paard* 'a wild horse', *een vuil paard* 'a dirty horse', etc.) and there are also indefinite contexts in which the indefinite determiner is absent (*mooi paard heb je daar!* 'what a beautiful horse you've got there!'). Although the storage of chunks might be an initial stage in the discovery of the adjectival inflection system and might explain some correct productions, it is not an economical strategy. A more economical solution is to apply a rule that states:

(1) a. use a bare form in an attributive, indefinite, neuter, singular context
/Ø/ → [+attr, -def, +neuter, +sg]
b. in all other attributive cases, use a schwa
/ə/ → [+attr]

Note that the rule only produces correct neuter bare forms of the adjective if a learner knows that the gender of a noun is neuter (or that the noun is used with the determiner *het*). Errors in adjectival inflection in neuter contexts may therefore either reflect that learners have not yet discovered the exceptional rule, or that they have incorrectly assigned common gender to the noun. One way to find out which explanation is correct is to look at determiner assignment and adjectival inflection within the same neuter nouns. If neuter nouns are produced with a common determiner consistently, this indicates that an individual has not assigned neuter gender to that noun (or has not discovered a gender feature at all). In those cases, the exceptional rule for adjectival inflection does not apply either, whether acquired or not. If a noun, on the other hand, is consistently used with a neuter determiner, this provides a good testing ground for the presence of the rule for adjectival inflection. If those 'stable neuter nouns' are combined with bare forms of the adjective, the rule clearly has been acquired. If, on the other hand, 'stable neuter nouns' are inflected with a schwa, the rule has not been learned yet.

Several authors (Hawkins & Franceschina, 2004; Blom et al., 2008b; Tsimpli & Hulk, 2013) propose that children who acquire Dutch as their first language start without a specification of gender on nouns. As described above, there are some morphological and semantic cues in the Dutch gender system, but many of these cues involve word classes or derivations that are not typically used in child-directed speech and therefore not very salient in children's input (an exception is the diminutive, which is quite frequent in child-directed speech, Souman & Gillis, 2007). It is therefore plausible that children start with a stage in which a gender feature has not yet been discovered and that the first correctly produced neuter determiners are based on lexical storage of chunks. According to a study on Dutch determiner assignment and adjectival inflection (Blom et al., 2008b) a rule for adjectival inflection does, however, seem to be present from very early on (Figure 5.1).

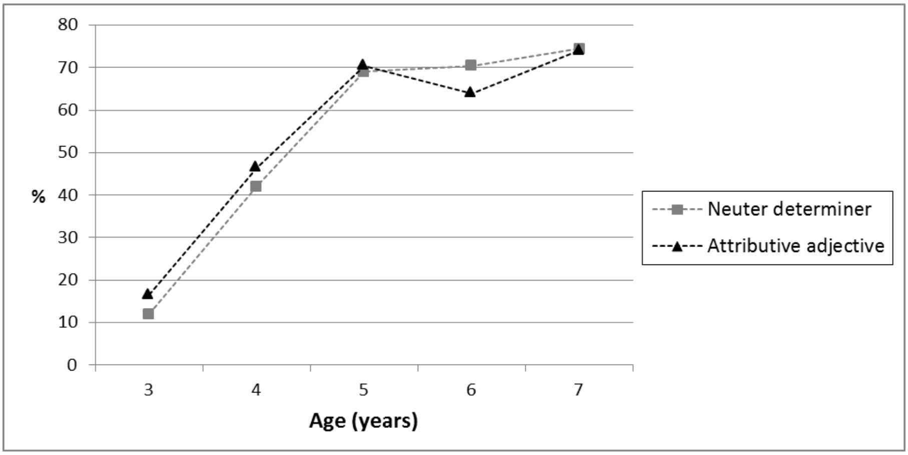


Figure 5.1. Development of neuter definite articles and attributive adjectival inflection for stable neuter nouns in monolingual children (Blom et al., 2008b, p. 320)

As can be seen from Figure 5.1, from three years onwards, the number of attributive adjectives that are correctly inflected follows the same course as the number of neuter nouns for which they use a neuter determiner consistently. It thus seems as if children, once they use the correct determiner *het* for a noun they also correctly inflect the adjective. Thus, an abstract gender feature and the exceptional rule for adjectival inflection seem to be present from early on.[4] What develops over time is the number of nouns that are stored with a neuter determiner or that are assigned a neuter gender feature. Children do not reach attainment levels of 90% correct neuter gender assignment until age 7 (Weerman et al., 2011, Tsimpli & Hulk, 2013). In most other languages (e.g., Spanish, French, Greek, German, Czech and Russian) 90% levels of correct gender assignment are reached much earlier (Tsimpli & Hulk, 2013; Janssen, 2016). The first appearance of Dutch neuter determiners is also later than in German, which also has neuter gender (Mills, 1985). Dutch neuter gender

[4] Different criteria have been used to determine when the system has been acquired. Some authors, (e.g., Tsimpli & Hulk, 2013 – based on Brown, 1973) take 90% correct assignment of neuter gender to nouns as the threshold for acquisition of grammatical gender in Dutch. Others (e.g., Blom et al., 2008b; Weerman et al., 2011) look at whether neuter gender is expressed consistently across different grammatical elements.

thus seems to form an exceptionally difficult case, and hypotheses regarding the acquisition of grammatical gender need to be language-specific. In the following section, the literature on grammatical gender in Dutch SLI will be reviewed and the hypotheses regarding the acquisition of grammatical gender in the younger and older SLI groups and TD groups will be specified.

5.2 Grammatical gender: literature review and hypotheses

As was described in the previous section, a Dutch grammatical gender feature seems to be discovered early in typical development, but ultimate attainment of neuter gender takes a considerable time. Until a fairly late age, common forms are overgeneralized to neuter contexts, both in the use of definite determiners and in the inflection of attributive adjectives. The late acquisition of neuter gender in Dutch is plausibly related to the lack of transparency of the gender system (Corbett, 1991, p.87). Because children with SLI often have problems with aspects that are non-transparent, it is not surprising that Dutch grammatical gender is one of the grammatical aspects that have proven to be particularly difficult for children with language impairments.

The acquisition of Dutch gender was studied earlier in work on effects of bilingualism and second language acquisition, but work on language impairments started only a few years ago. Orgassa & Weerman (2008) were among the first to report that gender is vulnerable in Dutch children with SLI. In a study that tried to disentangle effects of SLI and bilingualism, children with SLI and child L2 learners were shown to make the same type of errors as younger monolingual TD peers (overgeneralisation of common to neuter), while adult L2 learners appeared to make a different type of error (overgeneralisation in both directions). SLI and bilingualism both had negative effects on the acquisition of grammatical gender: a group of bilingual children with SLI performed worse than both the monolingual SLI group and the bilingual TD group (Orgassa & Weerman, 2008; Orgassa, 2009).

Another study that looked at the acquisition of Dutch grammatical gender in language impaired and bilingual populations indicated that there are different stages in discovering the determiner system (Keij et al., 2012). Children with language impairments between 8 and 12 years of age seem to fossilize in the first stage of discovering the gender system - that is the stage in which they do not yet have a gender specification in their grammar but choose the common form as the default. Similarly, Weerman and colleagues (2011) found little improvement in grammatical gender assignment comparing a younger (6-8) and an older group of children (12;3 to 13;3) with language impairments. They also found persistent overgeneralization of common determiners and overt adjectival inflections with neuter nouns. Children with SLI do not only make more substitution errors in assigning grammatical gender, but also more often omit articles (van Ewijk & Avrutin, 2010), which might be evidence for avoiding a gender choice. In error analyses, it is therefore important to take different types of mistakes into account.

Within this study, every grammatical variable was tested by using both a comprehension and a production task, as was described in §2.3. Within the production task, factors were manipulated that could possibly influence performance when processing demands are increased (see §2.3 for a discussion of how processing load is hypothesized to increase with these factors). For grammatical gender, these factors were noun frequency, the distance between dependent elements and information complexity. We varied the distance between determiner and noun (either adjacent, or with one or two elements in between) and the number of adjectives that had to be inflected. Frequency effects of nouns on gender assignment have been previously found in L2 learners of Dutch (e.g., Unsworth, 2008; Blom & Vasić, 2011). Neuter gender forms are used more often correctly with frequent nouns than with infrequent neuter nouns. The influence of the distance between dependent elements has not been investigated before for Dutch grammatical gender, but has been shown to influence performance in other grammatical aspects like subject verb agreement (e.g., Kaan, 2002).

Based on the literature discussed above, and our general hypotheses presented in §2.4, the expectation is that there will be persistent group differences for the judgement and production of neuter determiner assignment and adjectival inflection in neuter contexts (see Figure 5.2 for a summary of the hypotheses for grammatical gender). Development of neuter determiner assignment in SLI is expected on the basis of the fact that determiners can be produced correctly by storing determiner and noun together in the lexicon, as was discussed in §5.1. The older SLI group is therefore expected to outperform the younger SLI group because they are assumed to have a larger lexicon and deeper lexical knowledge. For determiner assignment, an interaction between the group and age effects is therefore hypothesized to be present (with the difference between SLI and TD becoming smaller in older groups). However, the difference between SLI and TD is not hypothesized to disappear entirely. Whether children with SLI will ultimately discover a grammatical gender feature and learn the rule for adjectival inflection is hard to predict. Previous studies seem to indicate that this feature is not discovered before age 12 in SLI (Keij et al., 2012), and the rule for adjectival inflection has also been shown to be difficult up to that age (Weerman et al., 2012). It is, however, not clear what happens after 12 years of age.

HYPOTHESES GRAMMATICAL GENDER

* Persistence grammatical problems
determiner assignment & adjectival inflection (neuter) (§5.4.1 - §5.4.3)
judgement
production } effect Group & Age (SLI<TD) – effect Group still present in adolescence

* Variability in performance (§5.4.4) & (§5.4.5)
determiner assignment & adjectival inflection (neuter)
linguistic context → effect size interacts with Group (larger in SLI than in TD)

* Relation variability and processing abilities (§5.4.6)
correlations between linguistic context effect and processing measures

Figure 5.2. Schematic illustration of the hypotheses for grammatical gender

As was already postulated in §2.4, variability in performance is hypothesized, dependent on the linguistic context in which grammatical gender is tested. Although all groups will show variability to some extent, effect sizes are hypothesized to be larger in the SLI groups compared to the TD groups. Whether this is due to (impaired) processing abilities is tested by computing correlations. The final hypothesis is that correlations will exist between information processing measures and task/context effects for grammatical gender. In the next section, the tasks used for testing grammatical gender will be described.

5.3 Task descriptions

5.3.1 Item selection

Ten common and ten neuter nouns were selected and matched on the number of syllables and on animacy (two animate nouns for every gender). The nouns were selected on the basis of relative frequency in the vocabulary of 4- to 6-year-olds, as rated by teachers (de Glopper, Damhuis, de Boers, & Kienstra, 1992). Half of the nouns had a high frequency and half had a low frequency. High frequency was defined as ‘occurring between nr. 1 and nr. 200 on the frequency word list; low frequency was defined as ‘occurring after nr. 1500 on the word list’. Table 5.3 shows the selected nouns and their frequency rankings in the word list. As can be seen, the average frequency ranking of the common and neuter words is similar.

Eight different adjectives were selected for the elicitation of adjectival inflection in indefinite contexts. The selected adjectives were used to the same extent with the nouns of the two genders to avoid effects of frequency of the adjectives (de Glopper et al., 1992) or their phonological characteristics (semivowel, plosive, nasal, liquidate endings) (see Table 5.4). Phonological characteristics were controlled for since in other inflectional domains like subject verb agreement, the phonological properties of the coda appeared to have an effect on the production of suffixes (e.g., Marshall & van der Lely, 2007; Song et al., 2009).

Table 5.3. Noun selection grammatical gender

Common high frequent		Freq.	Neuter high frequent		Freq.
1	*poes* 'cat'	19	11	*paard* 'horse'	89
2	*boom* 'tree'	93	12	*huis* 'house'	16
3	*fiets* 'bicycle'	95	13	*raam* 'window'	100
4	*ballon* 'balloon'	58	14	*potlood* 'pencil'	90
5	*schommel* 'swing'	169	15	*cadeau* 'present'	129
Average word list nr.		87	Average word list nr.		85

Common low frequent		Freq.	Neuter low frequent		Freq.
6	*draak* 'dragon'	1693	16	*hert* 'deer'	1607
7	*bijl* 'axe'	1847	17	*net* 'fishing net'	1770
8	*helm* 'helmet'	1930	18	*zwaard* 'sword'	1727
9	*raket* 'rocket'	1823	19	*eiland* 'island'	2362
10	*gitaar* 'guitar'	1562	20	*zadel* 'saddle'	1878
Average word list nr.		1771	Average word list nr.		1869

Table 5.4. Adjective selection grammatical gender task

	Freq.	Final consonant	With common	With neuter
blauw/blauwe 'blue'	856	semivowel	4	4
rood/rode 'red'	668	plosive	3	3
groot/grote 'big'	215	plosive	3	3
klein/kleine 'small'	190	Nasal	3	3
bruin/bruine 'brown'	1475	Nasal	3	3
groen/groene 'green'	1165	Nasal	3	3
wit/witte 'white'	929	Plosive	3	3
geel/gele 'yellow'	930	Liquidate	2	2

5.3.2 Judgement task

A judgement task was constructed to test the knowledge of gender in a relatively passive way as was described in §2.3. In this task, participants had to judge whether the right gender had been assigned to a noun by either accepting or rejecting noun phrases. The noun phrases consisted of combinations of a correct or erroneous definite determiner and a noun (for determiner assignment) or an indefinite determiner with a correct or erroneous inflected adjective and a noun (for adjectival inflection). Every noun was presented once with a correct and once with an incorrect determiner (e.g., *het/*de paard* 'the.N/*the.C horse'). Similarly, each noun was presented once with a correctly inflected and once with an incorrectly inflected adjective (e.g., *een groot/*grote paard* 'a big-N/*big-C horse'*)*. In total, participants were presented with 80 gender items *(*see Appendix A for the total list of items). The items for grammatical gender were mixed with the items for subject verb agreement (see §6.3.2 for a description of the subject verb agreement task). These items functioned as fillers for the gender items (and vice versa). In total, participants had to judge 152 items.

The judgement task was presented by telling participants a short story about a little girl, Layla, who comes from another country and tries to learn Dutch but still makes some mistakes. The following text is a translation of the introduction:

> *'This is Layla. Layla comes from abroad, and she is trying to learn Dutch. But sometimes she still makes mistakes. In a moment, you will hear sentences that Layla spoke and you have to say whether they sound right or wrong. If a sentence sounds right, you press the green button. If it sounds wrong, you press the red button. You are thus pretending to be teacher. Don't pay attention to the pictures, those are not important. It is only important to listen whether it sounds good or not. We will start with some practice items'.*

Layla's sentences were aurally presented via speakers and were accompanied by a semantically matching picture on the screen to facilitate processing. Because the test was administered in E-prime on a touchscreen laptop, answers were automatically recorded. Participants only had to press the green or red buttons on the touchscreen to respond (see Figure 5.3. for an illustration of the screen). The test started with a practice session of 11 items to familiarize the participants with the task.

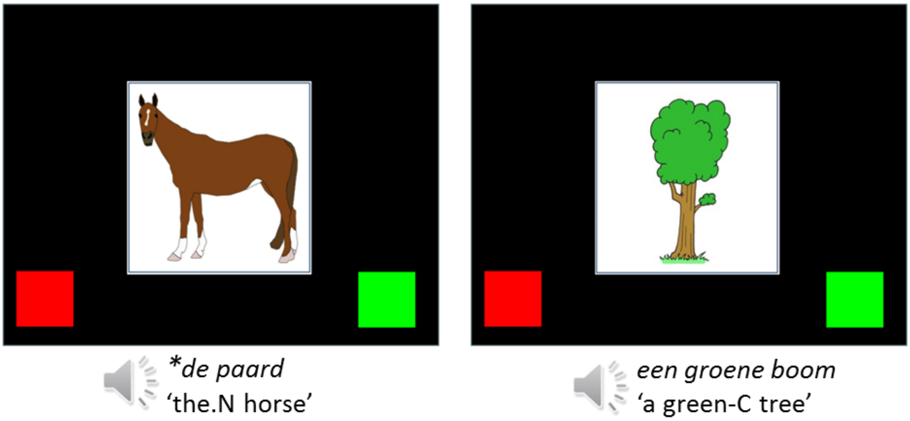


Figure 5.3. Illustration of the judgement task for grammatical gender

The judgement task took around 15 to 20 minutes per subject (including fillers), but the task was split into two in order to avoid effects of decreasing attention or fatigue. The data from the judgement task for grammatical gender was scored automatically (either 0 or 1).

5.3.3 Elicitation task

To test the production of grammatical gender, an elicitation task was constructed on the basis of previous Dutch gender tasks (Blom, Orgassa, & Polišenská, 2008a; Unsworth et al., 2014). Participants had to finish sentences started by the experimenter. Adjustments were made in order to manipulate frequency and context, and per context several items were elicited. For every noun, the definite determiner was elicited four or six times. Determiners were elicited twice in a context where the determiner and noun had to be produced together without any elements in between (DET Adjacent). Similarly, they were elicited twice in a context where an adjective had to be used to specify the noun (DET 1 ADJ). For every frequency category (*high common, low common, high neuter, low neuter*), two words were chosen for which definite determiners were elicited twice in the context of two intervening adjectives (DET 2 ADJ) to test the effect of a more complex context for determiner assignment. The following words were chosen: *boom* 'tree', *ballon* 'balloon', *helm* 'helmet', *raket* 'rocket', *huis* 'house', *potlood* 'pencil', *net* 'fishing net', *zadel* 'saddle' (see Figure 5.5 for examples of each level of complexity). In sum, a total of 96 items for the elicitation of definite determiners were constructed.

Table 5.5. Examples of elicitations of determiner assignment in different contexts

	Picture	Experimenter	Target
DET adjacent		*Kijk, een vogel en een huis!* 'Look, a bird and a house!' *De vogel vliegt over…* 'The bird is flying over…'	*het huis* 'the house'
DET 1 ADJ		*Kijk, twee huizen!* 'look, two houses' *De bloem staat voor…* 'the flower is in front of…	*het rode huis* 'the red house'
DET 2 ADJ		*Kijk, vier huizen!* 'look, four houses' *De hand pakt…* 'the hand is picking…'	*het kleine, rode huis* 'the small, red house'

The items for adjectival inflection in indefinite contexts were constructed in a similar way: for all nouns, an adjective was elicited twice in an indefinite context where only one adjective had to be produced to denote the noun (1 ADJ). For the same selection of 2 words per category, adjectival inflection was elicited twice in an indefinite context where two adjectives had to be produced to denote the noun (2 ADJ) (see Table 5.6 for examples of the two contexts). For adjectival inflection, a total number of 56 items were administered per child (a full list of elicitation cues and target answers for both determiner assignment and adjectival inflection can be found in Appendix B). As was the case in the judgement task, the items for the elicitation of grammatical gender (determiners and adjectives) were mixed with the items for the elicitation of subject verb agreement (see §6.3). The gender and the subject verb agreement items functioned as fillers for each other. In total, the elicitation task for gender and subject verb agreement took around 30 minutes. The task was, however, administered in two parts of 15 minutes, to avoid effects of fatigue or loss of attention. The sentences were elicited with the help of pictures, presented in Powerpoint. Subjects were told that pictures would appear on the screen and that the experimenter would start a sentence which they had to complete. For each target type, practice items were included to familiarize the subject with the procedure of the task. In Table 5.6, the elicitation contexts for the different target types are illustrated with an example of the picture used, the elicitation cue produced by the experimenter, and the target utterance.

Table 5.6. Examples of elicitation of adjectival inflection in different contexts

	Picture	Experimenter	Target
1 ADJ		*Kijk, twee huizen!* 'look, two houses' *Dit is een….* 'this is a….'	*blauw huis* 'small-N house'
2 ADJ		*Kijk, vier huizen!* 'look, four houses' *Dit is een….* 'this is a….'	*groot, rood huis* 'big-N, red-N, house'

In coding the scores for the accuracy of determiner assignment, productions of definite determiners or adjectives with a noun other than the target noun or productions of indefinite instead of definite determiners were regarded as missing data and eliminated from further analyses since they could not strictly be scored as incorrect. Accuracy analyses included correct determiners (score 1) and incorrect determiners (score 0). Determiner omissions were separately coded, since they might be an indication of problems with gender or problems with the obligatory nature of determiners in general (Eyer & Leonard, 1995; van Ewijk & Avrutin, 2010). In the analysis of the adjectival inflection, productions of adjectival inflection in a definite context or with a different noun than the target noun were coded as missing data and excluded from further analyses. In the case of omission of the adjective or production of an additional adjective, the item was coded separately and also excluded from the accuracy analysis. Accuracy analyses therefore included correctly inflected adjectives (score 1) and incorrectly inflected adjectives (score 0).

5.4 Results

In this section, the results of grammatical gender will be discussed, starting with the results of the determiner assignment (§5.4.1), followed by the results of adjectival inflection (§5.4.2). Both sections first discuss the results of the judgement task, and then focus on the results of the production task, in order to answer the question whether grammatical problems in this domain are persistent. In order to find out whether a gender feature has been acquired, the results of determiner assignment and adjectival inflection will also be discussed in combination (§5.4.3). In the sections that follow variability in performance will also be addressed. The results of the judgement and the production task are combined require to look at task effects

(§5.4.4) and the effects of the linguistic context are evaluated (§5.4.5). In the final section (§5.4.6), the correlation between the outcomes on grammatical gender and the processing measures will be discussed to answer the question whether fluctuations in performance (task and context effects) can be explained by (limited) processing capacities.

5.4.1 Results determiner assignment

Judgement of determiners

The results of the judgement of determiner assignment are shown in Figure 5.4. For each group, the range of accuracy percentages is represented by the boxes and the whiskers (the bold line in the middle of each box represents the median). Common and neuter determiners are represented separately in striped dark and solid light bars respectively.

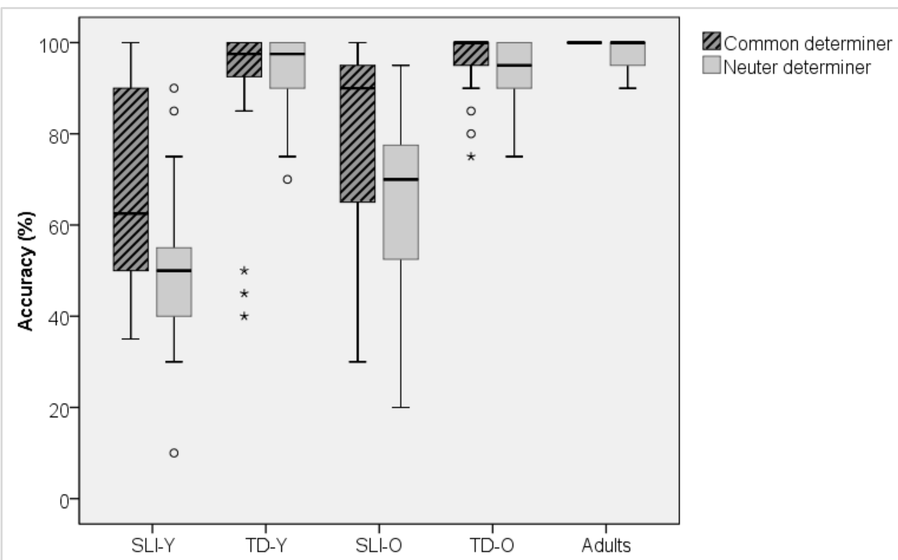


Figure 5.4. Accuracy (%) in the judgement of determiner assignment (common and neuter) per group

The graph shows, first of all, that the judgement task was an easy task for adults, since their performance was at ceiling (100% correct judgement of common determiners and 97% correct judgement of neuter determiners).[5] There were thus no unexpected or undesired task effects.

As the graph shows, the SLI groups had a large variance in their accuracy scores, but their scores did not overlap to any great extent with those of the TD groups. Although some children and adolescents with SLI performed much better than others, they almost all performed worse than their TD peers. For the percentage of correctly judged determiners, ANCOVAs with IQ as a covariate revealed main effects of Group for both common nouns ($F(1,120) = 32.67$, $p < .001$, $\eta_p^2 = .214$) and neuter nouns ($F(1,120) = 184.18$, $p < .001$, $\eta_p^2 = .605$). Main effects for Age were also found ($F(1,120) = 8.64$, $p = .004$, $\eta_p^2 = .067$ for common gender and $F(1,120) = 14.71$, $p < .001$, $\eta_p^2 = .109$ for neuter gender). An interaction between Group and Age was absent in common gender items, but present in neuter gender items ($F(1,120) = 8.22$, $p = .005$, $\eta_p^2 = .064$). Post-hoc analyses revealed that the difference between SLI and TD in the judgement of neuter gender items was still present in the older group ($F(1,58) = 54.09$, $p < .001$, $\eta_p^2 = .483$).[6]

The statistical analyses show that differences between SLI and TD are clearly present in the judgement of grammatical gender on determiners. Differences are found for both common and neuter gender, although the difference is much larger in the neuter domain, as expected. Difficulties in the judgement of neuter determiners persist, although they reduce in subjects with SLI with age, as was hypothesized.

The low scores on judgement of neuter determiners in the younger SLI group (an average accuracy of 51%) may seem to indicate that performance was at chance level in this group. This idea is requires some nuance if we look at the results of the judgement task in more detail. Table 5.7 shows the overall results of the judgement task for common and neuter nouns, now further divided into correct acceptance of correct forms and correct rejection of erroneous forms of the determiners. Both SLI groups mainly had problems in rejecting incorrect forms of the determiners, both in common and in neuter contexts. They thus accepted both correct and incorrect forms. This can be interpreted as an indication of their representations being unstable or as a reflection of insecurity about making linguistic judgements (possibly due to awareness of their language impairment). Whatever the explanation,

[5] The errors that constituted the 3% error rate in the judgement of neuter determiners in adults did not show any consistencies. Errors were made in different words and by different adults. They therefore seem to reflect incidental errors, probably due to flaws in attention that mainly had an effect on the non-default context (neuter).

[6] As was mentioned in §3.4, the analyses were repeated with a smaller sample in which the participants with very few rejections (<10) were eliminated. Five subjects in the SLI-Y group appeared to meet this criterion. Outcomes remained the same when these subjects were excluded.

the detailed results show that the low accuracy rates are not due to chance performance (i.e. just pressing a button without listening) but to genuine difficulties in rejecting erroneous forms.

Table 5.7. Accuracy (%) in the judgement of determiner assignment per group, divided over correct acceptance and correct rejection

	N	Common N = 20	Accept	Reject	Neuter N= 20	Accept	Reject
SLI-Y	32	69.69	92.50	46.88	51.09	79.06	23.13
TD-Y	32	91.41	93.44	89.38	93.75	95.63	91.88
SLI-O	31	79.03	86.45	71.61	65.65	83.55	47.74
TD-O	30	96.67	97.67	95.67	94.17	95.33	93.00
Adults	14	100	100	100	97.79	97.14	96.43

A possible adjustment of judgement scores, to correct for a bias towards accepting sentences would be to compute A' or D'-scores. According to Rispens & Been (2007), A' and D'-scores are based on a formula that takes into consideration the number of hits (correct acceptance of correct items) and the number of false alarms (the number of rejections of correct items). However, since the bias towards accepting sentences rather than rejecting them was only present in the SLI groups and not so much in the TD groups, a transformation was not considered necessary.

Production of determiners

Figure 5.5 presents the results of the production of determiners. Again, boxes and whiskers represent the range of accuracy percentage scores (the horizontal bar in the middle of each box represents the median) and for each group the production data of common and neuter determiners are presented separately. The graph shows that none of the groups had difficulties with the production of common determiners with common nouns (the few errors that were made will be discussed in more detail below). Because common gender is the default in Dutch, it is not surprising that this form was produced without any problems. In the production of neuter gender, on the other hand, large differences appeared: both the younger and the older SLI group performed much worse than their TD peers.

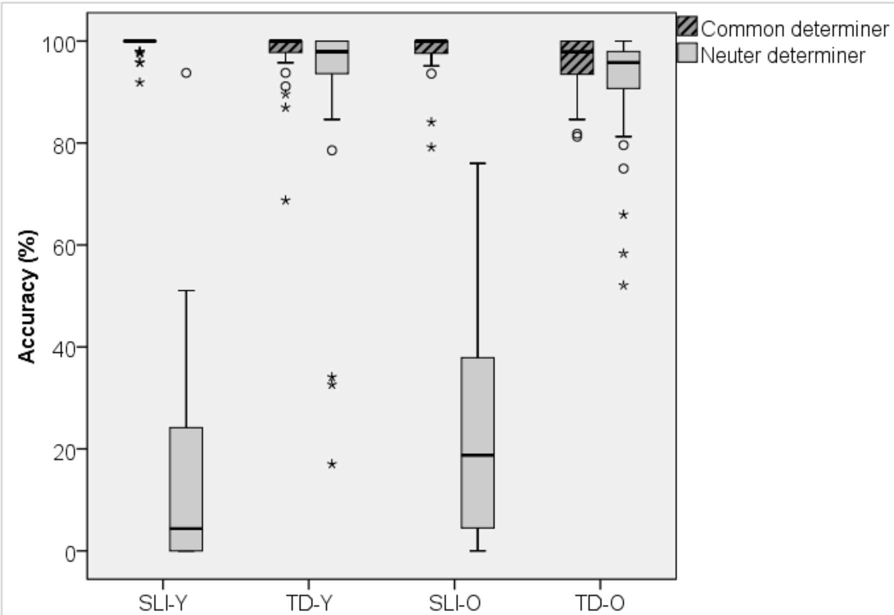


Figure 5.5. Accuracy (%) in the production of determiner assignment (common and neuter) per group

There was considerable variance in the SLI groups in the production of neuter determiners, but overall the scores of the SLI groups did not overlap to any great extent with those of the TD groups. Some children with SLI thus produced more correct neuter determiners than others, but none of them reached the performance level of the TD children. When the performance of the SLI and TD groups was compared statistically with ANCOVAs with IQ as a covariate, a small main effect for Group was found in the production of common determiners ($F(1,120) = 5.41$, $p = .02$, $\eta_p^2 = .043$), while large effects appeared in the comparison of neuter determiner production ($F(1,120) = 352.29$, $p < .001$, $\eta_p^2 = .746$). Age did not have a significant effect, and interactions between Group and Age were absent. Children with SLI thus performed significantly below TD children in the production of neuter determiners, and their performance did not improve significantly with age (in contrast to our hypotheses). Problems in the production of grammatical gender in determiners thus clearly persist into adolescence.

A detailed look at the data of the production task provides interesting additional information. Table 5.8 shows the total number of elicited determiners, the number of missing data points (these were often productions with a different noun), the number of indefinite determiners that were produced (Indefinite), the number of determiner omissions (Omission) and the number of determiners that were counted as analysable in the accuracy analysis (Analysable). The analysable productions are further divided into the percentage of correct and incorrect determiners (i.e. using the opposite gender).

It is particularly interesting to look at the omission rates in the different groups. Table 5.8 shows that both SLI groups produced more omissions of the definite determiner than the TD groups, both in common and in neuter contexts. Furthermore, the SLI groups less often produced an incorrect determiner in common contexts. This indicates that children with SLI relied more on the default definite determiner than children with a typical development. It furthermore shows that children with a typical development were more aware of the fact that a determiner is obligatory and cannot be omitted.

Table 5.8. Detailed results common and neuter determiner assignment per group

		Total	Missing	Indefinite	Omission	Analysable	Correct	Incorrect
	Group	N	N	N	N	N	N	N
Common	SLI-Y (N=32)	1536	40	0	93	1403	1392	11
	TD-Y (N=32)	1536	51	13	25	1447	1411	36
	SLI-O (N=31)	1489	6	28	100	1355	1320	35
	TD-O (N=30)	1392	6	14	21	1399	1338	61
Neuter	SLI-Y (N=32)	1536	31	7	92	1406	201	1205
	TD-Y (N=32)	1536	31	21	15	1469	1331	138
	SLI-O (N=31)	1488	30	19	88	1380	338	1042
	TD-O (N=30)	1392	1	2	15	1426	1294	132

In sum, the results on the judgement and production data of determiners show that differences between SLI and TD are present and persist into adolescence. Especially in neuter determiner assignment, large differences were found. In the judgement task, the older SLI group performed significantly better than the younger SLI group and the difference with age-matched peers decreased, as expected. Such an 'improvement' was not present in production. We will come back to this difference between tasks in §5.4.4. First, we will describe the outcomes for grammatical gender as expressed in adjectival inflection.

5.4.2 Results adjectival inflection

Judgement of adjectives

Figure 5.6 shows the results of the judgement of adjectival inflection in an indefinite context. Again, boxes represent the range of accuracy scores in percentages (with the horizontal bar in the middle of each box representing the median), and for each group the accuracy for adjectival inflection with common and neuter nouns is represented separately in blue and green respectively. As was the case in determiner assignment, adults did not have any problems performing this task. Again, the scores of the SLI groups showed considerable variance, but - especially with the neuter nouns - there was little overlap between the scores in the SLI groups and the scores in the TD groups.

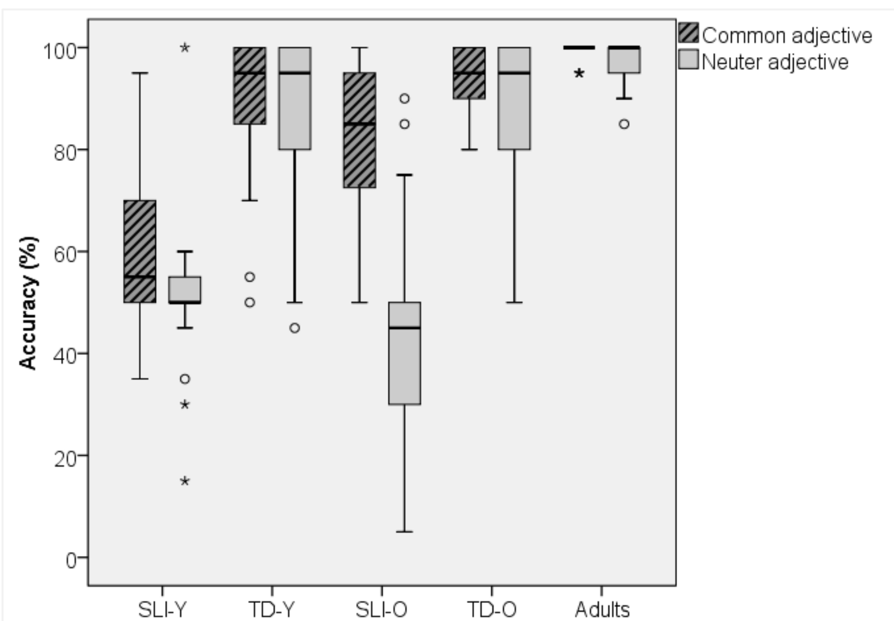


Figure 5.6. Accuracy (%) in the judgement of adjectival inflection in indefinite contexts (common and neuter) per group

In a comparison of the performance of the SLI groups to the performance of the TD groups (again, with ANCOVAs with IQ as a covariate), a significant main effect of Group on the judgement of adjectives appeared, both with common nouns ($F(1,120) = 65.43$, $p < .001$, $\eta_p^2 = .353$) and with neuter nouns ($F(1,120) = 167.11$, $p < .001$, $\eta_p^2 = .582$). Effect sizes were larger for adjectives in neuter contexts, as expected.[7] The factor Age showed a main effect in the judgement of common adjectives only ($F(1,120) = 28.99$, $p < .001$, $\eta_p^2 = .195$). Post-hoc analyses showed that it were mainly the SLI-subjects that caused this significant effect of Age (for SLI: $F(1,60) = 26.41$, $p < .001$, $\eta_p^2 = .306$ and for TD: $F(1,59) = 4.78$, $p = .033$, $\eta_p^2 = .437$). As was mentioned in §3.4, the analyses were checked with a smaller sample in which the participants with very few rejections (<10) were eliminated. Five subjects in the SLI-Y group appeared to meet this criterion. When these subjects were deleted from the analyses, the significant effect of age in the SLI-group also disappeared ($p = .065$).

As was the case with the judgement of determiners, separation of correct acceptance and correct rejection showed that the SLI groups mainly have difficulties in rejecting erroneous forms, while the TD groups performed very similarly in accepting and rejecting items. Adjustment of scores was therefore not performed, because a bias towards accepting sentences was not present in the whole group of participants.

Differences between SLI and TD were thus also clearly present in the judgement of grammatical gender in adjectival inflection. In comparison to the TD groups, both SLI groups showed differences in deciding whether adjectives are correctly inflected. This was true for both common and neuter nouns, but - as expected - effect sizes were larger with adjectives that modify neuter nouns. While the older SLI group performed significantly better than the younger SLI group in the judgement of adjectives in common contexts, their scores in neuter contexts were not significantly different (the median was in fact lower, see Figure 5.6). It thus seems as if the older SLI group relied even more on the default form when it comes to adjectival inflection, while the judgement of grammatical gender in determiners suggested improvement over time in SLI in the neuter domain. We will come back to this point in the following section, when we combine the results of determiner assignment and adjectival inflection (§5.4.3).

[7] Repetition of the analyses with a smaller sample in which the participants with very few rejections (<10) were eliminated (N=5) gave similar outcomes.

Production of adjectives

In Figure 5.7 the results of the production of adjectival inflection are shown. The boxes represent the range of percentage accuracy scores for the different groups. The graph shows that none of the groups had difficulty with the production of adjectives with common nouns (although there were some subjects who had accuracy rates of around 80%). In the production of adjectives in neuter contexts, there were again large differences between groups. The high number of outliers requires some explanation. This might reflect the dichotomous nature of the rule for adjectival inflection: it is either acquired or not. The outliers in the SLI group were those few individuals with SLI who seemed to have acquired a gender feature (see §5.4.3 for a more detailed discussion). Apart from the overlap in scores between the outliers, the SLI groups and the TD groups did not overlap. The SLI groups performed much worse than their TD peers on the production of the adjective in neuter contexts.

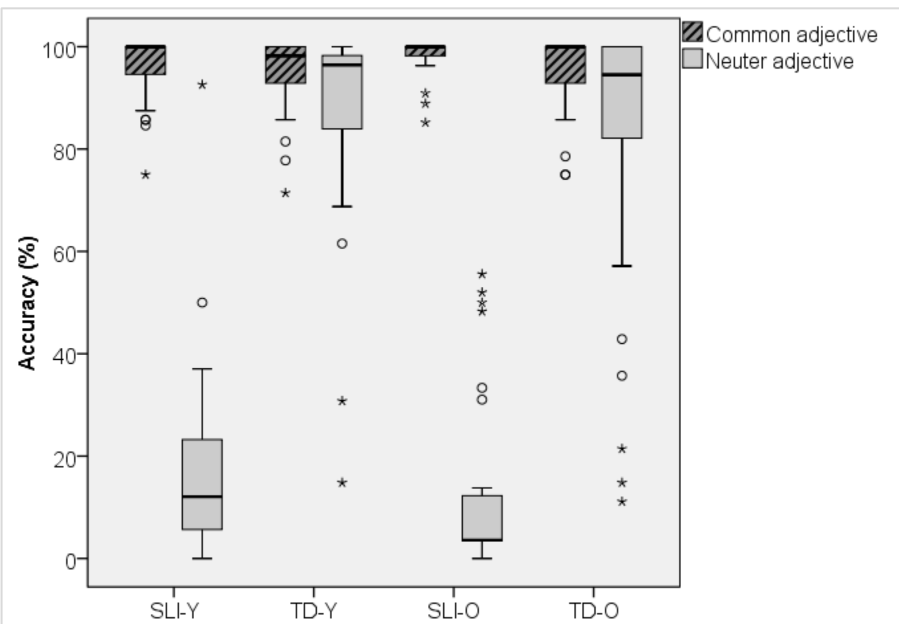


Figure 5.7. Accuracy (%) in the production of adjectival inflection in indefinite contexts (common and neuter) per group

When the performance of the SLI and TD groups was compared statistically with the help of ANCOVAs with IQ as a covariate, no significant effect of Group was found in the production of adjectives with common nouns, while a large main effect of Group appeared in neuter nouns ($F(1,120) = 299.51$, $p < .001$, $\eta_p^2 = .714$). Age was not found to have a significant effect on the scores, and interactions between Age and Group were absent. As was the case in the judgement of adjectival inflection, the older SLI group performed even slightly worse than the younger SLI group, although not significantly. Also in the production of adjectives, subjects with SLI thus seemed to rely heavily on the default form.

In sum, the results on the judgement and production data of adjectives showed that differences between SLI and TD are present and persist into adolescence. Especially in the neuter domain, large (persistent) differences were found, as was the case for grammatical gender expressed in the definite determiner. Whereas the judgement and production of grammatical gender in definite determiners seemed to improve over time in SLI (although not significantly in the production task), the judgement and production of grammatical gender in adjectives seemed to decline. Regarding grammatical gender in adjectives, individuals with SLI thus seemed to rely more heavily on the default over time. The dichotomy in results between determiners and adjectives raises the question what the 'improvement' in determiner assignment actually reflects: do children with SLI have (some) stable neuter representations or do neuter determiners randomly occur due to lexical storage? This question will be addressed in the following section.

5.4.3 Combining the results from determiners and adjectives

From the previous results where determiner assignment and adjectival inflection were considered separately, we cannot conclude whether an abstract gender feature has been assigned. In theory, adjectival inflection can only be correctly applied if gender has been stored correctly (and the rule for adjectival inflection has been learned). In order to determine this, the results of the determiner assignment task and the adjectival inflection task need to be combined. The best indication of the gender children have stored for a noun in their lexicon is their choice of the definite determiner. The neuter nouns for which the correct definite determiner *het* was used consistently (that is in ≥ 3 out of 4 test items for that noun) were counted. These nouns will be referred to here as *stable neuter* nouns. Table 5.9 indicates the number of neuter nouns elicited per group (10 per participant), the number of stable neuter nouns in those groups, and the proportion of stable neuter nouns for which the adjective was correctly inflected in an indefinite context.

Table 5.9. Elicited neuter nouns, stable neuter nouns and the percentage of correct adjectival inflection per group

	Neuter nouns elicited	Stable neuter nouns (consistent neuter det.)		Stable neuter nouns with correct adjectival inflection
	N	N (% of elicited nouns)		% (of stable nouns)
SLI-Y	320	33	(10)	54
TD-Y	320	287	(90)	93
SLI-O	310	50	(17)	42
TD-O	300	273	(91)	84

As can be seen from Table 5.9, the number of stable neuter nouns was very low in the two SLI groups (10% in the younger group and 17% in the older group) while the number of stable neuter nouns was high in the TD groups (both around 90% of the number elicited nouns). This is not surprising in the light of the large difference between the SLI and the TD groups evident in the production of neuter determiners.

Theoretically, it is possible that SLI subjects have an abstract gender feature specified on nouns (and have discovered the rule for adjectival inflection), but only have a small subset of nouns that are assigned *neuter* gender (or are classified as a *het*-word). To examine this possibility, the proportion of stable neuter nouns for which the adjective was correctly inflected in an indefinite context was calculated (see again Table 5.9). In the TD groups, the percentage of stable neuter nouns for which the adjective was correctly inflected in a singular, indefinite context was quite high (93% for the younger TD group and 84% for the older TD group). If the neuter definite determiner was used consistently, TD children also tended to correctly inflect the adjective. This indicates that in both TD groups a gender feature is present and the rule for adjectival inflection has been learned by the vast majority. In the SLI groups, on the other hand, the proportion of correct adjectival inflection was quite low (around or just under half). Their low performance on the gender tasks was thus certainly not only due to having a small subset of nouns that are assigned neuter gender. On top of difficulties in storing nouns as neuter gender or as *het*-words, subjects with SLI thus seem to have problems with the rule for adjectival inflection.

Table 5.9 does not indicate individual variation in the number of stable neuter nouns nor in the correct application of the rule for adjectival inflection. In Table 5.10, this information is provided. The number of children that produced at least 1 or more stable neuter nouns is indicated, and the number of children that produced correct adjectival inflection with a stable neuter noun is shown. As can be seen, almost all the TD children and adolescents had at least 1 stable noun and also applied the adjectival rule correctly with these nouns (as was found before by Blom et al., 2008b, see Figure 5.1). In the SLI groups, stable neuter nouns were however produced by only 10 younger children, which is less than a third, and 17 adolescents (just over half). Even fewer children also seemed to have discovered the rule for adjectival inflection – less than a sixth in the younger SLI group and less than a quarter in the older SLI group. Some subjects with SLI thus discovered the gender feature and applied the rule for adjectival inflection correctly to nouns they stored as neuter (or as *het*-word), but the majority of subjects with SLI did not seem to acquire a gender feature, or seemed to struggle with the rule for adjectival inflection.[8]

Table 5.10. Number of children per group that produced stable determiners and correct adjectival inflection with stable neuter nouns

	Stable neuter nouns (1 or more word)	Adjective correctly inflected with stable neuter noun
	N (children)	N (children)
SLI-Y (N=32)	10	5
TD-Y (N=32)	31	30
SLI-O (N=31)	17	7
TD-O (N=30)	30	30

In sum, the number of stable neuter nouns was low in the SLI groups compared to the much higher number in the TD groups. With stable neuter nouns, TD subjects often produced correct adjectival inflection – indicating that the rule for adjectival inflection has been learned and an abstract gender feature was present. In the SLI group, adjectives were, however, inflected correctly with only half of the stable neuter nouns. Furthermore, the few stable neuter nouns were produced by only a small minority of subjects with SLI, while almost all TD subjects produced at least one stable neuter noun. When the percentage of correct adjectival inflection with a stable neuter noun is considered, this number is even lower, around a half. Only a few individuals with SLI thus seemed to have acquired the feature that expresses grammatical gender (or to have stored nouns as *het*-nouns in the lexicon) and even

[8] The few children who discovered the rule also represented the outliers in Figure 5.7.

fewer of them had acquired the rule for adjectival inflection. The majority of children with SLI had not discovered a gender feature nor acquired the rule for adjectival inflection. The problem with grammatical gender is thus multi-faceted.

Curiously, the analysis of the stability of neuter determiner assignment revealed that some children showed a reversed pattern. One child in the younger SLI group (SLI-Y.10) and three adolescents in the older SLI group (SLI-O.103, SKLI-O.110, SLI-O.126) produced the correct form of the adjective for (some) neuter nouns, but did not produce the correct form of the definite determiner (or at least not consistently). Such remarkable behaviour was noted before by Weerman and colleagues (2011), who described one child in their older SLI group that showed similar patterns. They ascribed this phenomenon to problems in the spell-out of grammatical knowledge. In order to produce the correct form of the adjective, a neuter gender feature has to be present. In the production of the definite determiner, this feature is – however – not always correctly spelled out. Because these children do produce neuter determiners with some (other) neuter nouns, and also correctly inflect the adjective with a schwa in common contexts, Weerman and colleagues suggested that the spell-out rules do not seem to be missing. Their explanation for the reversed stability is the failure to implement rules. In the next sections, we will discuss other arguments for the idea that knowledge is not always implemented in performance in SLI.

5.4.4 Effect of task

While the difference between SLI and TD decreased with age for the judgement of neuter definite determiners, such an 'improvement in SLI' was not found in the production data. Our hypothesis that the older SLI group would perform better than the younger SLI group on the production of neuter determiner assignment, because they could make more use of the storage of determiner and noun in the lexicon, was only borne out in the judgement task, not in production. Such an effect of task seemed to be absent in the TD groups, as Figure 5.8 illustrates.

To test whether the effect of task was significantly different in SLI and TD groups, repeated measures were computed with IQ as a covariate and the judgement and production scores for neuter determiners as dependent variables. The type of task influenced the scores to a high extent, and the interaction between Task and Group was highly significant ($F(1,120) = 107.63$, $p < .001$, $\eta_p^2 = .473$). This indicates that performance in the SLI group was affected differently by the type of task than performance in the TD group (with production being more difficult than judgement). Caution regarding these conclusion is, however, needed because the TD groups performed at ceiling.

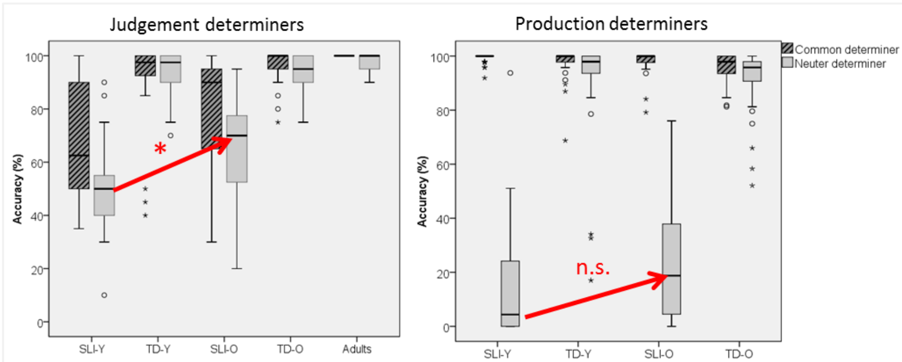


Figure 5.8. Comparison of neuter determiner assignment between SLI-Y and the SLI-O group in perception versus production

Although we cannot be entirely sure that similar task effects would not also have shown up in the TD groups if they had not performed at ceiling, the ‘improvement’ of determiner assignment between younger and older subjects with SLI seems to indicate that some knowledge is gained over time, although this is not reflected in performance on the production task. In §5.4.6, we will investigate whether the difference between judgement and production is correlated to processing abilities.

5.4.5 Effects of linguistic context

The production task for determiner assignment was constructed such that two factors could be measured as possible influences on performance: the frequency of nouns (high or low) and the distance between the determiner and the noun. The hypothesis was that these factors would affect performance in all groups, but would have a larger effect in the SLI groups than in the TD groups. Table 5.11 shows the results of the accuracy percentages in the production of common and neuter determiners, divided into the mean accuracy rate for highly frequent nouns (H) and for nouns with a low frequency (L).

As Table 5.11 shows, determiners were produced correctly more often with high frequency nouns than with low frequency nouns in all groups; this was true for both for common and neuter gender. Repeated measures (with two levels for high and low frequency) revealed a significant effect of Frequency on the production of the neuter determiner within the same items ($F(1,120) = 36.79$, $p < .001$, $\eta_p^2 = .235$). There was a significant interaction with Age ($F(1,120) = 14.66$, $p < .001$, $\eta_p^2 = .109$) but not with Group ($p = .107$) (despite the fact that the effect of frequency seems larger in the older SLI group than in the older TD group).

Table 5.11. Accuracy (%) in the production of determiners (common and neuter) for high and low frequency items across the four groups

	N	Common % (SD)	H	L	Neuter % (SD)	H	L
SLI-Y	32	99.21 (1.80)	99.28	99.16	14.37 (21.37)	15.21	13.60
TD-Y	32	97.28 (6.20)	99.13	95.34	90.50 (21.15)	92.13	89.02
SLI-O	31	97.47 (4.64)	99.30	95.71	24.21 (22.07)	31.96	14.93
TD-O	30	95.56 (5.37)	99.29	91.78	90.75 (12.61)	93.84	85.98

Noun frequency thus influenced performance on the production of neuter definite determiners, but the effect was not significantly different in SLI compared to TD. The hypothesis that this factor would cause variability in SLI was not supported.

The other complexity factor that was manipulated in the production task for determiners was the distance between determiner and noun. Determiners were elicited in a context where determiner and noun were positioned close together (Adjacent) and in contexts where they were separated by one or two adjectives (1 ADJ or 2 ADJ). The Adjacent-context was hypothesized to be the most simple, while the contexts with one or two intervening adjectives were considered more complex. Table 5.12 shows the scores for common and neuter determiner production again, now divided over the three contexts.

For neuter gender, we found a clear difference in the production of the neuter determiner as a result of the distance between determiner and noun in the SLI groups, but not in the TD groups. As was already shown, the SLI groups performed poorly in general in their production of neuter determiners, but their scores dropped substantially if the context was more complex. Repeated measures (with three complexity levels for 0, 1 or 2 adjectives in between determiner and noun) revealed a significant effect of Complexity on the production of the neuter determiner within the same items ($F(1,118) = 28.20$, $p < .001$, $\eta_p^2 = .193$). There was a significant interaction with Group ($F(1,118) = 22.06$, $p < .001$, $\eta_p^2 = .157$), but not with Age ($p = .318$). Post-hoc analyses revealed that it was the SLI group where the complexity factor influenced the results ($F(1,60) = 46.09$, $p < .001$, $\eta_p^2 = .434$). The TD-groups did not show any effect ($p = .748$). The linguistic context thus caused variability in performance in individuals with SLI but not in TD.

Table 5.12. Accuracy (%) in the production of determiner assignment (common and neuter) for contexts with increasing complexity (Adjacent, 1 ADJ, 2 ADJ) per group

		Common	Adjacent	1 ADJ	2 ADJ	Neuter	Adjacent	1 ADJ	2 ADJ
	N	%	%	%	%	%	%	%	%
SLI-Y	32	99.21	97.47	100	100	14.37	16.94	10.24	3.52
TD-Y	32	97.28	96.24	97.66	93.66	90.50	89.20	87.44	87.48
SLI-O	31	97.47	96.87	96.63	96.79	24.21	29.94	16.31	4.40
TD-O	30	95.56	97.13	95.11	88.18	90.75	85.24	87.86	85.47

For adjectival inflection, similar factors were used as in the production task: noun frequency varied (high or low) and children either had to inflect one or two adjectives to make the context more complex (see §5.3.3 for a detailed description of the factors that were manipulated). Table 5.13 shows the results of the production of adjectival inflections for common and neuter nouns, further divided into the mean accuracy rate (%) for highly frequent nouns (H) and nouns with a low frequency (L).

Table 5.13. Accuracy (%) in the production of adjectival inflection (common and neuter) for high (H) and low frequency (L) items across the four groups

	N	Common % (SD)	H	L	Neuter % (SD)	H	L
SLI-Y	32	95.97 (6.25)	96.05	95.93	16.90 (18.24)	17.54	16.38
TD-Y	32	95.11 (7.18)	96.98	93.59	87.88 (19.62)	88.88	87.52
SLI-O	31	98.29 (3.65)	99.08	97.51	12.52 (17.17)	16.18	8.78
TD-O	30	95.22 (7.47)	98.31	92.16	81.57 (27.69)	86.67	76.39

In all groups, adjectives were inflected more often correctly with high frequency neuter nouns than with low frequency neuter nouns, as can be seen in Table 5.13 and is reflected in the statistical analyses. Repeated Measures showed a main effect of Frequency ($F(1{,}120) = 13.95$, $p < .001$, $\eta_p^2 = .104$). There was a significant interaction with Age ($F(1{,}120) = 7.41$, $p < .007$, $\eta_p^2 = .058$), with older groups showing larger effects of frequency. No interaction between Frequency and Group was found ($p = .740$). Frequency thus affected performance on the production of adjectives in all groups, but was not significantly different in SLI compared to TD.

The other factor that was manipulated in the production task for adjectival inflection was the number of adjectives that had to be inflected. Subjects either had to inflect one or two adjectives, and the latter context was hypothesized to be more complex. Table 5.13 shows the scores for adjectival inflection in common and neuter contexts again, now divided over the two contexts.

Table 5.14. Accuracy (%) in the production of adjectival inflection in indefinite contexts (common and neuter) with increasing complexity (1ADJ, 2ADJ) per group

	N	Common % (SD)	1ADJ	2ADJ	Neuter % (SD)	1ADJ	2ADJ
SLI-Y	32	95.97 (6.25)	96.88	96.74	16.90 (18.24)	19.14	3.57
TD-Y	32	95.11 (7.18)	97.27	90.63	87.88 (19.62)	88.56	81.88
SLI-O	31	98.29 (3.65)	97.93	98.39	12.52 (17.17)	11.25	6.45
TD-O	30	95.22 (7.47)	97.92	91.96	81.57 (27.69)	78.69	82.22

As can be seen in Table 5.13, the number of adjectives that had to be inflected influenced the accuracy of adjectival inflection with neuter nouns in all groups (although the direction of the effect was reversed in the TD-O group). Repeated measures revealed that this complexity factor influenced the results significantly, although the effect was relatively small ($F(1{,}120) = 8.72$, $p = .004$, $\eta_p^2 = .068$). There was a significant interaction with Group ($F(1{,}120) = 4.94$, $p = .028$, $\eta_p^2 = .040$) and the interaction with Age approached significance ($F(1{,}120) = 3.54$, $p = .062$, $\eta_p^2 = .029$). Post-hoc analyses were therefore computed with Repeated Measures for SLI and TD separately (significance level was adjusted to .025 with a Bonferroni-correction). The analyses revealed that the complexity only significantly affected performance in the SLI-groups ($F(1{,}62) = 13.48$, $p < .001$, $\eta_p^2 = .179$). In the TD groups, effects of this linguistic context factor were not significant ($p = .576$).

5.4.6 Correlations between performance effects and processing abilities

To investigate whether the variation in performance in the SLI groups that was reported in the previous sections can be linked to limitations in processing capacity, difference scores were calculated and correlations with processing measures were computed. Difference scores were calculated by extracting the complex context scores from the easy context scores. Correlations were hypothesized for nonword repetition and sentence repetition, since these measures showed persistent limitations in SLI (as was discussed in §4.5). Correlations with other measures of processing were, however, computed. Table 5.15 shows the values of Spearman's rho and the corresponding *p*-values. Significance level was reduced to .01 with a conservative Bonferroni-correction for multiple comparisons (.05/number of tests).

Table 5.15. Correlations between effects of task/linguistic context and processing measures (Spearman's rho)

	Task effect (judge-produce)	Complexity DET (adjacent-2 ADJ)	Complexity ADJ (1ADJ-2ADJ)
Sentence rep.	r = -.61 (p < .001)	r = -.32 (p < .001)	n.s. (p = .111)
Nonword rep.	r = -.57 (p < .001)	r = -.35 (p < .001)	r = -.22 (p =.006)
Digit recall	r = -.33 (p < .001)	r = -.21 (p < .001)	n.s. (p = .463)
Visual recall	n.s. (p = .202)	n.s. (p = .866)	r = -.27 (p = .003)
Inhibition (EF)	n.s. (p = .819)	n.s. (p = .586)	n.s. (p = .871)

The hypothesis that nonword repetition and sentence repetition would correlate with the effects of task and linguistic complexity was confirmed for the effect of task (perception versus production) and the distance between determiner and noun (Adjacent versus 2 adjectives in between). Digit span also appeared to be correlated with variability due to those two factors. The ordering in strength of the correlations followed the proposed hierarchy in the amount of verbal load (see Figure 4.3 in §4.3): more verbally loaded processing tasks showed stronger correlations with grammatical performance. For the complexity factor in the adjectival inflection task, correlations appeared to be significant for nonword repetition but not for sentence repetition (and digit span). Furthermore, a significant correlation with visual span appeared for this factor. In §2.3, we assumed that the number of adjectives that had to be inflected potentially would add to the processing costs, but left open whether this would be due to the fact that more pictures had to be distinguished (visual processing load) or to the fact that children had to distinguish items on two scales, namely size and colour (conceptual load). The absence of a correlation with sentence repetition and the presence of a correlation with visual memory seems to indicate that it was a visual processing load rather than a linguistic processing load that triggered variability in performance. It remains, however, unclear why the same correlation with visual load did not show up in the correlational analyses with the complexity factor in determiner assignment. The complex context (2 ADJ) of this factor was elicited using items with the same visual complexity.

5.5 Conclusion and discussion

The results on grammatical gender have many different facets. The variable was investigated in common and neuter nouns, in judgement and in production, and in the assignment of determiners and the inflection of adjectives. The data of determiner assignment and adjectival inflection were also combined to investigate whether a gender feature is present and whether the rule for adjectival inflection is acquired. Finally, the effect of task and linguistic context on performance in SLI was examined, and linked to the processing measures that were selected in the previous chapter. In this section, the results are summarized and integrated with reference to the hypotheses presented in Figure 5.2.

The first hypothesis, that difficulties in performance on grammatical gender would be persistent in SLI, was confirmed. Persistent differences between SLI and TD were found for grammatical gender marking in determiners and adjectives. Both SLI groups had more difficulty than their TD peers in the production of definite determiners and in judging whether these were used correctly. Similarly, they had more problems determining whether adjectival inflections were correct in a judgement task and showed more inflection errors in production. For common gender, which is the default gender in Dutch, differences appeared to be small. Very large effects were, however, found for neuter gender, as expected. Children with SLI thus have persistent difficulties with grammatical gender, especially in the neuter domain.

The comparison between the results on determiners assignment and adjectival inflection revealed several interesting insights into the acquisition of grammatical gender in SLI. Although the older SLI group assigned neuter gender to determiners more often than the younger SLI group, correct adjectival inflection decreased. For adjectival inflection, individuals with SLI seemed to develop a strategy of relying more heavily on the default over time. This raises the question what the 'improvement' in determiner assignment in older subjects with SLI actually reflects: is a gender feature present? This question was answered by looking at the number of stable neuter nouns and by computing the proportion of correct adjectival inflection for those stable neuter nouns. Only a third of the children with SLI produced stable neuter nouns, and this number was reduced by half when adjectival inflection was taken into account. The number of individuals with SLI who produced correct adjectival inflection with stable neuter nouns did not increase significantly between the younger and the older group (7 individuals in the older group, opposed to 5 individuals in the younger group). The problem in SLI therefore seems to be multi-facetted: the majority did not acquire a gender feature at all: they either only produced the default form (*de* for definite determiners and –*ə*-inflections for attributive adjectives) or produced some neuter forms, but not consistently (representations remained unstable). Some individuals with SLI seemed to store nouns as being *het*-nouns, but only a few of them also had acquired the rule for adjectival inflection. The absence of differences in stable neuter gender assignment and in adjectival inflection between younger and older SLI groups seems to indicate that knowledge of grammatical gender fossilizes before adolescence.

These findings are in line with the existing literature on the acquisition of grammatical gender in SLI. Orgassa and Weerman (2008) already reported grammatical gender to be vulnerable in SLI, and indicated that children with SLI make the same type of errors as younger TD peers. Omission rates were, for instance, higher in SLI, which is characteristic of an early stage of article acquisition in TD (van Ewijk & Avrutin, 2010). Furthermore, overgeneralization errors were unidirectional (common for neuter), as is the case early stages of typical development. In this study, the unidirectionality was even stronger in the SLI groups than in the TD groups. TD children sometimes produced neuter determiners in common contexts while this rarely occurred in the SLI groups. This might reflect task effects in the TD group and a strong reliance on the default in SLI. Similar to the outcomes of the study by Weerman and colleagues (2011) and Keij and colleagues (2012) the children with SLI seemed to fossilize in early stages of gender acquisition. This study confirms their findings for an even older age group.

The second hypothesis was that performance on grammatical gender would show larger variability in SLI than in TD, dependent on the task or the linguistic context. Task effects in SLI were expected on the basis of previous findings by Keij (2009), who found improvement in the assignment of grammatical gender between 8 and 12 years of age in a knowledge task (forced choice), although this was not reflected in a production task. We found similar effects: while the older SLI group performed significantly better than the younger SLI group in the judgement of gender assignment in determiners, they did not perform significantly better in the production of determiners with the same nouns. Such effects of task were absent in the TD group. We must, however, be cautious regarding conclusions on the TD data. The TD groups scored at ceiling, which means similar task effects might be visible in scores that would show more variation. Correlational analyses do, however, support the idea that task effects were different in SLI: correlations between the judgement and the production scores for determiners and adjectives are stronger in the TD groups than in the SLI groups (r = .648 in TD, r = .433 in SLI), despite a smaller variance in scores. Individuals with SLI therefore revealed larger discrepancies between their judgement score and their production score.

In the tasks that were designed to test the production of grammatical gender, the influence of noun frequency and distance between determiner and noun was measured in the assignment of determiners. In adjectival inflection, performance was compared between items in which one adjective had to be produced and items in which two adjectives were required. Although noun frequency had a significant influence on the results, the effect was not larger in SLI than in TD. However, the other factors showed the hypothesized effect on performance. While scores for neuter determiner assignment in the SLI group dropped significantly when the distance between determiner and noun increased, no effect of this factor was found in the TD individuals. Similarly, the number of adjectives had a significant effect on performance in the SLI groups, but not in the TD groups. Although performance on adjectival inflection in neuter contexts was low in general in individuals with SLI, scores dropped significantly in more complex contexts. Previous studies that noted effects of distance or proximity between a noun and its dependent elements found these effects in processing or perception tasks (e.g. Keating, 2009). This study confirms that such factors also influence production.

The third and final hypothesis for grammatical gender was that effects of task or linguistic context would be related to processing abilities. In Chapter 4, we further specified this hypothesis into a relation with the processing measures from nonword repetition and sentence repetition tasks. These two measures showed persistent differences between SLI and TD and therefore potentially form an explanation for persistent differences in grammatical performance. The difference between the judgement and the production task, that was larger in SLI than in TD, correlated with sentence repetition and nonword repetition, as was hypothesized. Digit recall was additionally correlated with the difference between tasks, but inhibition and visual recall did not show any such relations. Similar significant correlations were found between the two processing measures and the effect of distance between determiner and noun. Subjects who showed a larger decrease in scores in a more complex task (production) or a more complex linguistic context (a greater distance between dependent elements), had weaker verbal processing abilities (sentence repetition, NWR and digit recall) and vice versa.

The context effect in the adjectival inflection task only partly confirmed our hypotheses. The effect of one or two adjectives that had to be inflected was found to be related to nonword repetition, but not to sentence repetition. Furthermore, an unexpected relation with visual recall appeared. As we already briefly discussed in §5.4.6, the absence of the hypothesized relation with sentence repetition and the presence of a relation with visual recall seems to indicate that it was the visual processing load of the more complex items that triggered variability in performance. However, similar relations would then be expected in the determiner assignment task, because the complex items in this task made use of the same visually complex pictures (see §5.3.3). The absence of similar relations in determiner assignment may indicate that visual recall only affects performance in grammatical aspects that are linguistically more complex (adjectival inflection rather than determiner assignment, because it requires gender assignment in combination with implementation of a grammatical rule).

In sum, grammatical gender is persistently difficult for individuals with SLI and knowledge of this grammatical feature seems to fossilize before adolescence. Although performance on determiner assignment and adjectival inflection was in general very low in both SLI groups, influences of the linguistic context in which performance was tested could still be measured. Performance was affected more negatively in the SLI groups, and this effect was related to (limited) verbal processing abilities. Although the problems with grammatical gender in Dutch subjects with SLI clearly indicate a problem at the representational level (knowledge fossilizes), their errors in the production of grammatical gender thus partly reflect problems in the implementation of knowledge. In Chapter 8, we will further elaborate on the theoretical and clinical implications of these findings.

6 Subject verb agreement

As fully discussed in §2.5, this thesis aims to answer the question whether the differences in grammatical abilities between SLI and TD that are often reported in childhood remain into adolescence. It also investigates whether performance in certain aspects of grammar is influenced by the type of task and linguistic context, and whether fluctuations in performance can be explained by variation in processing abilities. In the previous chapter, the grammatical abilities of children and adolescents were evaluated with respect to grammatical gender (Chapter 5). This chapter continues the evaluation of grammatical abilities in SLI but now looks at subject verb agreement.

Subject verb agreement, as was mentioned before in Chapters 1 and 2, is one of the grammatical domains that often appears to be problematic in SLI. To what extent this domain is problematic is, however, highly language-dependent. In languages with a rich inflectional system, children with SLI appear to have fewer problems in acquisition compared to the problems reported in languages with few inflectional distinctions in the verbal domain (Leonard et al., 1992). In the first section of this chapter, the paradigm of subject verb agreement in Dutch will be described (§6.1). This will be followed by a review of the literature on subject verb agreement in (Dutch) children with SLI and the specific hypotheses for this grammatical variable in the current study (§6.2). The tasks used to test comprehension and production of subject verb agreement are discussed in §6.3. In the final sections, the results are presented (§6.4) and then evaluated (§6.5).

6.1 The Dutch paradigm for subject verb agreement

Although many languages show agreement between the subject and the verb by means of verbal inflection, languages vary considerably in the features that are marked on the verb and the number of morphological distinctions made within the paradigm. Pro-drop languages, that is those languages that allow drop of the pronominal subject, often have a greater number of verb inflections than languages in which the subject is obligatory (Grela & Leonard, 2000; Leonard, 2000). Dutch is not a pro-drop language and, in the present tense, Dutch verbs are inflected for person and number, although the person distinction is only marked in the singular domain.

In the inflection of Dutch verbs, a distinction is made between 1st person singular, which appears with a bare verb stem (-ø), and 2nd and 3rd person singular, which are both marked by a suffix *–t* attached to the verb stem (see Table 6.1 for examples). With plural subjects, a suffix *–en* is added to the verb stem, regardless of whether the subject is 1st, 2nd or 3rd person plural.[1]

Table 6.1. The Dutch inflectional paradigm for verbs in the present tense (in non-inverted sentences)

	Person	Affix	Example	
Singular (SG)	1	***-Ø***	*Ik dans vaak*	'I often dance'
	2	***-t***	*Jij danst vaak*	'You often dance'
	3	***-t***	*Hij/zij/het danst vaak*	'He/she/it often dances'
Plural (PL)	1	***-en***	*We dansen vaak*	'We often dance'
	2	***-en***	*Jullie dansen vaak*	'You often dance'
	3	***-en***	*Zij dansen vaak*	'*They often dance'*

The form of the verb in second person singular contexts is dependent on the constituent order. Dutch is commonly typologized as an SOV/V2 language. That is, the basic word order is subject-object-verb (SOV), although finite verbs in declarative main clauses move to second position (V2), often resulting in an SVO order (Koster, 1975).[2] In subordinate clauses, on the other hand, the verb stays in final position (SOV). In SOV and SVO order sentences, the verb is inflected according to the paradigm described above. Sentences may, however, also start with another element than a subject, such as a topicalised adverb of time (e.g., *vandaag* 'today') or a question word. This causes inversion of the constituent order (VSO) and in these inverted sentences, the second person singular inflection *–t* is left out resulting in a bare stem (compare examples 1a and 1b).

(1a) *Vandaag dans jij*
Today dance-2.SG.INV you
'Today you dance'

(1b) *Jij danst vandaag*
You dance-2.SG today
'You dance today'

---

[1] The plural suffix *–en* is generally reduced to a schwa (-ə) in pronunciation.

[2] The typology of Dutch will not be discussed in any more detail here, since it is not within the scope of this thesis. For more details see the work by Koster (1975).

The Dutch inflectional system for verbs in the present tense shows a number of syncretisms, which may cause difficulties for the acquisition of the paradigm (Leonard, 2000). In non-inverted sentences, the *–t* suffix is used for both 2^{nd} and 3^{rd} person singular, whereas in inverted sentences the bare form is used for both 1^{st} and 2^{nd} person singular (as was discussed above). Furthermore, whenever a verb stem ends in a /*t*/ (e.g., [schiet] in *schieten* 'to shoot'*)* or a /*d*/, which is devoiced in final position (e.g., [rijd] in *rijden* 'to ride'), the distinction between a bare stem and a verb stem plus suffix *–t* is not evident in the spoken form. In these instances, children hear the same form for 1^{st}, 2^{nd} and 3^{rd} person singular. The most syncretic suffix is the *–en* suffix which is used for 1^{st}, 2^{nd} and 3^{rd} person plural, but also marks the infinitive in Dutch, often used with inflected auxiliaries and modals as is shown in examples 2a and 2b.

(2a)	*Ik ga dansen*	(2b)	*Ik wil dansen*
	I go-1SG.AUX dance-INF		I want-1SG.MOD dance-INF
	‘I will go dancing’		'I want to dance’

Another possible difficulty for acquiring the verbal paradigm is the fact that the auxiliaries that express tense and some auxiliaries for modality do not follow the exact same inflectional rules as lexical verbs. The auxiliaries for tense (*zijn* 'to be' and *hebben* 'to have') show little regularity in the singular domain, as can be seen in examples 3a and 3b.

(3a)	*Zijn*			(3b)	*Hebben*		
	1SG	*ik ben*	'I am'		1SG	*ik heb*	'I have'
	2SG	*jij bent*	'you are'		2SG	*jij hebt*	'you have'
	3SG	*hij/zij/het is*	'he/she/it is'		3SG	*hij/zij/het heeft*	'he/she/it has'

Some auxiliaries for modality do follow the paradigm for lexical verbs (e.g., *blijven* 'to keep'). However, in most modal auxiliaries no difference is made between 1^{st}, 2^{nd} and 3^{rd} person singular, which all have the same form (see example 4a with the verb *mogen* 'may (permission)').[3] Some modal verbs (*zullen* 'shall', *kunnen* 'can', *willen* 'want') allow an alternative form for second person singular (*jij zult* 'you shall', *jij kunt* 'you can', *jij wilt* 'you want'), in which a *–t* suffix is added and the vowel is altered (see example 4b). The 'modal paradigm' – if we can call it that - is therefore not entirely transparent.

---

[3] Following the categorization of auxiliaries and modal auxiliaries by Haeseryn et al. (1997).

(4a)	*Mogen*			(4b)	*2SG Alternatives*	
	1SG	*ik mag*	'I may'		2SG *jij zal/zult*	'you shall'
	2SG	*jij mag*	'you may'		2SG *jij kan/kunt*	'you can'
	3SG	*hij/zij/het mag*	'he/she/it may'		2SG *jij wil/wilt*	'you want'

Because modal verbs can also occur independently (with an implicit lexical verb as complement), as is shown in example 5, children come across independent finite verb forms of both the lexical and the modal paradigm, which may cause an even greater lack of transparency for children acquiring the inflectional rules.

(5) *Ik kan dat ook (doen*) ‘I can (do) that too’,

Despite the different inflectional paradigms and the many syncretisms in the Dutch verbal inflection paradigm, the system is acquired early in typical development. Children obtain high accuracy levels (90%) in subject verb agreement from three years onwards (Polišenská, 2010). In the acquisition of the verb inflection paradigm, children start with the use of (root) infinitival verb forms in sentence final position (see example 6). When they start using forms in second position (see example 7), they do so for a fixed set of verbs. Over time, the set of verbs that are used in finite form increases (Wijnen, 2000).

(6)	*mama eten*	(Fenna, 1;10)	(7) *zit water in*	(Tijn, 2;0)
	mommy eat.INF		sits.3SG water in	
	'mommy should eat'		'there is water in it'	

The next step in development is the use of more complex predicates (Wijnen & Verrips, 1998). In addition to an increase in the use of finite forms in second position (in terms of types and tokens), infinitives (INF) are now combined with a finite auxiliary (AUX) in second position. The onset of this third stage of using complex predicates seems to coincide with a decrease in the number of root infinitives, although root infinitives and finite forms of the same verb may co-occur in this stage (Wijnen, 2000).

Note that the first use of finite verb forms in early productions does not imply that these verb forms are correct in terms of agreement. Typical errors during acquisition are omissions or substitutions of inflections. Omission errors are believed to be more frequent in languages with a sparse verb morphology, while substitution errors are more often seen in rich verb inflection systems (Leonard, 2000). It should, however, be noted that it is not always possible to distinguish omissions from substitutions, as was previously discussed in §2.1. In Dutch, for instance, the 1st person singular has the same form as a bare stem. If a child uses a bare verb stem in a 3rd person singular context, it is therefore impossible to tell whether the inflection has been omitted or whether substitution of the 1st person singular form has taken place.

Dutch substitution errors usually have a clear direction: plural forms are not used in singular contexts, while singular forms are sometimes overgeneralized to plural contexts (de Haan, 1996). In the next section, the literature on subject verb agreement in Dutch SLI is reviewed. Hypotheses with respect to the acquisition of Dutch subject verb agreement in the younger and older SLI groups will be specified.

6.2 Subject verb agreement: literature review and hypotheses

As was described in the previous section, Dutch subject verb agreement is mastered early in typical development. Dutch TD children make errors by omitting and substituting verb inflections, but these errors show clear patterns: overgeneralization occurs from singular to plural contexts, not vice versa. The Dutch literature on grammatical abilities in children with SLI includes several studies on subject verb agreement.

In the study by Bol and Kuiken (1988) mentioned earlier in §2.1, it was noted that the spontaneous speech of children with SLI contained fewer first person singular verb forms in the present tense than the speech of younger language-matched TD peers. No differences were found in the number of third person singular verb forms that were produced. The authors indicate that differences in the frequency of specific verb forms between SLI and TD may indicate problems in verb morphology. However, the authors used a LARSP-type of analysis (Crystal et al., 1976) that only counted frequencies of inflections in correct contexts (a bare stem in first person singular contexts), without looking at grammatical errors (e.g., a bare stem in $2^{nd}/3^{rd}$ person singular contexts). Furthermore, they did not distinguish between the inflection of modals and lexical verbs.

De Jong (1999) was the first to study verb inflection in SLI in more detail, as was also mentioned in §2.1. From this study, based on elicited narratives from 35 children with SLI (mean age 7;8), three error types in the verb inflections of children with SLI became apparent: omissions, substitutions, and use of the infinitive (root infinitives). Omissions of agreement markers were seen in 3^{rd} person singular and plural contexts (note that narratives primarily trigger 3^{rd} person contexts). These errors occurred in both lexical verbs and in those modals/auxiliaries that follow the inflectional paradigm of lexical verbs. These omissions can, however, also be analysed as a 1st person singular substitutions since, as explained in the previous section, there is no distinction between the bare stem and the 1st person singular form (de Jong, 2002). Substitutions involved the use of singular verb forms in plural contexts. More specifically, this was the use of the third person singular affix *–t* in plural contexts, since the use of a bare stem was counted as an omission. The third type of error found was the use of infinitival verb forms at the end of the sentence without an inflected auxiliary, as shown in example 8.

(8) **En dan mama papa wakker maken*
And then mother father awake make.INF
'And then mother wakes up father'
(boy with SLI, 6;5, in de Jong, 1999, p. 72)

The example in (8) can be interpreted as the use of a root infinitival form of the verb in sentence-final position instead of a finite form in second position in the sentence (de Jong, 1999). Root infinitives have a nonfinite form but are often ascribed a finite interpretation. In that case, the sentence would be paraphrased as *'en dan maakt mama papa wakker'*. Another possible interpretation would, however, be that an auxiliary (*gaat* 'is going') is omitted in second position, suggesting the sentence could be paraphrased as following: '*en dan gaat mama papa wakker maken'* which translates into 'and then mother is going to wake up father' (de Jong, 1999). This latter interpretation was favoured by Klein (1974), who studied child-directed speech and argued that it is dominated by AUX + INF constructions. The third option of interpreting these examples as misplaced plural forms seems less probable, since all infinitival forms were produced with a singular subject (de Jong, 2002).

While the first two types of errors reported by de Jong (1999) (omissions and substitutions) still occur in younger TD children matched on MLU, the third type of error is remarkable since TD children with the same language age (MLU) no longer produce these errors. The use of infinitives is typically seen in a very early stage of typical development, when children produce two or three word utterances (Wijnen & Verrips, 1998). The persistent use of root infinitives in SLI has also been reported in other languages and has led to the hypothesis that children with SLI have a prolonged period in which tense marking is treated as optional and root infinitives are used next to finite forms (as was also discussed in §1.2). The name of this hypothesis is the Extended Optional Infinitives (EOI) Hypothesis (Rice & Wexler, 1996).

The EOI hypothesis was tested in Dutch subjects in a study by Wexler, Schaeffer and Bol (2004). They compared spontaneous language samples of 20 children with SLI (4;2 to 8;2 years old) to those of 47 younger TD peers (1;7 to 3;7). Their hypothesis was that children with SLI do have knowledge of the verbal inflectional paradigm, but do not know that verbs are obligatorily marked for agreement and tense. Children with SLI therefore keep using root infinitives. The number of infinitival root lexical verbs and the number of finite lexical verbs were counted (auxiliaries, modals and copulas were excluded from the analyses). The number of root infinitives appeared to decrease as a function of MLU in both SLI and TD children and as a function of age in TD children (83% for the one- to two-year-olds but declining to 7% in the three-year-olds). Overall, the percentage of root infinitives in the SLI group was 15%, but children with SLI with a low MLU of 2 to 3 morphemes (5 children) produced 30% root infinitives (27 out of 89 verbs), while those children with a higher MLU of 4 to 6 morphemes produced around 6% root infinitives (13 out of 238 verbs). The use of root infinitives did not seem to decrease with age in SLI, and they were still found at age 8. The study lacks statistical analyses, but from the trends in the data the authors conclude that children with SLI are still producing root infinitives at the age TD children have moved on from this acquisitional stage. Furthermore, because children with SLI have rates of finite verb forms that are comparable to MLU matched controls and place finite forms correctly in second position, the authors conclude that the paradigm of finite morphology and the syntactic rules have been learned.

As well as reporting on the occurrence of root infinitives, the study by Wexler and colleagues (2004) also gives the rate of agreement errors. Here auxiliaries, modals and copulas were included in the analysis.[4] The data shows that children with SLI produced between 3 and 12% incorrect inflections, again correlated with MLU. The highest error rates were found in the use of a plural ending with a singular subject. This contrasts with other studies on subject verb agreement in older children with SLI, where plural markings in singular contexts were rather infrequent (de Jong, 1999; Steenge, 2006; Orgassa, 2009; Weerman et al., 2011). The authors indicate that this high rate of errors is caused by children adding a schwa to the verb which makes the verb form sound plural-like (*je moet-ə die hebbe* 'you must.AUX-ə that have.INF', p. 187). The addition of a schwa may be triggered by the trochaic template of Dutch, in which strong syllables are followed by a weak one. The errors may thus well be the result of suprasegmental phonology rather than morphology. The authors conclude that agreement performance is very good, indicating that children know the verbal inflectional

[4] Note that auxiliaries, modals and copulas can have a deviating inflectional paradigm, as we described above. The amount of agreement errors may be influenced by these differences.

paradigm. According to them, the problem in SLI thus mainly seems to be that they do not leave the developmental stage at which finiteness marking is optional (Wexler et al., 2004).

Wexler and colleagues (2004) thus propose the use of infinitives in the language production of children with SLI as a core characteristic of their problems with verb inflections. Infinitives are judged to occur frequently and persistently until age 8, while other 'verb inflection errors' like omissions or substitutions are considered rather infrequent. The question is, however, whether the methods of analysis used in this study do not put disproportional emphasis on the incorrect use of infinitival forms in the language production of children with SLI. The analysis of root infinitives was performed separately from the analysis of other 'verb inflection errors', and only on the subset of lexical verbs. The occurrence of other 'verb inflection errors' was, on the other hand, calculated for the total number of verbs. Only trends were reported and no statistical analyses were done. Furthermore, it is debatable whether the prolonged use of root infinitives can be adequately described in terms of a prolonged period of optionality of finiteness marking.

As was described in §6.1, studies on Dutch typical development show that the stage in which infinitives are used predominantly cannot be interpreted in terms of optionality of finiteness marking. As de Jong (2002, p. 162) puts it: in the infinitival stage in Dutch typical development, there is "nothing optional about it. Nonfiniteness is obligatory". This stage is followed by a stage in which finiteness marking is applied consistently to an increasing but fixed set of verbs (no optionality there either). Root infinitives only co-occur with finite forms for the same lexical verb in the third stage, when children start using more complex predicates (AUX + INF) next to finite verbs in second position. It remains, however, unclear whether 'optionality' truly describes what happens in this developmental stage. With the first use of complex predicates (the use of infinitives in finite constructions with an auxiliary), the use of root infinitives rapidly decreases (Wijnen, 2002). Even if we could call the latter stage a period of optionality, it is not clear which developmental stage is extended in SLI: the stage in which some verbs are obligatorily finite and others are root infinitival, or the stage in which root infinite and finite forms of the same verb co-occur. These two points indicate that it remains unclear whether infinitival forms in the language production of children with SLI can be interpreted as support for the existence of an extended optional infinitive stage (finiteness marking is optional).

When studies are considered that compare the three error types in SLI in the same set of verbs, infinitives are reported as occurring less frequently than omissions and substitutions (de Jong, 1999; Steenge, 2006; Orgassa, 2009). A study of verb morphology in the Frog Story narrative compared 24 children with SLI aged between 7 and 9 years old to age-matched TD peers (Steenge, 2006). In the verb production of the children with SLI, omissions and substitutions occurred much more frequently than root infinitival forms. When compared statistically, the use of infinitives was not significantly higher in children with SLI than in the TD group. Omission of third person singular did, however, indicate clear group differences (Steenge, 2006; Verhoeven et al., 2011). Orgassa (2009) also looked at monolingual children with SLI between 6 and 8 years of age, but examined verb inflection errors in an experimental task. Root infinitives occurred so infrequently in this data set that statistical analyses could not be performed on this error type (Blom et al., 2013a). Although infinitival forms do occur in the language production of children with SLI to some extent, these studies indicate that it is not the most frequent and characteristic verb inflection error in this population. This suggests that Wexler and colleagues' claim for the importance of infinitival forms in children with SLI is unwarrantable.

The differences between SLI and TD children in the rate of verb inflection errors are substantial, although overall accuracy is usually fairly high. Inflection errors can even be used as a clinical marker with fairly high sensitivity and specificity (87% and 75% respectively) (Blom et al., 2013a). Within the finite verb productions of children with SLI, the use of a bare form is reported to be the dominant error type, as is the case in younger language-matched TD children (Steenge, 2006; Orgassa, 2009). In contrast to younger peers, the SLI group, however, also uses the third person singular *–t* suffix in first person singular and plural contexts (Orgassa, 2009). It is mainly this last type of error that seems to persist when children become older, although overall performance improves (Weerman et al., 2011). Interestingly, experiments on the perception of Dutch subject verb agreement showed that children with SLI only detected substitution errors, while omission errors were not noticed. Apparently, the most frequent errors in language production are also the hardest to notice for children with SLI (Blom et al., 2014). This indicates that problems with subject verb agreement are not restricted to production, but can also be found in perception or judgement tasks.

So far, we have only discussed the frequency of errors in subject verb agreement. We should, however, also mention that inflection of lexical verbs can be avoided by using (dummy) auxiliaries + infinitives. Dummy auxiliaries carry grammatical features but do not add any meaning to the sentence (Blom, van de Craats, & Verhagen, 2013b). An example is shown in (9) with the verb *gaan*, which has an inchoative meaning in adult Dutch ('the girl is going to read a book') but in child Dutch often does not carry the adultlike aspectual meaning.

(9) *het meisje gaat een boek lezen*
the girl goes.DUMMY-AUX read a book.INF
'the girl is reading a book'

The use of dummy auxiliaries is a stage in the typical acquisition of finiteness marking (Zuckerman, 2013). Constructions with a dummy auxiliary and an infinitive are regarded as less complex options than constructions with an inflected lexical verb, both in terms of processing costs and in terms of linguistic complexity. Auxiliaries are assumed to be stored as chunks instead of inflected in the production process, and auxiliary constructions do not require syntactic movement of the lexical verb to second position (de Jong et al., 2013). The latter argument only holds for matrix sentences, where the finite verb is in second position (as opposed to subordinate clauses in which the finite lexical verb is in final position). In typical development, the use of dummies is also restricted to matrix sentences (van Kampen, 1997; Zuckerman, 2001), which is an extra indication that dummies are (partly) used to avoid syntactic movement. Although not all studies report overuse of dummy auxiliaries in the production of children with SLI (Bastiaanse et al., 2002), some studies report significantly more use of these constructions in SLI subjects compared to younger TD children (de Jong, 1999; de Jong et al., 2013; Zwitserlood, 2014). The problems in subject verb agreement may therefore not only be found in inflectional errors, but also in the use of alternative strategies (de Jong et al., 2013).

The Dutch studies on subject verb agreement in SLI have reported variability in accuracy dependent on the linguistic context in which performance is tested, as was the case in English subjects with SLI (Bishop, 1994) (see §2.3). In the study by Weerman and colleagues (2011), accuracy was, for instance, lower in novel verbs than in existing verbs. Furthermore, accuracy was reported as lower in main clauses compared to subordinate clauses (Weerman et al., 2011; Blom et al., 2013a; Blom et al., 2014). The influence of sentence type was tested with an elicitation task in which the subject of the subordinate clause was already provided and children had to finish the sentence. The task thus mainly tested the effect of the position of the verb and corresponding movement operations. The relative ease of verb inflections in subordinate clauses suggests that verb inflections are easier in sentence final position than in second position, probably due to the difference in movement operations (Weerman et al., 2011).

As was described in §2.3, verb inflection accuracy has also been shown to be influenced by the phonological properties of the verb stem (Marshall & van der

Lely, 2007; Song et al., 2009; Blom et al., 2014). In the study by Blom and colleagues (2014), omissions of the *–t* suffix were observed as more frequent with verb stems ending in an obstruent than with verb stems ending in a sonorant, both in SLI and in TD. This is in line with English studies, where effects of obstruent versus non-obstruent codas were found for past tense inflections (Oetting & Horohov, 1997; Johnson & Morris, 2007). In TD children, a difference in third person singular marking was also found between two obstruent contexts (fricatives and plosives). Errors were more frequent after plosives than fricatives, while no such difference was found in the SLI group. The authors indicate that this may well be an effect of the complexity of the verb codas because the set of verbs ending in a plosive contained one verb ending in a consonant cluster, while the set of verbs ending in a fricative did not (Blom et al., 2014). In fact, many studies have shown effects of the complexity of a coda on verb inflections before (Marshall & van der Lely, 2007; Song et al., 2009).

In sum, the literature indicates that the inflectional errors and avoidance strategies of children with SLI generally mirror the error patterns in the production of younger TD children, which seems to indicate delayed acquisition rather than deviant development. Although the representation of the inflectional paradigm appears to be weak in SLI, the large number of correctly produced forms in SLI indicates that children are capable of learning these inflections (de Jong, 1999). Errors are however still reported until a fairly late age, indicating that subject verb agreement is a vulnerable area in SLI. Inflection accuracy seems to be also related to the syntactic context in which agreement is elicited and to the phonological characteristics of the verb stem. Next to errors in inflection, subjects with SLI seem to avoid inflections by using dummy auxiliaries instead.

On the basis of the review of the literature, differences between SLI and TD are expected to be present and to be persistent (see Figure 6.1 for a schematic illustration of the hypotheses in this chapter). We expect group differences in accuracy between the SLI and TD groups in the judgement and production of subject verb agreement. Furthermore, we expect the older SLI group to perform better than the younger SLI group, although scores are not expected to reach ceiling at adolescence. A significant interaction between Age and Group is therefore hypothesized, but differences between the older SLI and the older TD group are still expected.

HYPOTHESES SUBJECT VERB AGREEMENT

* Persistence grammatical problems (§ 6.4.1 & § 6.4.2)

judgement
production
Effect Group & Age (SLI<TD) – effect Group still present in adolescence

* Variability in performance (§ 6.4.3)

linguistic context → effect size interacts with Group (larger in SLI than in TD)

* Relation variability and processing load (§ 6.4.4)

correlations between linguistic context effect and processing measures

Figure 6.1. Schematic illustration of the hypotheses for subject verb agreement

We furthermore expect to find variability in performance dependent on the linguistic context in which subject verb agreement is produced, and we expect this variability to be larger in the SLI group. Within the production task, we manipulated the sonority and the complexity of the verb stem coda (phonological context factor). Furthermore, verb inflections were tested in main clauses and subordinate clauses (syntactic context factor). All factors have previously been shown to affect performance on verb inflection. The influence of the syntactic context factor was, however, tested using a different method from previous task designs (e.g., Blom and colleagues, 2008a). In contrast to previous designs, in which subordinate (relative) clauses were elicited by providing participants the subject of the clause and asking them to finish the sentence (*dit is de vrouw die…*'this is the woman that….') participants in this study were only presented with a main clause and had to construct a subordinate clause themselves (*de jongen ziet dat…*'the boy sees that…'). In this way, the influence of the complexity of the syntactic structure on subject verb agreement was tested. Note that we cannot separate the effect of syntactic movement/position of the verb and the effect of complexity of the sentence structure in our task design. The results of this study can, however, be compared with the outcomes of the studies discussed above, to discover the effect of syntactic structure.

Although all groups will show variability to some extent, effect sizes are hypothesized to be larger in the SLI groups compared to the TD groups. We therefore expect to find interactions between Group and the effect of phonological or syntactic context.

Finally, we expect that the effects of context are related to processing capacity. Inserting inflections after phonologically more complex verb stem codas or in sentence structures that involve movement of the verb (main clauses) is thought to be more costly in terms of information processing, as previously described in §2.4. Correlations between the context effects and the processing variables that showed persistent differences between SLI and TD (as described in §4.5) are therefore expected. In the next section, the tasks used for testing subject verb agreement will be described.

6.3 Task descriptions

As indicated earlier (see §3.4), each variable was tested in a perception and a production task. For subject verb agreement, a judgement task and two production tasks were used. The verbs used were the same across the different tasks. The selection of verbs is described in §6.3.1. Then the judgement task (§6.3.2) and the two elicitation tasks (§6.3.3) are presented.

6.3.1 Verb selection

For the subject verb agreement tasks, six transitive verbs where chosen that were likely to be familiar to children and could be captured in a picture. As was described in §3.4, the phonological characteristics of the verbs were varied. The last consonant of the verb stem differed in sonority (sonorant > fricative > plosive) and the complexity of the verb stem coda varied (1 or 2 consonants). All verbs had monosyllabic verb stems to avoid effects of word length. Table 6.2 shows the selected verbs and their corresponding numbers (V1 to V6) for convenience in the presentation and discussion of the results. The verb *aaien* ‘to stroke′ was used in the practice items.

Table 6.2. Verb selection subject verb agreement task

	Sonorant	Fricative	Plosive
Simple coda	*kam* ‘comb’ (V1)	*lees* ‘read’ (V3)	*bak* ‘bake’ (V5)
Complex coda	*film* ‘film' (V2)	*poets*‘polish' (V4)	*drink* ‘drink’ (V6)

In the categorization of these different verbs into simple or complex stem codas it is important to note that phonological processes that change complex codas into simple ones may play a role. The consonant cluster in the verb stem *film* ‘film′ is, for instance, often realized as /*fIl-əm*/ in production (Booij, 1995, p. 127), by inserting a schwa between the two consonants of the verb stem coda. When a *–t* suffix is added to the verb stem *drink* ‘drink’, the /*k*/ can be omitted from the inflected form /*drIŋkt*/ in production without being noticed by listeners (Booij, 1995, p. 137). It is therefore hard to tell whether this will be produced as a complex coda or not. With respect to the complex coda in *poets* ‘polish', pronounced as /*puts*/, it is also important to remark that in some phonological analyses, the combination /*t*/ and /*s*/ is seen as one consonant: an affricate /*ts*/ (Brosnahan, 1959). In that case the question arises whether *poets* is a verb stem with a complex coda. Because there is individual variation in all three cases, it is not possible to categorize the verb stems definitively as complex. However, complexity might affect the representation or processing of these verbs. These remarks will therefore be kept in mind and taken into account once we start describing the results. The categorization into plosive, fricative and sonorant endings is less debatable, and effects of this variable on the number of errors in subject verb agreement have already been attested (Blom et al., 2014).

6.3.2 Judgement task

As was described in §3.4, a judgement task was constructed for every variable to test the perception of grammatical aspects. The judgement task for subject verb agreement had the same design and procedure as the judgement task for gender (see §5.3.2). Participants were presented with short phrases, consisting of a pronominal subject, a verb and an object. The inflection of the verb was either correct or incorrect, and participants had to judge whether the phrase was correct or not by pressing a green or a red button on the screen. The items were accompanied by a semantically matching picture to facilitate processing. Figure 6.2 illustrates the computer screen used for the presentation of the items in the judgement task. For a more detailed description of the procedure, see §5.3.2.

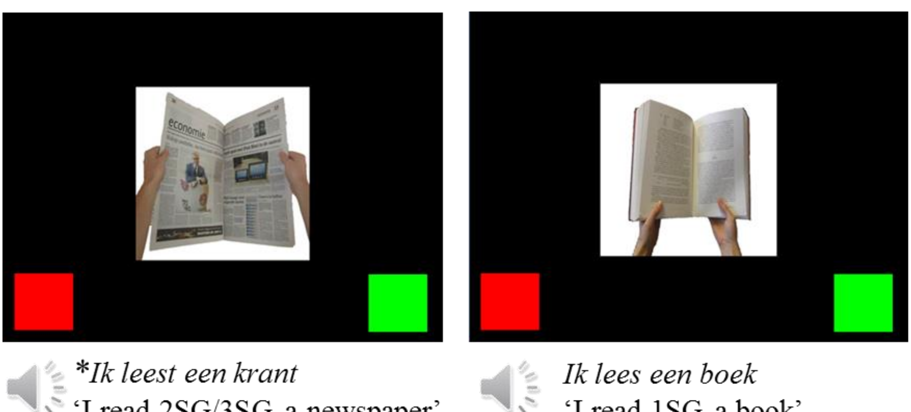


Figure 6.2. Examples of auditory stimuli and accompanying screen shots used with the judgement task for subject verb agreement

Every verb was tested with both correct inflections and incorrect inflections. Since substitutions of plural suffixes in singular contexts neither occur in typical development (Polišenská, 2010) nor in the language production of subjects with a language impairment (de Jong, 1999; Orgassa, 2009; Weerman et al., 2011), this type of error was not included. All other possible error types were included, resulting in a total of 12 items per verb. Examples of these 12 items for the verb *lezen* ‘to read’ are shown in Table 6.3.

Table 6.3. Correct and incorrect inflections for the verb *lezen* ‘to read’ as presented in the judgement task for subject verb agreement

	-Ø	*-t*	*-en*
1SG	*ik lees* ‘I read’	* *ik leest* ‘I read’	
2SG	* *jij lees* ‘you read’	*jij leest* ‘you read’	
3SG	**hij lees* ‘he reads’	*hij leest* ‘he reads’	
1PL	**wij lees* ‘we read’	**wij leest* ‘we read’	*wij lezen* ‘we read’
2PL	**jullie lees* ‘you read’	**jullie leest* ‘you read’	*jullie lezen* ‘you read’

Because these 12 correct and incorrect inflections were tested with six different verbs, the whole judgement task for subject verb agreement consisted of 72 items: 30 correct inflections and 42 incorrect inflections (see Appendix C for the total list of items). Two items had to be excluded, since it appeared that adults could not judge these items correctly (items were only included in the analyses if at least 10 out of 14 adults judged the item correctly). Both items concerned the omission of the *–t* suffix in 2nd and 3rd person singular for the verb *poetsen (*jij poets je tanden* 'you brush your teeth' and **zij poets haar schoen* 'she polishes her shoe'). We will come back to the properties of this verb *poetsen* in §6.4.2. The total number of items was therefore 70 instead of 72 (30 correct inflections and 40 incorrect inflections). The number of correct judgements was scored automatically (either 0 or 1). The total number of correct responses was counted and converted into percentages correct. As previously mentioned in §5.3.2, the items for subject verb agreement were randomly mixed with the items for grammatical gender. The items of the two tasks served as each other's fillers. The judgement task took in total around 15 to 20 minutes per subject (including fillers), but the task was split into two subtests in order to avoid effects of decreasing attention or fatigue.

6.3.3 Elicitation tasks

To elicit the production of subject verb agreement, two elicitation tasks were constructed, partly based on previous task designs for Dutch (Blom et al., 2008a and Unsworth et al., 2014), as was the case for the gender tasks (§5.3.2).

Elicitation game

In an elicitation game, items involving first, second and third person singular (1SG, 2SG and 3SG) and first and second person plural (1PL and 2PL) were used. The experimenter told the participant that they would play a little card game, and introduced a third participant, Kim, to the game. For child participants, Kim was a doll sitting on the table. With adolescent participants, a photo of a girl was used instead. Figure 6.3 illustrates the setup of the elicitation game. The experimenter, the participant and Kim were each seated at a different side of the table and had a pile of cards in front of them with pictures showing the actions expressed by the six verbs to be elicited. The pictures showed the actions with an anonymous mouth or hand (or both).

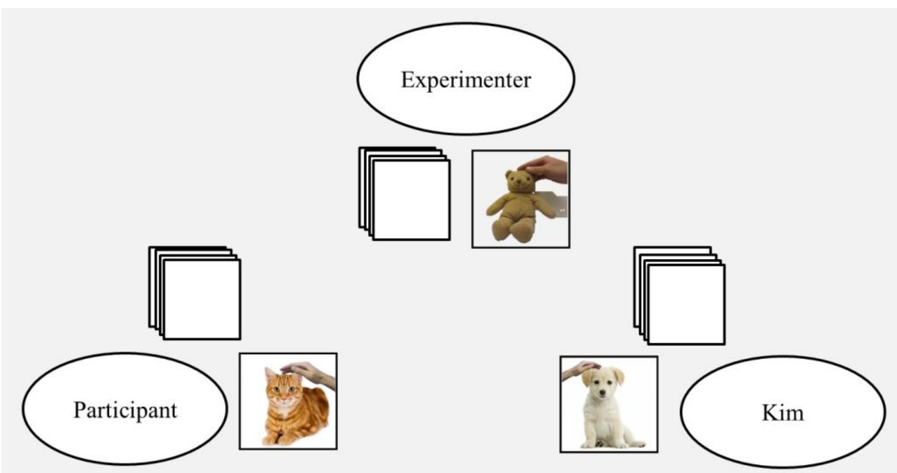


Figure 6.3. Setup of the elicitation game for subject verb agreement

Before the game started, the experimenter showed and named all the actions presented on the cards in infinitival form. The verb *aaien* 'to stroke' was added as a practice item.

> *hier zie je poetsen, lezen, drinken, bakken, filmen, kammen en aaien*
> 'here you see brushing, reading, drinking, baking, filming, combing and stroking'

The experimenter distributed the three piles of cards and explained that they all (i.e. the experimenter, the participant and Kim) would turn over a card from their own pile at the same time. The experimenter turned the cards in front of Kim. In the practice phase, the experimenter explicitly asked the participant to name everyone's actions.

wat doe jij, wat doe ik, wat doet Kim?
'what are you doing, what am I doing, what is Kim doing?'

These questions elicited the 1SG, 2SG and 3SG inflections. If two players had the same action on their card, the participant was instructed to describe them together in order to elicit 1PL and 2PL inflections. Table 6.4 illustrates the cues in the first three practice trials and the different target answers. There were seven practice items, but the task was started earlier when participants seemed to understand the task and all contexts (1SG, 2SG, 3 SG and 1PL, 2 PL) had been practised. Many participants only needed three or four practice items to grasp the idea and start the task. After the practice phase, the cues were left out to avoid giving away the correct inflections on the verb *doen* 'to do'.

Table 6.4. Examples of questions experimenter and target items in first three practice trials

Experimenter:	*Wat doe jij, wat doe ik, wat doet Kim?* 'What are you doing, what am I doing, what is Kim doing?'		
Target:	*Ik aai de beer*	'I stroke the bear'	*1 SG*
Target:	*Jij aait de poes*	'you stroke the cat'	*2 SG*
Target:	*Kim aait de hond*	'Kim strokes the dog'	*3 SG*
Experimenter:	*Nu doen wij hetzelfde en Kim doet iets anders. Kan je vertellen wat we doen?* 'Now we are doing the same and Kim is doing something different. Can you tell me what we do?		
Target:	*wij aaien de hond*	'we stroke the dog'	*1 PL*
Experimenter:	*Nu doen Kim en ik hetzelfde en jij doet iets anders. Wat doen we?* 'Now Kim and I are doing the same and you are doing something different. Can you tell me what we do?'		
Target:	*jullie aaien de kat*	'you stroke the dog'	*2PL*

For every context (1SG, 2SG, 3SG, 1PL, 2PL), an inflected form was elicited twice with every verb, adding up to a total of 60 items. A few extra items had to be elicited because there were three players and not every turn could be used efficiently to elicit only target answers. The extra items were, however, removed from the analyses. A full list of the elicited items and the pictures that were used can be found in Appendix D and E.

In production tasks, it is not always possible to obtain the target answer. In the case of the elicitation of subject verb agreement, participants sometimes produced a (dummy) auxiliary with an infinitive instead of using an inflected lexical verb (e.g., *Kim gaat de beer aaien* 'Kim is going to stroke the bear'). In these cases, additional cues were given to obtain a target answer with an inflected lexical verb. The participant was asked to give a response without the auxiliary (*kun je het ook zeggen zonder gaat* 'can you tell me without using the word go') or was asked to only use the lexical verb (*kan je het zeggen met alleen aaien* 'can you tell me only using 'stroke'). This requires metalinguistic awareness and some younger participants did not understand these cues. In those cases the experimenter gave an extra phonological cue by starting the answer and only pronouncing the first phonemes of the verb, not giving away the inflection (*Kim a…* 'Kim str…'). Participants also sometimes used different lexical verbs than the target ones. Instead of *poetsen* 'to polish', participants, for instance, sometimes used the verb *schrobben* 'to scrub'. In those cases, the participant was simply reminded what the action on the picture had been called when it was introduced at the beginning of the game (*dit was poetsen* ‘this was called polishing’). They were then asked to repeat the item. The number of cues was adjusted to the needs of the participant. When a participant seemed to get frustrated, the experimenter would continue with the next item. If the first answer of the participant was a root infinitive, no additional cues were given.

Answers including an auxiliary (also modals) were counted and will be reported on separately, but were not included in the analyses for the accuracy of subject verb agreement. If inflection of the target verb had been prompted by providing an extra cue, this answer was included in the analyses. If no target response was obtained, the item was coded as missing data and excluded from the analysis.

Participants had to produce at least 30 out of 60 target verbs in order to be included in the analyses. Sometimes, the experimenter mixed up the piles of cards, which changed the contexts of the paradigm and resulted in some missing items (5 items per child). This happened especially in the younger TD group where tests were administered with help of assistants. The data was then coded as missing, but with a different score in order to be able to distinguish between missing items due to experimenter error and missing items due to child factors. Table 6.5 provides an overview of the total number of elicited items and the number and percentage of missing items due to experimenter error. This data was excluded from further analyses.

Table 6.5. Overview of the elicited items and missing items due to experimenter error

		Elicited items	Missing items due to experimenter error	
	N	N	N	%
SLI-Y	32	1920	16	0.83
TD-Y	32	1920	190	9.90
SLI-O	31	1860	6	0.32
TD-O	30	1800	10	0.56

For accuracy analyses, the numbers of correct and incorrect inflections were counted (score 1 and 0). The incorrect inflections were also analysed in more detail to have information on the type of errors. Additionally, the number of auxiliary + infinitive constructions was counted. All types of auxiliary constructions were taken into account: both dummy auxiliaries and non-dummy auxiliaries and both bare infinitives and infinitives in combination with aspect markers (*te* + INF or *aan het* INF). Analyses of the effect of phonological properties of the verb stem coda were carried out on the amount of correct and incorrect responses in 2SG and 3SG contexts, because effects are hypothesized to show up with *–t* inflections, and not with *–ø* or *-en* inflections.

Elicitation task with different syntactic contexts

In a second elicitation task for subject verb agreement, the influence of syntactic context on subject verb agreement was tested (see §2.3 for a list of the manipulated factors in the production tasks). Inflections were elicited in the missing context 3rd person plural (3PL), in combination with third person singular (3SG) because the highest frequency of errors in subject verb agreement has been found in these contexts, as was described in §6.2. Different syntactic constructions were elicited in order to examine the impact of the complexity of syntactic context.

Verb inflections were tested in main clauses and in subordinate clauses. As was described in §6.2, an alteration in the task design of previous studies was made. In contrast to the design by Blom and colleagues (2008a), in which the subject of the subordinate (relative) clause was provided in the elicitation cue, participants in this study had to construct the subordinate clause from scratch. On top of the influence of syntactic context in terms of movement of the verb (which was the focus of Blom and colleagues), this study thus tested the influence of complexity of the sentence structure on the production of verb inflection. The two factors are not separable in our results, but by comparing the results of this study to the outcomes of the task constructed by Blom and colleagues, the effect of sentence structure can be evaluated.

For the same set of six verbs, inflections were elicited twice per context (3SG and 3PL) in two different sentential structures (main clause and subordinate clause). A third sentence type with an inverted verb and subject was included in order to be able to report on accuracy in verb placement. Those items will not be reported on here and were excluded from the accuracy analysis of subject verb agreement. Per sentence structure, 24 items were elicited (2 items x 2 contexts x 6 verbs). In total, the elicitation task for subject verb agreement in different syntactic contexts consisted of 72 items. However, as already described in §5.3.3, the items for the elicitation of subject verb agreement were mixed with the items for the elicitation of grammatical gender (determiners and adjectives). The gender items served as fillers for the subject verb agreement task and vice versa. The elicitation task was administered in two sessions of 15 minutes.

As in the elicitation task for grammatical gender, sentences were elicited with the help of pictures presented in Powerpoint. Pictures presented men and women performing the actions expressed by the verb with contrasting objects (combing a cat versus a dog, filming a flower versus a lamp, reading a newspaper versus a book, polishing a shoe versus brushing teeth, baking pancakes versus eggs, drinking water versus milk).[5] Participants had to complete sentences started by the experimenter that referred to the pictures on the screen. For each sentence type, a few practice items were included to train the participants to give target-like responses. In Table 6.6, examples of the pictures, the experimenter cues and the target utterances are shown for every sentence structure. A full list of elicitation cues and target answers for subject verb agreement can be found in Appendix F.

[5] In Dutch, 'brushing teeth' and 'polishing shoes' translate into the same verb *poetsen.*

Table 6.6. Examples of elicitation of subject verb agreement in different sentence contexts

Category	Experimenter	Target
Main Clause 	*De vrouw….* 'the woman….' *en de man….* 'and the man…'	*leest een boek* 'is reading a book' *leest een krant* 'is reading a newspaper
Inversion 	*Op dit plaatje….* 'in this picture….' *En op dit plaatje….* 'and in this picture….'	*leest de vrouw een boek* 'the woman is reading a book' *leest de man een krant* 'the man is reading a newspaper
Subordinate clause 	*De jongen ziet dat….* 'the boy sees that….' *En hij ziet ook dat….* 'he also sees that ….'	*de vrouw een boek leest* 'the woman is reading a newspaper' *de man een krant leest* 'the man is reading a book'

Similar to the elicitation game, additional cues were provided if the participant produced a non-target answer such as an auxiliary + infinitive or an inflection of a non-target lexical verb. If the extra cue prompted a target-like answer, this answer was included in the analyses. If no target response was obtained after providing a cue, the item was coded as missing data and excluded from the analysis. For the analysis of the accuracy of subject verb agreement in different syntactic contexts, the answers were coded as correct (score 1) or incorrect (score 0). Additionally, the number of auxiliary + infinitive constructions was counted. All types of auxiliary constructions were taken into account: both dummy auxiliaries and non-dummy auxiliaries and both bare infinitives or infinitives with progressive aspect markers (*te* + INF or *aan het* INF).

6.4 Results

The results of the judgement task will be presented first (§6.4.1), followed by the results of the elicitation game and the elicitation task (§6.4.2). The influence of the phonological and syntactic contexts will be evaluated in §6.4.3, to determine whether individuals with SLI show (a higher) variability in performance. Subsequently the correlation between the outcomes on subject verb agreement and the processing measures will be examined to answer the question whether fluctuations in performance (context effects) are explained by (limited) processing capacities (§6.4.4).

6.4.1 Judgement

Figure 6.4 shows the results of the judgement task for subject verb agreement (in tables and figures abbreviated as SVA). For each group, the range of accuracy scores in percentages is represented by the boxes and whiskers. The bold line in the middle of each box is the median. The graph shows that adults performed at ceiling (98.5%).[6] The TD groups showed some variance in performance, with scores between 80% and 100% correct in the younger TD group and between 90% and 100% in the older TD group. There was a large variance in scores in the SLI subjects, especially in the younger SLI group, where scores ranged from 40% to 100%. The overlap between the younger SLI and the younger TD group was small, indicating that although some children with SLI performed much better than others, they almost all performed worse than their TD peers.

ANCOVAs with the number of correct judgements of subject verb agreement as dependent variable, with Group and Age as fixed factors and with IQ as a covariate (see §3.2 for a rationale) showed main effects for Group ($F(1,118) = 79.71$, $p < .001$, $\eta_p^2 = .403$) and Age ($F(1,118) = 36.52$, $p < .001$, $\eta_p^2 = .236$). Group and Age also interacted significantly ($F(1,118) = 25.41$, $p < .001$, $\eta_p^2 = .177$), as expected. Post-hoc analyses revealed that the difference between SLI and TD was still present in the older group, although much smaller than in the younger groups ($F(1,57) = 14.82$, $p < .001$, $\eta_p^2 = .206$).[7,8] As was the case in the judgement of gender, the performance of the younger SLI group seemed to be only just above chance level (61.5% on average), whereas a more detailed analysis indicates a pattern.

---

[6] The errors that constituted the 1.5% error rate in the judgement of subject verb agreement in adults did not show any consistencies. Errors were made in different verb contexts and by different adults. They therefore seem to reflect incidental errors, probably due to lapses in attention.

[7] As was mentioned in §3.4, the analyses were repeated with a smaller sample in which the participants with very few rejections (<10) were eliminated. Five subjects in the SLI-Y group appeared to meet this criterion. Outcomes remained the same when these subjects were excluded.

[8] The same results were obtained with binary logistic regression analyses.

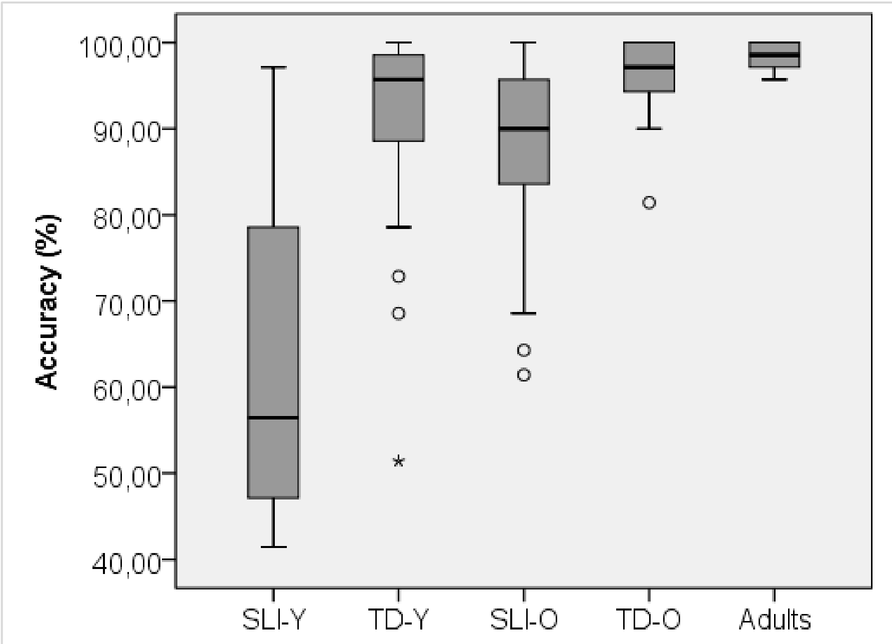


Figure 6.4. Accuracy (%) in the judgement of SVA per group

Table 6.7 shows the results of the judgement task, divided into correct acceptance of correct inflections and correct rejection of incorrect inflections. As was the case in the grammatical gender task, the younger children with SLI mainly had problems in rejecting the incorrect inflections of the verbs. They accepted correct inflections quite accurately (around 84%), but were not accurate in rejecting incorrect forms (around 45%). This may indicate that their representations are unstable, or might alternatively reflect insecurity in making linguistic judgements, as has been discussed earlier in §5.4.1. An explanation in terms of chance performance (pressing a button randomly) can however be ruled out. Low accuracy rates were caused by difficulties in rejecting erroneous inflectional forms. As indicated before in §5.4.1, transformations of the scores into A' or D' scores to correct for a bias towards accepting sentences rather than rejecting them were not performed since this bias was not present in TD subjects and was mainly present in the younger SLI group, as was the case for the judgement of gender. In the older SLI group, the acceptance-rejection rate was more equal.

Table 6.7. Accuracy (%) in the judgement of SVA per group, divided over correct acceptance (hits) and (incorrect acceptance)

	N	Total (SD) (n = 70)	Accept (SD) (n = 30)	Reject (SD) (n = 40)
SLI-Y	32	61.52 (16.84)	83.95 (12.60)	44.69 (29.11)
TD-Y	32	91.52 (11.03)	91.94 (13.05)	91.21 (14.96)
SLI-O	31	87.37 (10.45)	90.11 (12.22)	85.32 (11.76)
TD-O	29	96.21 (4.15)	96.44 (4.54)	96.03 (5.02)
Adults	14	98.47 (1.63)	99.05 (2.42)	98.04 (2.44)

In sum, the hypotheses about the judgement data of subject verb agreement were confirmed. The data showed differences between individuals with SLI and their TD peers in their ability to judge whether subject verb agreement is correct or not. Older participants scored better than younger participants, and the significant interaction showed that the difference between SLI and TD became smaller with age. Although the older SLI seemed to have gained knowledge of subject verb agreement (or gained confidence to make judgements), differences between SLI and TD persisted in adolescence, as expected.

6.4.2 Production

As described in §6.3.3, the production data for subject verb agreement was gathered in two separate tasks: an elicitation game and an elicitation task in which syntactic contexts were varied. In this section, the results of the elicitation game will be discussed first, followed by a discussion of the general results of the elicitation task. The influence of the phonological and syntactic factors that were varied in the two tasks will then be described in the next section (§6.4.3).

Elicitation game

Figure 6.5 shows the accuracy percentage scores in the production of subject verb agreement in target verbs for the different groups. As described before, the boxes and whiskers represent the range of scores, and the bold line in the middle of the boxes represents the median. Three subjects were excluded from the analyses of the elicitation game data because they did not produce enough target verbs (30 target verbs were the minimum). These three children produced zero, or only 4 or 5 target verb inflections. Two were in the younger SLI group and one was a younger TD child. The figure shows that the younger SLI group displayed a large range in accuracy, with some children performing at ceiling and others achieving a low score. The older SLI group performed much better than the younger SLI group, but still made some mistakes.

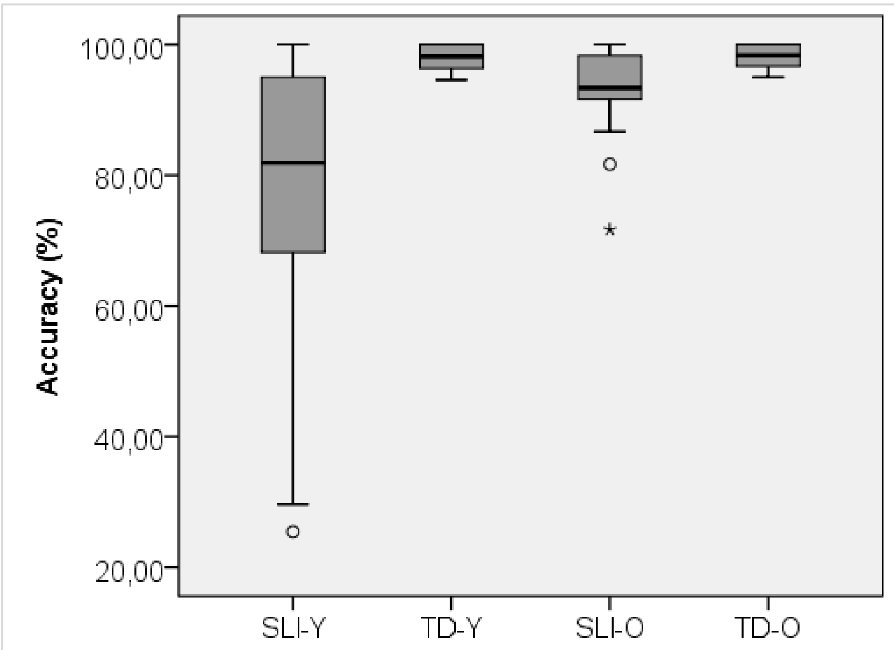


Figure 6.5. Accuracy (%) in the production of SVA per group.[9]

As the figure shows, neither of the two TD groups scored completely at the 100%-level either. However, a closer look at the performance with the different verbs (Table 6.8) indicates that the errors in the TD group were mainly made with one particular verb: the verb *poetsen* 'to polish' (V4). The two items that were removed in the judgement task because an error in inflection was not noticed by adult listeners (**jij poets je tanden* 'you brush your teeth' and **zij poets haar schoen* 'she polishes her shoe') appeared to be vulnerable to *-t* omissions in typical production as well. Apparently, the sequence of the three consonants /t/, /s/ and /t/ is rather difficult in both perception and production. Note that this also influenced the

[9] We did not consider adult data necessary in this task, since subject verb agreement is acquired early in typical development.

transcription of these items. It was sometimes hard to judge whether a child omitted the –*t* suffix or produced it rather quickly. In case of doubt, the experimenter always gave the participant the benefit of the doubt and transcribed and coded the item as correct. Items were therefore only coded as incorrect if errors were clearly identifiable.

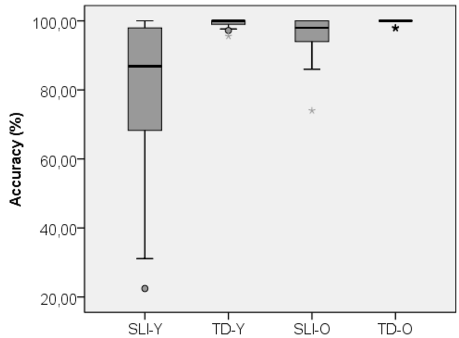
Table 6.8. Accuracy scores (%) in the production of SVA per group, divided over different verbs

	N	V1 *kammen* 'to comb’	V2 *filmen* 'to film'	V3 *lezen* 'to read'	V4 *poetsen* 'to polish'	V5 *bakken* 'to bake'	V6 *drinken* 'to drink'
SLI-Y	30	79.58	79.40	80.77	63.12	83.68	77.55
TD-Y	31	99.64	96.42	98.92	91.86	100.00	98.57
SLI-O	31	97.42	97.74	94.19	80.32	95.16	96.42
TD-O	30	99.63	100.00	100.00	88.67	99.67	99.33

The disproportional difficulty with V4 *poetsen* ‘to polish’ is not due to the complexity of the verb stem coda nor to its sonority since the scores for this verb are worse than the scores for the other verbs with similar characteristics. An additional phonological or articulatory factor seems to play a role. Perhaps, an influential phonological factor is the fact that the consonants in the verb stem *poets* 'polish' do not follow the Sonority Sequencing Principle (Carlisle, 2001, p. 24), which states that 'the sonority of consonants must decrease towards the edges of a syllable' (Booij, 1995, p. 24). Adding an inflectional –*t* suffix to a verb stem that does not follow this principle may be harder. Alternatively, it might be an effect of repetition of the same consonants with another consonant in between (CON_{A} – CON – $\text{CON}_{\text{A})}$, more specifically the repetition of two plosives with a fricative in between (PLOS – FRIC – PLOS) or perhaps even more specifically the repetition of the same plosive (PLOS_{A} – FRIC – PLOS_{A}). What exact factor caused the particular difficulties with this verb cannot be answered in this study, since these particular phonological effects were not controlled for. Because this verb generated deviant outcomes, it was removed from the statistical analyses. Table 6.9 shows the accuracy scores of the different groups without the results of V4. As can be seen, both TD group performed at ceiling once the V4-items were removed. Performance in the SLI group was less affected by excluding the V4-items.

Table 6.9. Accuracy scores (%) in the production of SVA per group (without V4)

	N	Average accuracy % (SD)
SLI-Y	30	80.33 (21.73)
TD-Y	31	99.34 (1.20)
SLI-O	31	96.18 (5.70)
TD-O	30	99.73 (0.72)



ANCOVAs with the number of correct judgements of subject verb agreement as dependent variable, with Group and Age as fixed factors and with IQ as a covariate showed main effects for Group ($F(1,117) = 29.07$, $p < .001$, $\eta_p^2 = .199$) and Age ($F(1,117) = 11.06$, $p = .001$, $\eta_p^2 = .086$). Group and Age interacted significantly ($F(1,117) = 14.60$, $p < .001$, $\eta_p^2 = .111$). Post-hoc analyses revealed that the difference between SLI and TD in the judgement of subject verb agreement was still present in the older group, although much smaller than in the younger groups ($F(1,58) = 7.7$, $p < .007$, $\eta_p^2 = .117$).[10]

Interestingly, an analysis of the accuracy in subject verb agreement in the different contexts of the paradigm (1SG, 2SG, 3SG, 1PL and 2PL) revealed an important qualitative difference between the older and the younger SLI group. While the younger SLI group made mistakes in all contexts, (most frequently in the plural contexts), the older SLI group performed well in plural contexts and only made some mistakes in the singular contexts. The distinction between SLI-Y and TD-Y is therefore independent of the context of the paradigm, while the distinction between SLI-O and TD-O is only visible in the singular domain. Table 6.10 shows the average accuracy scores in the different contexts of the paradigm per group.

[10] With binary logistic regressions, similar effects of Group and Age were found. The interaction between Group and Age was, however, not significant, supposedly due to the fact that the younger TD group already performed at ceiling. Because we hypothesized either absence of an interaction, or (in case of a significant interaction) presence of a difference between SLI and TD in the older groups, the difference between linear and logistic tests is not problematic (in both cases, the hypotheses are confirmed).

Table 6.10. Accuracy scores (%) in the production of SVA distributed over different contexts

	N	1SG	2SG	3SG	1PL	2PL
SLI-Y	30	88.30	80.20	81.76	77.23	72.89
TD-Y	31	99.36	99.68	99.36	99.32	98.55
SLI-O	31	95.16	93.80	94.80	98.39	98.71
TD-O	30	99.67	99.33	100.00	99.67	100.00

If we consider the type of errors made (Table 6.11), the most frequent error was the use of bare forms (*ø) in all contexts where this was visible, i.e. second and third person singular contexts (**jij bak/*hij bak)* and plural contexts (**wij bak/*jullie bak*). Other mistakes consisted of the erroneous use of the third person singular *–t* suffix in first person singular contexts (**ik bakt)* and in plural contexts (**wij bakt, *jullie bakt*). The use of the plural *-en* suffix in singular contexts barely occurred and root infinitives (RIs) were found only 9 times in the younger SLI group (produced by six different children).

Table 6.11. Types of errors in the production of SVA distributed over different contexts per group

		**-Ø*	*2SG*	*3SG*	*1PL*	*2PL*	**-t*	*1SG*	*1PL*	*2PL*	**-en*	*2SG*	*RI*
	N	N	N	N	N	N	N	N	N	N	N	N	N
SLI-Y	30	164	49	47	34	34	94	32	29	27	1	1	9
TD-Y	31	7	1	2	2	2	2	2			0		0
SLI-O	31	38	19	16	2	1	21	15	3	3	0		0
TD-O	30	3	2		1		1	1			0		0

In sum, the production data of the elicitation game showed persistent differences between SLI and TD in the production of subject verb agreement, as expected. Children with SLI performed worse than their TD peers and, although their accuracy levels increased with age, the production of subject verb agreement was still significantly different in adolescence. While younger subjects with SLI mainly made mistakes by using bare forms or *–t* suffixes erroneously in plural and singular contexts, older subjects made relatively few errors in the plural context, but persisted in substituting inflections in singular contexts (the use of a bare form in 2nd and 3rd person singular, and the use of a *–t* suffix in first person singular). Although overall performance was quite high in subjects with SLI, subject verb agreement thus remains an area of difficulty – as was hypothesized.

As was discussed in §6.2, problems in subject verb agreement may not only be found in inflectional errors, but also in the use of alternative strategies. The use of a (dummy) auxiliary + infinitive is such an alternative strategy, since inflection of the lexical verb is 'avoided'. Children with SLI are sometimes reported to use these constructions significantly more often than younger TD children (de Jong, 1999; de Jong et al., 2013). Answers including an auxiliary were excluded from the analyses for accuracy of subject verb agreement discussed above, but were counted separately. Table 6.12 shows how often participants produced inflection of the target verb, in contrast to using auxiliary constructions (both in numbers and in percentages of the total number of elicited items with the correct procedure). Note that these were the auxiliary constructions that continued after several cues to elicit a target construction (§6.2.3). All auxiliary + infinitive constructions are combined in this table (dummy auxiliaries, modal auxiliaries, auxiliaries with a progressive aspect marker).

Table 6.12. Overview of target responses and auxiliary constructions

		Elicited items	Target construction		Auxiliary construction	
	N	N	N	%	N	%
SLI-Y	32	1904	1689	88.71	215	11.29
TD-Y	32	1730	1640	94.80	90	5.20
SLI-O	31	1854	1854	100	0	0
TD-O	30	1790	1790	100	0	0

The table shows that the use of non-target auxiliary constructions despite numerous attempts to elicit a target construction only occurred in the younger groups, and twice as much in the SLI group compared to the TD group (while the TD children received fewer cues to obtain target constructions, because they were tested by research assistants who less often provided an extra elicitation cue). Older subjects did sometimes use an auxiliary, but could be instructed to inflect a target verb. Although the production of an auxiliary construction is an 'avoidance strategy' that is seen in all younger subjects, it thus primarily characterizes the language production of the subjects with SLI. We will return to this point in more detail in the next section (§6.4.3) when we discuss the influence of syntactic context.

Elicitation task with different syntactic contexts

The second elicitation task tested verb inflection in two verb contexts: 3SG and 3PL (the missing context in the elicitation game). The results of the verb *poetsen* 'to polish' were removed on the basis of the same rationale as discussed above when we discussed the outcomes of the elicitation game. Table 6.13 shows the overall performance on subject verb agreement in the elicitation task with different syntactic contexts, and shows the number and type of errors per group. The outcomes are similar to the outcomes of the elicitation game (Table 6.9), as expected. Errors in verb inflections were especially found in the younger SLI group. Although the older SLI group performed better than the younger group, their performance was not yet TD-like. The errors in the younger SLI group mainly consisted of using a *–t* suffix in 3PL contexts and some bare stems in both 3SG and 3PL contexts. Again, the plural marker *–en* was (almost) never overgeneralized to singular contexts.

Table 6.13. Overview of overall performance elicitation task (3SG,3PL) and type of errors per group

		Target items	Accuracy	**ø*	**-t*	**en*
				3S, 3PL	*3PL*	*3SG*
	N	N	%	N	N	N
SLI-Y	32	887	87.03	22	91	2
TD-Y	32	1088	96.63	4	4	
SLI-O	30	1104	96.38	11	29	
TD-O	30	1179	99.91	1		

An ANCOVA with accuracy as dependent variable, Group and Age as fixed factors and IQ as covariate confirmed the similarity in outcomes between the two elicitation tasks for subject verb agreement. Five subjects in the younger SLI group were excluded from the statistical analyses of the elicitation task because they did not produce enough target verbs (< half of the items). There was a significant main effect of Group ($F(1,114) = 24.50$, $p < .001$, $\eta_p^2 = .177$ for group) and Age ($F(1,114) = 7.70$, $p = .006$, $\eta_p^2 = .063$) and a significant interaction effect between the two ($F(1,114) = 7.61$, $p = .007$, $\eta_p^2 = .063$). Although the difference between SLI and TD decreases with age, it is still significant in the older groups ($F(1,57) = 7.02$, $p = .01$, $\eta_p^2 = .110$). As expected, these outcomes are in line with the outcomes of the first task, the elicitation game.

6.4.3 Effects of linguistic context

As was described in §6.2, two phonological factors and one syntactic factor were manipulated in the subject verb agreement production tasks, to investigate the effect of linguistic context on production accuracy. The influence of the phonological factors is evaluated using a subset of the data of the elicitation game (2SG, 3SG). The influence of syntactic structure is evaluated in the elicitation task, that included 3SG and 3PL contexts.

Phonological context

As described in §6.2 and §6.3.3, the phonological factors that were manipulated were the sonority and the complexity of the verb stem coda. Although six verbs were selected with verb codas ending in sonorants, fricatives and plosives, only the sonorant and plosive endings could be analyzed since one of the verbs with a fricative verb stem ending – the verb *poetsen* 'to polish' – appeared to behave differently from the rest and had to be excluded from the analyses (see previous section). Because effects of phonological context are mainly expected for contexts where a *–t* suffix is added to the verb, and not for contexts that require a *–ø* suffix or a plural *-en* ending, the analyses for phonological context were performed on the items for second and third person singular only (2SG and 3SG). Each of the levels (simple versus complex and sonorant versus plosive) therefore contained a subset of 8 items per child (2 verbs per context, elicited twice in both 2SG and 3SG). If children produced less than half of the target verbs in one of the contexts (<4), their data had to be excluded from the analyses for phonological context factors. For the influence of the complexity of the verb stem coda, three participants had to be excluded on this basis: two participants in the younger SLI group and one in the younger TD group. For the influence of sonority, two younger TD participants and one younger SLI participant had to be excluded. Analyses of the effect of phonological factors of the verb stem coda were carried out on the number of correct and incorrect responses. Table 6.14 shows the average percentage of correct production for the different phonological contexts.

Table 6.14. Accuracy (%) in the production of subject verb agreement in different phonological contexts

	N	Simple coda % (SD)	Complex coda % (SD)	N	Sonorant coda % (SD)	Plosive coda % (SD)
SLI-Y	29	84.58 (25.60)	80.69 (27.24)	30	87.52 (24.11)	77.88 (28.16)
TD-Y	30	100.00 (0.00)	98.75 (3.81)	29	100.0 (0.00)	98.71 (3.87)
SLI-O	31	95.45 (11.21)	95.56 (8.87)	31	99.14 (3.35)	91.93 (9.94)
TD-O	30	100.00 (0.00)	99.17 (3.17)	30	100.00 (0.00)	99.17 (3.17)

As can be seen in Table 6.14, almost every group obtained (slightly) higher scores in the context of a simple verb stem coda compared to contexts with complex verb stem codas. Similarly, higher scores were obtained after a sonorant compared to a plosive ending in all groups. The differences are small in the TD groups, and larger in the SLI groups.

Repeated measures were performed with two levels for the phonological factors Complexity and Sonority (simple versus complex and sonorant versus plosive). Age and Group were included as between subject factors and IQ as a covariate. The repeated measures showed no significant effect of Complexity of the verb stem ($F(1,115)= 3.01$, $p = .086$, $\eta_p^2 = .025$). Sonority of the verb stem on the other hand did influence performance significantly ($F(1,115) = 15.11$, $p < 001$, $\eta_p^2 = .116$). There is a significant interaction between Sonority and Group ($F(1,114) = 7.32$, $p = .008$, $\eta_p^2 = .060$), which means that the influence of this phonological factor is different in the two groups. The direction of this interaction is clear from the numbers presented in Table 6.14: the effect is larger in the SLI group than in the TD group. Subjects with SLI are thus more influenced by the sonority of the verb stem coda than TD subjects. An interaction between Sonority and Age was absent ($p = .228$), indicating the influence was not significantly different in the older groups than in the younger groups.

In sum, phonological factors do seem to play a role in correct production of subject verb agreement. The specific difficulties evoked by the verb *poetsen* 'to polish' already indicated that the phonological properties of the verb might influence performance, both in TD and in individuals with SLI. Another phonological factor that influences performance on the production of subject verb agreement, specifically in the SLI groups, is the sonority of the verb stem. For subjects with SLI, suffixation with second or third person singular –*t* is more often correct if the verb stem ends in a sonorant, compared to verb stems that end in a plosive. The complexity of the verb stem on the other hand, does not seem to play an important role. Here, we have to recall the remarks made in §6.3.1 on the complexity of the complex verb stem codas selected in this study. Both the sonorant and the plosive complex verb stem coda (*/fIlm-t/* and */drIŋk-t*) can be altered to simple ones in production by insertion or omission of a phoneme (*/fIllem-t/* and */drIŋ-t/*), without being noticed by listeners. It is therefore questionable whether the absence of an effect of complexity in this study reflects the influence of the complexity of verb stems accurately. However, the fact that performance on subject verb agreement is different depending on the type of verb stem, and the fact that this factor mainly affects performance in SLI is noteworthy and in line with our hypotheses.

Syntactic context

As was described in §6.3.3, the effect of syntactic context was investigated in a separate elicitation task for 3SG and 3PL contexts only. The general outcomes on this task were discussed in the previous section. Here, we will report on the items that contrasted subject verb agreement in main clauses (MC) and in subordinate clauses (SC). Recall that the items of the verb *poetsen* 'to polish' were eliminated because this verb showed deviating scores in all groups in 2nd and 3rd person singular contexts, as we saw in the analyses of the production data of the elicitation game (see previous section). In total, 40 items per participant formed the basis for investigating the effect of syntactic structure (2 clause structures x 2 contexts (3SG/3PL) x 5 verbs x 2 elicitations).

A first interesting influence of syntactic context can be found in the analysis of the missing data (see Table 6.15). The percentage of missing data was higher in the younger groups than in the older groups, and higher in the SLI groups than in the TD groups. It is also very obvious that target responses were predominantly obtained in main clauses, while subordinate clauses triggered far more unanalyzable responses. In comparison to the elicitation game, target answers were harder to elicit in this task because the syntactic context was more constrained. As well as the elimination of auxiliary constructions and verb substitutions, the intended contrast of main and subordinate clauses resulted in a few more situations for data exclusion. Target answers, for instance, had to include an object (or other element) next to a subject and a verb in order to be able to distinguish between main clauses where the verb comes before the object and subordinate clauses which have the verb at the end. Similarly, verb placement in subordinate clauses had to be correct in order to be considered as a subordinate clause. Items without an object, or without a subordinate word order had therefore to be excluded from the analyses.

Table 6.15. Overview of target responses and missing data in two syntactic contexts of the elicitation task per group (MC=main clause; SC=subordinate clause)

	N	Total elicited items		Target construction		Missing	
		MC N	SC N	MC %	SC %	MC %	SC %
SLI-Y	32	640	640	90.94	47.50	9.06	52.50
TD-Y	32	640	640	91.87	78.28	8.13	21.72
SLI-O	30	600	600	100.00	84.00	0.00	16.00
TD-O	30	600	600	99.67	96.83	0.33	3.17

Table 6.16 shows the different types of missing data in the two syntactic contexts per group. As was the case in the elicitation game, younger TD children were tested by assistant experimenters, who less often cued the child to inflect the target verb or produce an object. This caused the higher amount of missing data due to verb substitutions and answers without an object in the younger TD group (in the other groups, these answers did not occur less frequently, but they were more often corrected after repetition of the experimenter cue). Because the younger TD children did not receive that many repeated elicitation cues, caution in interpreting their auxiliary constructions as ‘avoidance’ is required.

This overview of the types of missing data in the different syntactic contexts reveals some interesting facts. Word order alterations in subordinate contexts caused a significant amount of missing data in the SLI groups (and some missing data in the TD groups). Instead of a subordinate word order, with the verb at the end of the sentence, children often produced a main clause word order after the elicitation cue [*hij ziet dat….*] *de man drinkt melk* '[he sees that…] the man drinks milk'. Even when participants were reminded that they had to finish the sentence started by the experimenter, they often came up with a main clause word order in subordinate contexts. The number of word order alternations in this context indicates that this sentence type is rather difficult to construct. This goes against the idea that subordinate clauses provide an easier context because they do not involve movement of the verb to second position of the sentence and follow the Dutch basic word order (see §6.2).

Table 6.16. Overview of the types of missing data in the two syntactic contexts per group (MC=main clause; SC=subordinate clause)[11]

		Aux + INF		Verb substitution		No object		Word order		Other	
		MC	SC	MC	SC	MC	SC	MC	SC	MC	SC
	N	N	N	N	N	N	N	N	N	N	N
SLI-Y	32	39	153	2	14	2	9	4	147	11	31
TD-Y	32	14	36	14	23	23	24	0	35	1	21
SLI-O	30	0	21	0	3	0	6	0	57	0	6
TD-O	30	0	3	0	3	0	1	0	12	2	0

Furthermore, the younger SLI group produced many auxiliary + infinitive constructions, as was the case in the elicitation game. Surprisingly, auxiliary constructions were more often used in subordinate contexts than in main clause contexts. This contrasts with earlier reported findings on the use of dummy auxiliary constructions in TD and SLI. As was discussed in §6.2, previous studies indicated that dummy auxiliary + infinitive constructions are found more often in main clauses than in subordinate clauses in typical development (van Kampen, 1997; Zuckerman, 2001) and in SLI (de Jong et al., 2013). This was interpreted earlier as support for the idea that dummy auxiliary constructions are used to avoid syntactic movement of the finite lexical verb. The outcomes of this study challenge the idea that it is only the avoidance of syntactic movement that triggers the use of (dummy) auxiliary constructions. In our elicitation task, all groups used more auxiliary constructions in the sentence type that did not involve movement. Note that the reported number of auxiliary constructions in this study does not only contain dummy auxiliaries, but also modals and auxiliary constructions with aspect markers (*'is/zijn aan het* INF' or *'zit te* + INF'). A more detailed overview of the different types of auxiliary constructions is therefore provided in Table 6.17.

[11] In the SLI-Y group, there were some answers with auxiliaries in the subordinate clause context that also contained a word order alteration. The sum of the different missing data categories therefore exceeds the 52.5% missing data reported in Table 6.15.

Table 6.17. Different types of auxiliary constructions in the two syntactic contexts per group (MC=main clause; SC=subordinate clause)

		Dummy Aux (*gaan*)		Modals (*wil*)		Aux (*zijn*) + 'aan het'		Aux (*zit)* + 'te'	
	N	MC N	SC N	MC N	SC N	MC N	SC N	MC N	SC N
SLI-Y	32	10	61	0	0	13	81	16	11
TD-Y	32	0	0	3	0	11	36	0	0
SLI-O	30	0	0	0	0	0	19	0	2
TD-O	30	0	0	0	0	0	2	0	1

Although the majority of the auxiliary + infinitive constructions in the SLI groups consisted of auxiliaries + infinitives with a progressive aspect marker ('*te*' + INF, *'aan het'* + INF), the auxiliary constructions of the younger SLI group still contained many dummy auxiliary constructions with the verb *gaan* 'to go' (*de jongen ziet dat de vrouw een boek gaat lezen* 'the boy sees that the woman is going to read a book').[12] The table seems to reveal a developmental pattern in SLI: while the younger SLI group still used dummy auxiliaries (example 10), which is characteristic of an earlier stage in typical development, as was described in §6.2, the older SLI group no longer uses this strategy.

(10) *hij gaat lezen*
'he is reading'
(dummy auxiliary + INF)

(11) *hij is aan het lezen*
'he is reading'
(auxiliary + aan het + INF)

When older subjects with SLI used an auxiliary construction, they used the construction with the progressive aspect marker *'aan het'*, which was also the

[12] Critics could argue that it is hard to tell whether the verb *gaan* ‘to go’ should be interpreted as a dummy (not adding any meaning to the sentence) or, by this age, is intended as the adultlike aspectual marker for future. The fact that there is a difference in the number of these constructions in the two sentence contexts, while both sentence contexts involved the same set of pictures is an argument for the first interpretation (dummies). The same argument holds for the use of the two auxiliary + aspectual marker + INF construction (‘*is/zijn* aan het INF’ & ‘*zit te* INF’), in which the auxiliary construction does not seem to be used to express a specific meaning.

strategy that was sometimes used by the children in the younger TD group (example 11). It thus seems as if there is a reason other than syntactic movement operations for producing (dummy) auxiliaries. We will discuss an explanation for these contrasting outcomes in the conclusion of this chapter (§6.5).

As was just mentioned, the number of target responses varied considerably between syntactic contexts, especially in the SLI groups. If participants missed more than half of the items in one of the two conditions (<10 items per context), their scores were excluded from this analysis. For this reason, 16 children in the younger SLI group, 5 children in the younger TD group and 1 child in the older SLI group were excluded from the repeated measures. Table 6.18 shows the overall accuracy for subjects with sufficient data, the number of target items that were elicited and the mean accuracy scores for the two syntactic contexts per group.

Table 6.18. Overview of accuracy (%) in different syntactic contexts per group (MC=Main clause; SC=Subordinate clause)

Subjects with sufficient data			Target items		Accuracy	Accuracy
		Total accuracy	MC	SC	MC	SC
	N	% (SD)	N	N	N	N
SLI-Y	16	93.05 (9.33)	583	304	96.88 (5.73)	88.06 (16.26)
TD-Y	27	99.55 (1.13)	587	501	99.40 (1.72)	99.67 (1.75)
SLI-O	29	97.07 (3.48)	600	504	97.76 (4.93)	96.21 (5.54)
TD-O	30	99.91 (0.47)	598	581	99.84 (0.91)	100 (0.00)

As can be seen from Table 6.18, the younger SLI subjects showed a large difference in inflection accuracy between main clauses and subordinate clauses, and the older SLI group showed a minor difference. For subjects with SLI, verb inflection appeared to be easier in main clauses than in subordinate clauses. In the two TD groups, the scores in the two syntactic contexts can be considered the same. A repeated measures, with the two syntactic contexts (MC and SC) as dependent variables, with Group and Age as between subject factors and with IQ as covariate showed a significant main effect of the syntactic context (F(1,97) = 10.91, p = .001, η_p^2 = .101). Furthermore, interactions existed between the effect of the syntactic context * Group (F(1,97) = 10.79, p = .001, η_p^2 = .100) and the syntactic context *

Age ($F(1,97) = 7.55$, $p = .007$, $\eta_p^2 = .072$). Finally, a significant three-way interaction was found between the syntactic context and Age * Group ($F(1,97) = 6.46$, $p = .013$, $\eta_p^2 = .062$). The statistical outcomes show that syntactic context influenced performance to a larger extent in the SLI group than in the TD group and to a larger extent in younger children than in older children. The three-way interaction reflects the fact that syntactic context mainly influenced the results in one particular group: the younger SLI group. If we perform post-hoc analyses, the difference between SLI and TD in the effect of syntactic context was not significant in the older groups ($p = .220$). The hypothesis that syntactic context would influence performance in SLI more so than in TD was therefore only confirmed for the younger participants.

In sum, syntactic context does play a role in the production of subject verb agreement. First of all, all children appeared to have higher rates of missing data in subordinate contexts than in main clause contexts, and all participants seemed to apply avoidance strategies (auxiliary constructions) more often in subordinate clauses than in main clauses. As expected, these rates were much higher in the SLI group, mainly in the younger group. The direction of the difference was, however, unexpected. Avoidance strategies were more often applied in subordinate contexts than in main clause contexts. In the target answers, a similar effect of syntactic contexts was shown: errors in verb inflection were found more often in subordinate clause contexts than in main clause contexts. In line with our hypotheses, an effect of syntactic context was thus present, but the direction of the effect was different from what had been found in previous studies on the influence of syntactic context. We will discuss this issue in more detail in the conclusion (§6.5).

6.4.4 Correlations between performance effects and processing abilities

As we saw in the previous section, the phonological context (sonority of the verb stem coda) and syntactic context influenced performance on subject verb agreement especially in the SLI groups. To investigate whether the effect size of these factors in the SLI groups can be linked to limitations in processing capacity, difference scores and correlations with processing measures were computed. Difference scores were calculated by computing the difference between the scores in the complex contexts and the scores in the easy contexts. Correlations were hypothesized to exist between the context effects and the nonword repetition and sentence repetition scores, since these measures showed persistent limitations in SLI (as was discussed in §4.5). Correlations with other measures of processing were, however, also computed. Table 6.19 shows the values of Spearman's rho and the corresponding significant *p*-values.

Table 6.19. Correlations between effects of linguistic context and processing measures (Spearman's rho)

	Sonority (sonorant–plosive)	Synt. complexity (MC-RC)	Synt. complexity (MC-RC) in younger groups
Sentence rep.	*r* = -.27 (*p* = .004)	*r* = -.25 (*p* = .012)	*r* = -.44 (*p* <.001)
Nonword rep.	*r* = -.31 (*p* = .001)	*r* = -.20 (*p* = .045)	*r* = -.50 (*p* <.001)
Digit recall	n.s. (*p* = .191)	*r* = -.22 (*p* = .028)	*r* = -.54 (*p* < .001)
Visual recall	n.s. (*p* = .077)	n.s. (*p* = .091)	*r* = -.30 (*p* <.048)
Inhibition (EF)	n.s. (*p* = .264)	n.s. (*p* = .188)	n.s. (*p* = .689)

After correction of the significance level to .01 (Bonferroni-correction: .05/number of tests), which has to be performed if multiple comparisons are computed, our hypotheses are only confirmed for the phonological context effects. None of the correlations with syntactic complexity is significant. However, we saw in §6.3 that the syntactic factor mainly affected the younger SLI group and had little influence on the scores in the older SLI group. When the younger SLI-group is analysed separately, the correlational analyses reveal the expected pattern: highly significant correlations appear between the verbal processing measures and the influence of syntactic context. The influence of linguistic context factors reported in the previous sections was thus related to processing abilities.

6.5 Conclusion and discussion

In this final section, we will summarize the findings from the previous tasks, discuss whether the hypotheses for subject verb agreement were borne out and whether the findings are in line with the literature discussed in §6.2. We expected to find a difference between SLI and TD, both in the judgement and in the production of subject verb agreement. We furthermore expected that this difference would be larger in younger children, and become smaller as children grow older. The difference between SLI and TD was therefore expected to interact with age – but we still expected to find significant difference between SLI and TD in the older groups. Within the production of subject verb agreement, fluctuations in performance were expected, depending on the linguistic context. These effects were expected to be larger in SLI than in TD. Finally, these fluctuations were hypothesized to be explained by differences in processing ability.

If we summarize the findings discussed in §6.4, most of the hypotheses were borne out. Differences between subjects with SLI and TD participants were found in both the judgement and the production of subject verb agreement. Although accuracy rates were not very low in individuals with SLI, they did make errors in judging or producing verb inflections, regardless of age. In accordance with most of the previous studies on subject verb agreement in SLI (de Jong, 1999; Steenge, 2006; Weerman et al., 2011), production errors mainly consisted of omissions and substitutions of inflectional markers. Substitutions had a clear direction, as reported in previous studies: singular markers were used in plural contexts but the opposite pattern did not occur. Root infinitives were rarely found. This is in accordance with the findings by Steenge (2006) and Orgassa (2009) but in contrast with the findings by Wexler and colleagues (2004). These contrasting findings may be caused by differences in subject ages or in data type (spontaneous versus experimental) but may alternatively stem from the drawbacks in the experimental design of the study by Wexler and colleagues discussed in §6.2.

As hypothesized, the difference between SLI and TD was dependent on age, with larger differences in the younger groups and smaller differences in the older groups. Although adolescents with SLI seem to have gained knowledge of subject verb agreement, as was previously found by Weerman and colleagues (2011), performance was still not age-appropriate in adolescence and differences were thus persistent.

Variability was hypothesized on the basis of linguistic factors manipulated in the production tasks. This was not confirmed for every linguistic factor. Phonological complexity of the verb stem coda did not appear to influence performance significantly, as had previously been found for other languages (Marshall & van der Lely, 2007; Song et al., 2009) and was suggested to be of influence in Dutch as well by Blom and colleagues (2014). Suffixation did not seem to be to be easier for children with SLI if the verb stem consisted of only one consonant, compared to verb stems with two consonants. It is however questionable whether the effect of complexity was properly tested in this study, since some of the complex verb stem codas can be altered to simple ones in production by insertion or omission of a phoneme. The previously found effect of the sonority of the verb stem (Blom et al., 2014) was confirmed in the current study. Suffixation was harder when the verb stem coda ended in a plosive, and this effect was larger in SLI than in TD, as expected. Sonority is, however, not the only phonological factor that plays a role in suffixation. In the current study, a specific verb (*poetsen* 'to polish') had to be excluded from the analyses because it elicited deviant outcomes in all groups. It is, furthermore, hard to categorize verbs in terms of phonological properties, since pronunciation is subject to phonological processes and individual variation. Whether the phonological factors manipulated in this and other studies are purely measuring the intended phonological factors is therefore questionable. Nevertheless is the fact that the sonority of the verb stem coda caused variability in performance in the SLI group noteworthy and in line with our hypothesis.

The hypothesis that syntactic factors would influence performance was confirmed. Accuracy rates differed depending on the syntactic context, but his was found mainly in the younger SLI group, The direction of the effect was, however, different than in previous studies. Weerman and colleagues (2011) found higher accuracy rates in SLI in subordinate contexts than in main clause contexts, while this study found the opposite effect. Similarly, the rate of (dummy) auxiliary constructions was in this study higher in subordinate contexts than in main clause contexts while previous studies found the opposite pattern in typical development (van Kampen, 1997; Zuckerman, 2001) and in SLI (de Jong et al., 2013). In previous studies, the differences between main clauses and subordinate clauses in terms of accuracy and avoidance strategies were often interpreted as an effect of differences in syntactic movement: (dummy) auxiliary constructions were interpreted as avoidance of movement of the finite lexical verb (de Jong et al., 2013) and suffixation would be more difficult in main clauses due to syntactic movement operations (Weerman et al., 2011).[13] This explanation is challenged by the results of the current study.

There were differences between the design of this study and previous test designs, that may be important in explaining the contradictory results. As was described in §6.2 and §6.3, this study used a slightly different task design than was used previous studies. In contrast to the design by Blom and colleagues (2008a), in which the subject of the subordinate (relative) clause was provided in the elicitation cue (see example 12), participants in this study had to construct the subordinate clause as a whole (see example 13).

(12)	Cue:	*dit is de vrouw die…*	Target:	*een zonnetje*$_O$ *tekent*$_V$
		'this is the woman that…'		'draws a sun'

[13] Blom & Baayen (2013) suggest that the higher accuracy of verb inflection in final position in subordinate clauses as opposed to second position in main clauses may also be explained in terms of time constrains.

(13) Cue: *de jongen ziet dat…* Target: *de vrouw*$_S$ *een zonnetje*$_O$ *tekent*$_V$
'the boy sees that…' 'the woman draws a sun'

The difference in findings between this study and previous studies may therefore plausibly stem from differences in the amount of syntactic structure that had to be produced by the participants. Producing the whole dependent sentence might require more processing effort than finishing the dependent sentence in which the subject is already provided. This could explain why the subordinate structure is more often avoided if it has to be constructed from scratch and why more errors are made within the subordinate clauses constructed by the participants themselves. The contrasting findings may thus be explained by the fact that the facilitative effect of the verb in final position is overshadowed by the negative effect of having to construct a more complex sentence. The results of previous studies are therefore not necessarily contradicted but nuanced.

Nevertheless, this study shows that (dummy) auxiliary constructions are not just used to avoid movement of the lexical verb. Alternatively, an explanation in terms of avoidance of inflection seems more plausible. Auxiliaries are often assumed to be stored as chunks instead of morphologically computed during the production process. Next to being more 'economical' in terms of syntactic movement, auxiliary constructions are therefore also less 'costly' in terms of morphological operations. This may explain why they are chosen as alternative options more often in contexts that involve a higher processing load. Which context that is, depends on the situation, as we saw above. The fact that the TD children in this study also seemed to favor auxiliary constructions if they had to construct a syntactically more complex structure (in terms of embedding, not in terms of movement) further supports this idea. This brings us to the final hypothesis that was tested in this thesis: the relation between (variability in) grammatical performance and processing abilities.

The final hypothesis for subject verb agreement, namely that effects of linguistic context would correlate with (verbal) processing measures, was confirmed. Variability in performance dependent on the phonological context in which verbs had to be inflected appeared to be correlated with sentence repetition and nonword repetition. The correlational analyses between the effects of syntactic context and the (verbal) processing measures also revealed strong correlations within the younger SLI group (recall that the syntactic factor affected performance mainly in this group). The idea that fluctuations in performance in SLI are explained by (limited) processing capacities was therefore supported.

In conclusion, the rules for subject verb agreement have generally been mastered by adolescents with SLI, but persistent differences were still found. Adolescents with SLI improved in terms of accuracy and the amount of finite lexical verbs produced (instead of auxiliary constructions, which were often used by younger subjects with SLI), but still made some inflection errors in the singular domain. Errors occurred more often in phonologically and syntactically complex contexts and these influences of linguistic context appeared to be related to (limited) processing abilities. This confirms the idea that grammatical aspects that have been acquired in SLI are still vulnerable to loss in performance when the processing load increases. The theoretical and clinical implications of these findings will be further discussed in Chapter 8.

7 Relative clauses

This thesis aims to answer the question whether the differences in grammatical abilities between SLI and TD that are often reported in childhood remain into adolescence, as was discussed in more detail in §2.5. It also investigates whether performance in certain aspects of grammar is influenced by the type of task and linguistic context, and whether fluctuations in performance can be explained by variation in processing abilities. The previous chapters discussed the grammatical abilities of children and adolescents with respect to grammatical gender and subject verb agreement. This chapter will focus on relative clauses, which is the third and final grammatical aspect to be investigated in this thesis (see §2.4). In contrast to the first two aspects, which have been the topic of several previous studies in Dutch children with SLI, the acquisition of the Dutch relative clause is fairly new.

In general, the use of complex syntax is still in development in adolescence. It becomes progressively more important with age as it is required for an adequate expression of complex thoughts and ideas (Scott & Stokes, 1995; Marinellie, 2004). As discussed in §2.2.1, children with SLI are often reported to have problems with complex syntactic structures (e.g., Schuele & Nicholls, 2000; Novogrodsky & Friedmann, 2006; Hesketh, 2006; Delage, Monjauze, Hamann, & Tuller, 2008). These problems express themselves in a less frequent use of complex structures and in errors in the use of such structures. Both 'avoidance of' and 'errors in' complex clauses seem to persist into adolescence (Marinellie, 2004; van Groningen, 2010; Tuller et al., 2012; Zwitserlood et al., 2015a). The term complex syntactic structures encompasses different types of structures (e.g., complement clauses, passives, *wh*-questions, relative clauses) and problems in SLI are found across the different types. This study focusses on relative clauses, since the international literature on these structures in subjects with SLI is the most detailed.

The first section of this chapter describes the general characteristics of relative clauses and the construction and acquisition of relative clauses in Dutch (§7.1). In §7.2, the literature on relative clause construction in children with SLI is considered and the specific hypotheses for relative clause construction in Dutch subjects with SLI are presented. The tasks used to test comprehension, judgement and production of relative clauses will be described in §7.3, followed by a discussion of the results in §7.4 and the conclusion in §7.5.

7.1 Relative clauses

Relative clauses are constructed differently in different languages. Because syntactic complexity is dependent on a number of other linguistic factors (e.g., word order or case marking) which are also language-dependent, the structure and developmental trajectory of relative clauses vary across languages. In this section, the general characteristics of relative clauses are illustrated (using English as an example) (§7.1.1), followed by a description of the characteristics of the Dutch relative clause (§7.1.2) and a summary of the literature on typical acquisition of (Dutch) relative clauses (§7.1.3).

7.1.1 General characteristics of relative clauses

Relative clauses are subordinate clauses that modify the head of a noun phrase. Generally, relative clauses include a relative marker, which is a pronoun or another element that is co-indexed with the nominal head it modifies (in English *that*, but also *which, where, who, whose* etc.) (Matthews, 2007).[1,2]

The antecedent of the relative pronoun can fulfill different syntactic or thematic roles within the matrix clause, and the relative marker itself can also represent different syntactic or thematic roles within the relative clause (subject, object, adjunct). Relative clauses that modify a head noun can therefore be divided along two dimensions: a. the syntactic function or thematic role of the antecedent in the matrix clause, and b. the syntactic function or thematic role of the relative marker within the relative clause. Examples 1 to 4 illustrate this (Sheldon, 1974).

		MC	**RC**
(1)	The dog$_i$ (S) [that$_i$ (S) ate the cake] bit the girl	S	S
(2)	The girl grabbed the dog$_i$ (O) [that$_i$ (S) ate the cake]	O	S
(3)	The dog$_i$ (S) [(that$_i$) (O) the boy saw] ate the cake	S	O
(4)	The girl grabbed the dog$_i$ (O) [(that$_i$) (O) the boy saw]	O	O

[1] In English object relatives the relative marker is optional and can be left out in production.

[2] Co-indexation is generally shown by elements having the same subscript (NP$_{i,}$ relative marker$_i$).

In the literature, the terms subject relative clause and object relative clause are used in different ways and their use depends on whether the syntactic or thematic role is defined in the main clause (MC) or in the relative clause (RC). Some authors take the role of the antecedent in the main clause as the basis for a division into subject and object relatives (e.g., Sheldon, 1974). In general, however, the syntactic or thematic role of the relative marker in the relative clause is taken as the basis for classification. Sentences in which the relative pronoun takes a subject role (examples 1 and 2) are usually referred to as *subject relatives* (SR), while sentences in which the relative pronoun takes an object role (examples 3 and 4) are referred to as *object relatives* (OR) (e.g., Goodluck & Tavakolian, 1982; Schuele & Nicholls, 2000; Botwinik-Rotem & Friedmann, 2009; Zwitserlood et al., 2015a). The division along the other dimension - the syntactic function or thematic role in the main clause - is then often expressed by the sentential positions of the relative clause. For English, in which the word order in the main clause is SVO and relative clauses are post-nominal, this means a division between *center-embedded relatives* (examples 1 and 3) and *right-branching relatives* (examples 2 and 4). Both aspects seem to play a role in the acquisition of relative structures, but we will focus on the division between subject relatives and object relatives.

Besides co-indexation between the relative marker and the antecedent head noun, relative clauses involve a syntactic process called *wh*-movement. Similar to *wh*-questions, relative clauses with a transitive verb are assumed to be derived from an underlying representation in which constituents are ordered canonically (subject – verb – object for English transitive sentences). The syntactic process *wh*-movement denotes the movement of a *wh*-phrase or relative phrase from its original position to the beginning of the clause (in generative grammar denoted as spec-CP). Although the name of this process refers to question words or relative markers that start with *wh- (which, where, who, whose),* the same process occurs in sentences with other relative markers like *that* (Chomsky, 1977).

In clauses that involve *wh*-movement, the moved element is believed to leave a gap or trace at its original position in the sentence (Botwinik-Rotem & Friedmann, 2009). Online processing studies have shown that the relativized head is reactivated at the position of the gap or trace (Love & Swinney, 1996) and thematic role assignment is hypothesized to take place at this position (Friedmann & Novogrodsky, 2004).[3] Examples 5 and 6 demonstrate the position of the gap or trace

[3] The findings of reactivation at the gap position are based on object relatives in which the position of the gap matches with the point at which roles can be assigned in online processing. In English subject relatives, thematic role assignment cannot take place before the verb has been uttered and

and the movement that takes place. Examples 7 and 8 show that similar movement processes occur in subject and object *wh*-questions.

(5) The girl grabbed the dog$_i$ [that$_i$ __ ate the cake] (subject relative)

(6) The girl grabbed the dog$_i$ [that$_i$ the boy saw __] (object relative)

(7) [Which dog] __ ate the cake (subject *wh*-question)

(8) [Which dog] did the boy see __ (object *wh*-question)

As can be seen in examples 5 and 6, the syntactic or thematic role of the relative pronoun influences the distance over which *wh*-movement takes place and it determines whether movement takes place across another syntactic or thematic role or not. In the English examples, *wh*-movement takes place over a shorter distance in subject relatives than in object relative clauses (i.e. the trace and the relative marker are closer together in example 5 than in example 6). Furthermore, in subject relatives, movement takes place without crossing another syntactic role while in object relatives, movement takes place across the subject of the relative clause.

Cross-linguistically, it has been found that object relatives are harder to process, comprehend and produce than subject relatives (Brown, 1971).[4,5] Several explanations have been put forward to explain the disproportional difficulty of object relatives. Some authors seek the explanation in the properties of subjects and objects in general. Subjects are for instance hypothesized to be cognitively more

---

does therefore, strictly speaking, not take place at the trace position but after the verb, although the thematic role is assigned to the gap position.

[4] In some ergative languages, like Basque, the opposite pattern is found in prenominal relative clauses (i.e. object relatives are easier to process). See Laka (2013) for a discussion and explanation of the cross-linguistic findings.

[5] In Cantonese, an asymmetry between subject and object relatives is not present either. Relatives are prenominal in Cantonese, and subject relatives have a VOS-word order while object relatives have an SVO-word order. Children seem to take the first animate NP as the agent of the sentence, leading both to incorrect object interpretations in subject clauses and incorrect subject interpretations in object relatives (Ooi & Wong, 2014).

accessible (Keenan & Comrie, 1977) and they are relativized more often than objects (Keenan, 1975). The fact that in some languages the relative difficulty of object and subject relatives is absent or even reversed does not fit with this type of explanation (Laka, 2013; Ooi & Wong, 2014). Other explanations involve the structural characteristics of subject and object relatives. Object relatives would require more processing capacity because the constituent order is non-canonical (MacDonald & Christiansen, 2002) or because the distance between gap and antecedent is larger (O'Grady, Lee, & Choo, 2003). Several studies have suggested a relation between relative clause processing and processing abilities like working memory (Avrutin, 2000; Deevy & Leonard, 2004; Lewis, Vasishth, & van Dyke, 2006; Seidel, 2013).[6] However, an explanation in terms of canonicity or distance between antecedent and gap does not seem to suffice either, since not all non-canonical structures pose problems in processing (passives are for instance not that problematic) and in sentences with a similar distance between gap and antecedent, there appear to be a number of other factors modulating processing difficulty (Mak et al., 2006).

When the subject and object differ in number or in animacy, the processing difficulty of object relatives seems to decrease or even disappear (Mak et al., 2002; Traxler, Morris, & Seely, 2002; Mak et al., 2006; Adani et al., 2010). Similarly, whether the subject and object NPs are nouns, proper nouns or pronouns also influences performance (Warren & Gibson, 2002) and participants appear to be helped by clear verb semantics in the processing of object relatives (Mecklinger, Schriefers, Steinhauer, & Friederici, 1995) (see Kidd, Brandt, Lieven and Tomasello (2007) for a review of the factors that have been found to influence the processing of relative clauses). According to Belletti and Rizzi (2013), the difficulty with object relatives is explained by intervention effects in syntactic processing: the greater the similarity between an intervening element and the antecedent, the more difficulty in syntactic processing is hypothesized. Other authors like Kidd and colleagues (2007) argue that the object/subject asymmetry in processing difficulty is in fact explained by distributional and discourse regularities in the input and not by intervention per se. In naturalistic speech, object relatives typically have inanimate heads and often include a subject that is relatively easily accessible, like a proper noun or a pronoun (Fox & Thompson, 1990; Mak et al., 2006). When these two constraints are satisfied (e.g., '*I saw the rock he climbed')*, the asymmetry between subject and object relatives disappears (Mak et al., 2002; Kidd et al., 2007). If, on the other hand,

---

[6] Although some studies report non-significant correlations between the processing of object relatives and working memory (Yoon et al., 2015).

object relatives have inanimate subject and animate objects, the asymmetry appears again (Traxler et al., 2002; Mak et al., 2002). The disproportional difficulty of object relatives is therefore not explained by intervention effects (similarity of items), but by the typicality of the structure. In summary, the processing of a relative clause is based on an interplay between grammatical, semantic and discourse factors.

7.1.2 Dutch relative clause construction

Like English relatives, Dutch relative clauses start with a relative pronoun (*die/dat* 'that'), a relative marker (*wie/wat/welke/waar* 'who/what/which/where') or a preposition + relative marker (*met wie, door wie, waardoor, waarmee* 'with whom, by whom, by which, with which/whom').[7,8] There are a few archaic genitive forms (*wiens/wier* 'whose') but these are hardly used in present speech. The relative pronoun has to agree in grammatical gender with the head it modifies. The pronoun '*die*' is used for common gender heads while '*dat*' is used for neuter gender heads (see §5.1 for a discussion of grammatical gender in Dutch).

In contrast to English, in which the relative pronoun can be left out in object relative clauses, relative markers are obligatory in Dutch. Furthermore, relative clauses entail embedding, which in Dutch also requires a different word order, as was already discussed in §6.1. In embedded clauses, the finite verb is placed at the end (example 7), while in the main clause, the finite verb takes the second position in the sentence (example 8).

(7) *Dat is de man [die (S) een huisje (O) <u>tekent</u> (V)]* [embedded clause: SOV]
that is the man [that (S) a house (O) draws (V)]
'that is the man that draws a house'

(8) *[De man (S) <u>tekent</u> (V) een huisje (O)]* [main clause: SVO]
the man (S) draws (V) a house (O)
'the man draws a house'

Dutch relative clauses are structurally ambiguous for several reasons. Firstly, the word order is the same in subject and object clauses, in contrast to the English examples. Secondly, Dutch lacks a clear case-marking system to indicate which

---

[7] Haendler, Kliegl, & Adani (2015) examined the effect of accessibility of the subject DP and only found a facilitating effect of a subject pronoun in case of a 1st person singular subject and not in case of a 3rd person singular subject.

[8] Recall that whenever Dutch is mentioned, the standard form is being referred to. Much variation exists across different varieties of Dutch in the construction of relatives (Boef, 2008; 2012).

element takes which syntactic or thematic role, and thirdly, it does not allow pronoun retention or full NP repetition in relative clauses.[9] If the subject and object NP of a relative clause have the same features (in terms of animacy and number and being a full NP), and verb semantics does not reveal which NP is likely to have which role, the sentence is therefore also semantically ambiguous. See examples 9 and 10 for two identical syntactic structures with alternative interpretations. Because both the subject and the object NP are singular and animate, both are full NPs, and the verb semantics of *zien* 'to see' does not trigger expectations regarding thematic role assignment, the relative pronoun - in both cases referring to *prinses* ‘princess*’* - can either have a subject interpretation or an object interpretation.

(9) *Dit is de prinses$_i$ [die$_i$ (S) __ de ridder (O) ziet]* (subject interpretation)
This is the princess [that __ the knight sees]
'This is the princess that sees the knight'

(10) *dit is de prinses$_i$ [die$_i$ (O) de ridder (S) __ ziet]* (object interpretation)
This is the princess [that (O the knight (S) __ sees]
'This is the princess that the knight sees'

A subject interpretation is preferred in ambiguous situations (Mak et al., 2002), as is also the case in similar ambiguous relatives in German (Mills, 1985) and Italian (Belletti & Contemori, 2009). An object interpretation can, however, be forced as a result of a number of different characteristics of the sentence.[10] In sentences where subject and object show a difference in number - i.e. one is singular and the other plural - the interpretation of the roles in the relative clause follows from subject verb agreement. As is shown in (11), the plural marking on the verb indicates that the plural NP in the relative clause is the subject. The relative marker must therefore have an object role in contrast to the sentence in (12), which is structurally ambiguous. Note that the object interpretation is only possible after the verb has been processed, thus at the end of the sentence.

(11) *Dit is de prinses die de ridders <u>zie-n</u>* (object relative)
This is the princess.SG that the knight.PL see.PL
'This is the princess that the knights see'

---

[9] Although in some varieties of Dutch, a resumptive pronoun is allowed (Boef, 2008)

[10] See de Hoop & Lamers (2006) for an overview of ways to obtain distinguishability between subject and object.

(12) *Dit is de prinses die de ridder zie-t* (structurally ambiguous)
This is the princess.SG that the knight.SG see.SG
'This is the princess that the knights see'

Another situation in which the object interpretation is forced is when one of the NPs is a pronoun (or both are). In contrast to full NPs, some Dutch pronouns carry overt case marking. Examples 13 and 14 show how the interpretation of a relative clause can be disambiguated by the different forms of the pronoun in nominative or accusative case.

(13) *Dit is de prinses die hem ziet* (subject relative)
This is the princess that he[ACC] sees
'This is the princess that sees him'

(14) *Dit is de prinses die hij ziet* (object relative)
This is the princess that he[NOM] sees
'This is the princess that he sees'

Finally, an object interpretation of a structurally ambiguous relative clause can be required by the semantic or pragmatic characteristics of the arguments (Frazier, 1987; Mecklinger et al., 1995; Mak et al., 2006). Examples 15-18 show situations in which such semantic/pragmatic knowledge has this effect: thieves do not plausibly arrest policemen and ice creams cannot eat girls. Thematic fit with the semantic content of the verb and differences in animacy between subject and object can therefore also force an object interpretation.

(15) *Dit is de politieman die de dief oppakt* (subject relative)
This is the policeman that the thief arrests
'This is the policeman that arrests the thief'

(16) *Dit is de dief die de politieman oppakt* (object relative)
This is the thief that the policeman arrests
'This is the thief that the policeman arrests'

(17) *Dit is het meisje dat het ijsje eet* (subject relative)
This is the girl that the ice cream (inanimate) eats
'This is the girl that eats the ice cream'

(18) *Dit is het ijsje dat het meisje eet* (object relative)
This is the ice cream (inanimate) that the girl eats
'This is the ice cream that the girl eats'

As was discussed in the previous section, cross-linguistic studies indicate that object relatives are often harder to process than subject relatives, although the asymmetry is dependent on the language and the linguistic context. The structural ambiguity of Dutch relative clauses may pose an extra challenge for the comprehension and production of object structures (Sauerland et al., 2015). A corpus study of Dutch newspaper articles revealed that in these written texts subject relatives are in general far more common than object relatives (Mak et al., 2002). However, when only sentences with a transitive verb are considered, object relatives are more frequent (Kaan, 1997). If object relatives are used, the subject is almost always animate and the object inanimate. This matches the idea discussed earlier that object relatives typically have inanimate heads (Kidd et al., 2007).

In an early study of the comprehension of Dutch relative clauses, Frazier (1987) investigated the subject/object asymmetry in a self-paced-reading experiment using 12 ambiguous sentences with two singular animate NPs (*Jan houdt niet van de Amerikaanse die de Nederlander wil uitnodigen* 'John liked not the American who the Dutchperson wants to invite') and 12 unambiguous relatives with a difference in number between the two NPs (*Karl hielp de mijnwerkers die de boswachter vond/vonden* 'Karl helped the mineworkers who the forester found-SG/PL'). The ambiguous sentences and half of the unambiguous sentences were followed by the question: *wie wil wie uitnodigen?* 'who wants to invite who?', in order to check the interpretation of participants. Both the reading times and the answers to the interpretation questions revealed a clear preference for subject structures in Dutch, although the reading times were not significantly different between subject and object sentences (2440 ms. for object sentences, opposed to 2328 ms. for subject sentences). In ambiguous sentences, participants interpreted the head of the relative clause as a subject in 74% of the cases. Furthermore, unambiguous object relatives were incorrectly interpreted as subject relatives in 31% of the cases, while these role reversal errors occurred only 4% of the time in unambiguous subject relatives. With two animate NPs with similar features, participants thus initially assigned a subject analysis to object sentences and sometimes failed to revise this into a correct object analysis (it is unfortunately not clear whether some participants were responsible for all the errors or whether all participants made some errors). The fact that reanalysis failed in some object sentences presumably explains the absence of a significant difference in reading times between subject and object sentences, since reading times are only expected to be longer in the case of a revision (Frazier, 1987).

A more recent study on subject- and object *wh*-questions in adults supported the idea of an initial subject interpretation analysis and subsequent reanalysis into an object interpretation in sentences with two animate referents and a difference in number (*welke boerin heeft de prinsessen gewassen?* 'which farmer washed the princesses?' / *welke boerin hebben de prinsessen gewassen?* 'which farmer did the princesses wash?'). Eye-tracking data in a picture selection task showed that adults look at the picture with a subject interpretation before switching to the picture with an object interpretation when hearing object questions (Schouwenaars, 2012; Schouwenaars, van Hout, & Hendriks, 2014). Reanalysis started a little while after hearing the verb (that carried the disambiguation cue and comes in second position in *wh*-questions) and the second NP. Perceivers thus seem to be biased towards a subject reading of the sentence if both NPs are animate and only differ in number. Topicality of the antecedent or second NP did not correct this bias (Strangmann, Slomp, & van Hout, 2014).

However, in another study into the processing of Dutch center-embedded subject and object relatives in adults, the idea of an initial subject interpretation in the online processing of relative clauses was nuanced (Mak, 2001; Mak et al., 2002; 2006). As discussed before, Mak and colleagues (2002) did find a subject/object asymmetry in reading times of relative clauses with animate heads, but found that the difference in reading time disappeared when the object was inanimate (*vanwege het onderzoek moeten de inbrekers, die de computer gestolen hebben, nog een tijdje op het bureau blijven* 'because of the investigation, the burglars, who stole the computer, had to stay at the police station for some time' / *vanwege het onderzoek moet de computer, die de inbrekers gestolen hebben, nog een tijdje op het bureau blijven* 'because of the investigation, the computer, that the burglars stole, had to remain at the police station for some time'). In sentences where the antecedent of the relative marker is inanimate, participants thus do not seem to be biased towards a subject interpretation (Mak et al., 2002).

A subsequent study tested whether this result could be explained by participants having a bias towards a subject interpretation with animate antecedents and an object interpretation with inanimate antecedents of the relative marker. This hypothesis was not borne out since in those cases where both subject and object were inanimate, subject relatives had shorter reading times again (*volgens de folder moet de gel, die de lekkages verhelpt, in een keer werken* 'according to the brochure the gel, that repairs the leak, should work at once' / *volgens de folder moeten de* *lekkages, die de gel verhelpt, in een keer verdwenen zijn* 'according to the brochure the leak, that the gel repairs, should disappear at once').

The difference in animacy between subject and object thus seems to be important in processing relative clauses. When object relatives have typical characteristics (i.e. an animate subject and an inanimate object), they can be processed as fast as subject relatives (Mak et al., 2002; 2006). The fact that the asymmetry disappears if subject, object and the verb have certain characteristics challenges the idea that a subject interpretation is constructed from the start. Apparently, perceivers take into account the characteristics of the antecedent of the relative marker and the NP within the relative clause, before constructing a syntactic analysis. At which point expectations regarding the syntactic structure are made or revised remains open to debate (Mak et al., 2006) and falls beyond the scope of this thesis. For this thesis, it is important to realize that (problems in) the processing of relative clauses seems to be based on an interplay between grammatical, semantic and discourse factors. The subject/object asymmetry that is often reported in the literature might be an artefact of the way object relatives are tested in most studies. We will come back to this point in the conclusion and discussion of this chapter (§7.5).

In the production of Dutch relatives, a good alternative to the object relative clause is the passive construction. Compare examples 19 and 20, in which the animate referent ‘prinses’ *princess* is relativized by using an active object relative (example 19) or a relative with a passive voice (example 20). Using a passive circumvents the difficulty of having an animate object relativized since the passive voice turns the patient of the sentence into subject. The agent is optionally expressed by a prepositional phrase that precedes or follows the verb.

(19) *Dit is de prinses die de ridders zien*
'This is the princess that the knights see'

(20) *Dit is de prinses die gezien wordt (door de ridders)*
‘This is the princess that is seen (by the knights)’

If adults need to add information about an animate object, they often choose to produce a passive instead of a relative clause as was shown in a pilot study eliciting object relatives (Duinmeijer, 2011). In Italian, which shows the same structural ambiguity in relative clauses, the avoidance strategy of using a passive in object contexts is adopted by adults almost without exception (Belletti & Contemori, 2009). Similarly, the use of passives in animate object contexts was also reported by

Schouwenaars (2012) in the elicitation of wh-questions. Adults thus seem to avoid the construction of object structures – at least with animate antecedents. This is in line with the findings discussed earlier that, in Dutch written texts, object relatives occur almost exclusively with inanimate referents (Mak et al., 2002).

In sum, the literature on the processing and production of relative clauses in Dutch shows that object relatives are harder to process than subject relatives and are sometimes incorrectly processed as subject sentences, even in the presence of clear cues like subject verb agreement, case marking or thematic fit. With a difference in animacy between the two NPs (i.e. the subject being animate and the object being inanimate), the subject bias disappears. In production, this is also the most felicitous context for an object relative to occur. With animate objects, adults tend to avoid the object relativization by using a passive construction. The next section will briefly discuss how and at which age children acquire this particularly difficult aspect of Dutch in typical development.

7.1.3 Acquisition of relative clauses (typical development)

A number of studies across languages report that the use of relatives starts early in childhood, but the acquisition of relatives shows a gradual development over a long period of time (Hamburger & Crain, 1982). Several studies indicate that (TD) children start using relative clauses in spontaneous speech by age three, and are able to imitate or produce them in repetition or elicitation tasks around age four (e.g., Sheldon, 1974 for English; Schuele & Nicholls, 2000; Rothweiler, 1993; Thiel, Sanfelici, Koch, & Schulz, 2014 for German; Crain, McKee, & Emiliani, 1990 for Italian; Friedmann & Novogrodsky, 2004 for Hebrew). Relatives are, however, not fully acquired before age six (Roth, 1984 in Laka, 2013). Before relatives are acquired, children try to combine information using prepositional phrases, juxtaposition/coordination or reduced relatives as substitutes for the relative clause construction (Schuele & Nicholls, 2000), especially in contexts where the experimenter wants to elicit an object relative (Thiel et al., 2014).

Cross-linguistic findings also indicate that the different types of relative clauses are acquired at different ages. As mentioned before, right-branching relatives are easier to understand and acquired earlier than center-embedded relatives (e.g., Botwinik-Rotem & Friedmann, 2009). Furthermore, it has been attested that subject relatives are easier and acquired earlier than object relatives, both in comprehension and in production (e.g., Stewart & Sinclair, 1975; de Vincenzi, Adruino, Ciccarelli, & Job, 1999; Kidd et al., 2007; Friedmann, Belletti, & Rizzi, 2009; Thiel et al., 2014).[11] The asymmetry in acquisition pace between subject and object relatives extends to other types of sentences that involve *wh*-movement, like *wh*-questions (Tyack & Ingram, 1976; Avrutin, 2000; van der Meer, van Atteveldt, Coopmans, & Philip, 2001; Seidl, Hollich, & Jusczyk, 2003).

Little research has been conducted on the developmental trajectory of the Dutch relative clause. However the research done on the acquisition of Dutch *wh*-questions can be helpful in forming expectations about the results with relative clauses. Van Kampen (1997) described the development of the production of *wh*-questions in three children, and distinguishes three stages: in the first stage, children produce *wh*-questions without a *wh*-pronoun, with the verb slot restricted to *doen* 'to do' or *zijn* 'to be' (**doet beer nou?* 'does bear do?'). In the second stage, the set of verb forms used in *wh*-questions expands and the first instances of *what* appear. In the third stage, the *wh*-pronouns *what* and *where* occur with any finite verb, although they are still treated as optional (they are sometimes left out). Although *wh*-questions and relative clauses are similar in terms of the syntactic process *wh*-movement, they also differ in a number of respects (position of the verb, number of elements). The acquisition of *wh*-questions might therefore not be representative of the development of relative clauses. The stages might, however, be roughly similar (leaving out the relative marker, starting with a small set of verbs and expanding it to a larger set). According to van Kampen, Dutch *wh*-questions already appear quite early in language production, but the acquisition process requires different steps and is not completed before age seven.

A number of studies investigated the subject/object asymmetry in Dutch children's comprehension and production of *wh*-questions, although none of them took the effects of animacy and accessibility into account. Van der Meer and colleagues (2001) investigated the interpretation of ambiguous who-questions with an 'X tickles Y and Y tickles Z'-design. Children were asked *wie zei je dat Y kietelde?* 'who did you say Y was tickling/was tickling Y?'. Participants' answers indicated whether they interpreted this ambiguous question as a subject or an object question (the order of the events was counterbalanced). The results showed a clear subject/object asymmetry in these contexts. Children between four and six years old most often gave a subject answer, as would be the expected pattern in adults.

As mentioned earlier in §7.1.1, a study by Schouwenaars (2012) showed that Dutch adults are able to use a subject verb agreement cue to analyse a *which*-question as an object question. The same study showed that children between six

---

[11] Although it would be interesting to investigate the acquisition of relative clauses in languages like Basque or Cantonese, that show a reversed asymmetry between subject and object relatives (see §7.1.1).

and eight do not yet use this cue consistently and prefer a subject interpretation in most of the cases (84%) (Metz, van Hout, & van der Lely, 2010; Schouwenaars, 2012). For those object items that were correctly assigned an object interpretation, reaction times were much longer than for correctly assigned subject interpretations, indicating reanalysis took place. TD children thus show difficulty analysing a sentence as an object relative, both in their error patterns and in their reaction times.

Strangmann and colleagues (2014) tested whether an extra cue in terms of topicality would help children interpreting questions as object *wh*-questions. It appeared that, regardless of the agreement cue on the verb and the topicality of the antecedent, 5-year-olds interpret all question types as subject questions. Between 5 and 8 years of age, Dutch TD children have a strong bias for subject-first interpretations in *wh*-questions, and most of them are not yet able to use syntactic or pragmatic cues to reanalyse the sentence as an object relative (Strangmann et al., 2014). At age nine, children perform somewhat better, but not yet adult-like (Metz et al., 2010). It has not yet been established at which age Dutch children fully develop these capacities. Furthermore, it has not yet been tested whether children in this age range are able to interpret object structures in contexts where they naturally occur: with accessible subjects and inanimate objects.

As described in the previous section, Dutch adults tend to avoid object constructions in production by using a passive construction. Another question is, therefore, at what point children start using this avoidance strategy in object contexts. Schouwenaars (2012) elicited the production of wh-questions in children between six and eight years old and found the use of passives to be a clear strategy, although not yet to the same extent as in adults (Schouwenaars, 2012). According to Belletti and Contemori (2009), who studied the same avoidance strategy in Italian object relative clauses, children develop this strategy from five years onwards. Evidently, the use of the avoidance strategy also depends on the acquisitional path of the passive itself. Dutch children start producing passives between two and three years, although they start with short passives that often carry different morphology than adult passives. Full passives are not seen in spontaneous speech before age five (Verrips, 1996). A large cross-linguistic investigation of the acquisition of the passive showed considerable cross-linguistic variation in the development of the comprehension of full passive constructions. In Dutch, the passive construction was not fully acquired by age five, although accuracy levels in comprehension were already around 80% (Armon-Lotem et al., 2016). It remains undetermined at which age Dutch children will start using the passive as an avoidance strategy for object relatives, and at which point their use of the strategy will be adult-like.

In sum, relatively little is known about the acquisition of the Dutch relative clause in TD children. What can be postulated is that, like other complex syntactic structures such as *wh*-questions and relatives in other languages, Dutch relatives are acquired relatively late and show a gradual developmental path with some relative types being acquired earlier than others. Subject relatives are, for instance, acquired before object relatives. It remains undetermined at what age children are able to create an object analysis (with the help of syntactic, semantic or pragmatic cues) and at what age children will develop the strategy to avoid these structures in production. Up to eight years of age, children are generally not yet able to interpret object relative structures, although this has not yet been tested with object structures with an inanimate object and an accessible subject.

It could be argued that the ability to comprehend object structures is not very useful in Dutch since these structures rarely occur, are difficult to comprehend for adults and TD children, and tend to be avoided in ambiguous contexts. The following anecdote from a clinical situation the author experienced does, however, illustrate the occurrence of object relatives in a natural situation and also nicely illustrates what happens if the cues for an object relative interpretation are not picked up by the listener.

> *One day, a clinician tried to test fluency (retrieving words from the lexicon) in a child with a language impairment. She first asked the child to name as many words starting with a /p/ to test his ability to retrieve words from the lexicon based on phonological properties. Then she wanted to test how well the child could retrieve words from the lexicon based on semantic properties, so she asked the child: 'Noem dingen op die mensen eten' (which means either 'name things that people eat' or 'name things that eat people', although she intended the first object interpretation). She gave some examples (pizza, bread) to help the child start the task. To her surprise, the child started naming things that could eat people (tigers, crocodiles etc.) instead of things that people could eat (May 2014, Royal Dutch Kentalis).*

The example also shows that, in comparison to TD children, children with a language impairment may have extra difficulty in the comprehension or production of an object relative. Despite clear cues in the sentence and context (the object 'things' is inanimate and the clinician gives examples of things that are clearly 'eaten' by people instead of the other way around), the child interprets the sentence as a subject sentence. In the next section, the literature on problems with relative clauses in SLI is evaluated and hypotheses regarding performance on Dutch relative clause tasks are postulated.

7.2 Relative clauses in SLI: literature review and hypotheses

The gradual and relatively late acquisition of Dutch relative clauses in typical development suggests that these structures will be particularly problematic in SLI. Subordinate clauses/complex clauses in general have been shown to be problematic for this group (van der Lely, 1998; Marinellie, 2004; Hesketh, 2006; Tuller et al., 2012). Problems can be identified when there is either less frequent use of complex structures in spontaneous speech and elicitation tasks (Marinellie, 2004; Hesketh, 2006; Tuller et al., 2012) or when there are errors in the use of such structures. These two factors, i.e. complexity and grammaticality, seem to be negatively correlated. Children with SLI who use complex structures make more grammatical errors than children who use simpler sentence structures (van Groningen, 2010).

In the literature on SLI, problems in the comprehension and production of relative clauses have been widely attested in different languages (Hebrew: Friedmann & Novogrodsky, 2004; 2007; 2011; Novogrodsky & Friedmann, 2006; Italian: Contemori & Garaffa, 2010; Swedish: Håkansson & Hansson, 2000, English: Deevy & Leonard, 2004; van der Lely, 1998; Schuele & Nicholls, 2000; Schuele & Tolbert, 2001; Marinelli, 2004; Hestvik, Schwartz, & Tornyova, 2010; Greek: Stavrakaki, 2001; French: Jakubowicz, 2011; Tuller et al., 2012; Danish: Jensen de Lopéz, Sundahl Olsen, & Chondrogianni, 2012; Dutch: Zwitserlood et al., 2015a).

In production, studies using elicitation tasks report that relatives are rarely produced by young children with SLI between age four and six (Contemori & Garaffa, 2010). Over time, relatives can be elicited more often in SLI, but performance is not comparable to (language-matched) TD peers (Novogrodsky & Friedmann, 2006; Jensen de Lopéz et al., 2012). Especially object relatives are reported to be problematic for individuals with SLI. In object contexts, children with SLI produce more role reversal errors while TD children produce more passives (Jensen de Lopéz et al., 2012). Studies examining spontaneous speech also report that relatives are used less frequently by children and adolescents with SLI. Even in comparison to much younger TD peers (aged six, eight and eleven years) adolescents with SLI produce fewer relatives (Tuller et al., 2012).

In general, relatives are thus produced less frequently by individuals with SLI, both in elicitation and in spontaneous speech. Furthermore, their relatives contain more and different grammatical errors. The relative marker is, for instance, sometimes omitted by children with SLI (‘*point to the truck Ø knocked Bert over’*), while this error is not seen in typical development (Håkansson & Hansson, 2000; Schuele & Nicholls, 2000; Schuele & Tolbert, 2001). However, not all studies find these omission errors, presumably due to differences in the language studied or to the age of the subjects in the study (Hesketh, 2006; Novogrodsky & Friedmann, 2006).

Children with SLI also have difficulties comprehending relative clauses, especially when the structure is an object relative (Friedmann & Novogrodsky, 2004; Jensen de Lopéz et al., 2012). This difficulty with object relatives is much larger than attested in TD children (Jensen de Lopéz et al., 2012) and persists into adolescence (Friedmann & Novogrodsky, 2004). Adani and colleagues (2014) found that the comprehension of object relatives in children with SLI between nine and twelve was facilitated when there were discrepancies in number. Gender also helped, but to a lesser extent (Adani et al., 2010). Although most studies report problems in the comprehension of object relatives in SLI, the comprehension of subject relatives is not unproblematic either (Stavrakaki, 2001; Friedmann & Novogrodsky, 2004; Zwitserlood et al., 2015a). Stavrakaki (2001), for instance, reports that children with SLI between five and nine years perform at a lower level than language-matched TD peers (between three and five years) on subject relatives in an act-out task. Similarly, children with SLI between seven and eleven years of age were found to perform at the same level as their six-year-old TD peers in the comprehension of subject relatives in a picture-selection task (Friedmann & Novogrodsky, 2004).

There is only one study on relative clauses in Dutch children with SLI, that is the aforementioned study by Zwitserlood and colleagues (2015a). This study tested the effect of a metalinguistic and multimodal intervention program for subject relative clauses in a sample of 13 children with SLI between 6 and 10 years of age. Although the study does not have a control group, the children with SLI clearly had problems comprehending and producing subject relatives. Five hours of intervention appeared to have a significant effect on the production of subject relative clauses, but not on comprehension (Zwitserlood et al., 2015a).

In sum, cross-linguistic findings indicate problems in SLI with relative clause comprehension and production in general and with object relatives in particular. In comprehension, there is a strong subject bias (relatives are interpreted as subject relatives) and in production, relatives are less frequent and contain more errors. Although object relatives are harder than subject relatives, subject relatives are not unproblematic either. On the basis of the literature on SLI and the specific characteristics of Dutch, several hypotheses can be formed (see Figure 7.1 for a schematic illustration of the hypotheses for relative clauses).

HYPOTHESES RELATIVE CLAUSES

* Persistence grammatical problems (§ 7.4.1 - § 7.4.3)

comprehension
judgement
production } Effect Group & Age (SLI<TD) – effect Group still present in adolescence

* Difference between subject and object relatives in all tasks (§ 7.4.1 - § 7.4.3)

effect size larger in SLI than in TD

* Relation difference in contexts and processing load (§ 7.4.4)

correlations between effect of linguistic context and processing measures

Figure 7.1. Schematic illustration of the hypotheses for relative clauses

Similar to the hypotheses for grammatical gender and subject verb agreement in the previous chapters, we expect to find persistent differences between SLI and TD in the comprehension and production of relative clauses. Because complex structures are still in development between age six and ten, older subjects (between twelve and sixteen) are hypothesized to perform better than younger subjects. Main effects for Group and Age are therefore expected for the total accuracy scores on the different tasks testing the comprehension and production of relative clauses. Because problems with relative clauses have already been reported to be persistent in SLI in other languages, no interaction between Group and Age is expected (the difference between SLI and TD is not expected to become smaller in adolescence) .

We also expect to find variability in performance on the basis of the type of relative clause. More specifically, object sentences (OR) are predicted to yield lower scores than subject relatives (SR). This difference is expected to be larger in the SLI group since object sentences are reported to be particularly difficult for children with SLI, both in interpretation and in production. We thus expect an interaction between linguistic context (subject or object relative) and Group (SLI or TD) in the comprehension and production of relative clauses, with larger effect sizes in the SLI group.

We furthermore expect to find a (differential) effect of the type of cue in object contexts (subject verb agreement cues, animacy cues and verb semantics cues) on the comprehension of relative clauses. On the basis of previous studies using animacy cues, object relatives with an inanimate object are hypothesized to be easier to process than object relatives with an animate object. Furthermore, within object relatives with an animate object, clauses with clear semantic cues regarding thematic role assignment (due to semantic/pragmatic properties of the verb and the arguments) are expected to be interpreted more easily than object relatives with reversible thematic roles. The effects of these factors are hypothesized to interact with group membership: children with SLI are hypothesized to profit less from these cues than TD children.

Finally, because several studies have suggested or shown relationships between problems in complex syntax and information processing measures (as was discussed in §7.1.1), we expect to find correlations between the effects of the linguistic contexts in our tasks and the processing measures described in Chapter 4.

7.3 Task descriptions

To test knowledge and production of relative clause structures, three different tasks were designed. A comprehension and a production task were constructed following earlier studies in other languages. A third task – a judgement task – was additionally created to test the effect of animacy and semantic cues on the interpretation of Duch object relatives.

7.3.1 Comprehension task relative clauses (RCC)

To test the knowledge of relative clause structures in comprehension, children were auditorily presented with relative clauses and had to choose which of two pictures matched the relative clause the best. One of the pictures matched the sentence, and the other picture showed the same action but with reversed roles (e.g., a knight catching a princess and a princess catching a knight). The design of this test resembled a comprehension task for relative clauses designed by Friedmann and Novogrodsky for Hebrew (2004). Their task has some advantages over other types of comprehension tasks. Representing the same action with reversed thematic roles makes the restrictive relative clause 'this is the knight that… –' felicitous, since there are two knights, one performing the action and one not performing the action. Furthermore, in contrast to act-out-tasks, in which the elicitation cue often follows the action, the pictures and sentences are presented simultaneously in this picture-selection task.

The relative clause comprehension task (RCC-task) designed for this study contained four types of sentences. First of all, relative clauses with two animate singular full NPs were included, since these sentences have been tested cross-linguistically. However, as described in §7.1.2, Dutch relative clauses are ambiguous when subject and object are full NPs and have the same linguistic features, although a subject interpretation is the default interpretation. Ambiguous relative clauses will therefore be referred to as *SR-ambiguous* (see example 21). Because object relatives tend to be avoided in production by using a passive construction, subject relatives with a passive were added to test knowledge of the passive (*SR-pass*, see example 22). Figure 7.2 illustrates a picture pair used in the RCC-task for SR-ambiguous and SR-pass items. For both categories, six items were elicited. Participants heard the sentence and had to choose the corresponding picture.

(21) SR-ambiguous *dit is de prinses die de ridder vangt*
'this is the princess that catches the knight'

(22) SR-pass *dit is de prinses die wordt gevangen door de ridder*
'this is the princess that is being caught by the knight'

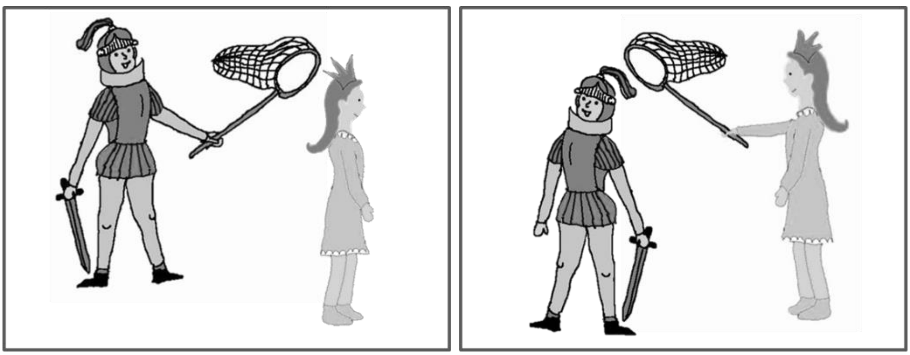


Figure 7.2. Example of a picture pair used for the SR-ambiguous and SR-pass items in the RCC-task[12]

As was also described in §7.1.2, an object interpretation can be forced in Dutch in different ways. By making the subject and object differ in number and using a verb indicating either singular or plural, participants are put in the position of being able to deduce which element in the sentence is the subject and which element is the object. To test the comprehension of such relative clauses, sentences with a difference in number between subject and objects were included (denoted as *SR-PL* and *OR-PL*, see examples 23 and 24). Because the hypotheses are mainly focussed on the distinction between subject and object relatives, twice as many items (12) were included for these sentence types. Figure 7.3 illustrates an example of a picture pair used for the elicitation of SR-PL and OR-PL items.

[12] Pictures for this task were drawn by the author.

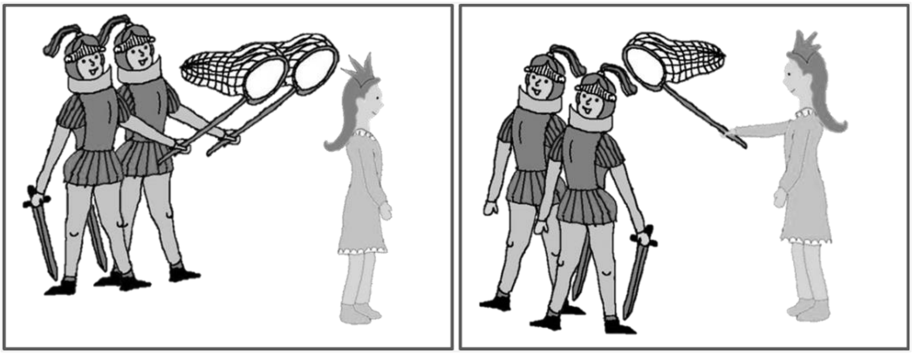


Figure 7.3. Example of a picture pair used for SR-PL and OR-PL items in the RCC-task

(23) SR-PL *dit is de prinses die de ridders vangt*
'this is the princess that catches the knights

(24) OR-PL *dit is de prinses die de ridders vangen*
'this is the princess that the knights catch'

Note that in order to perform well on the comprehension of OR-PL sentences, subject verb agreement rules have to be in place and the child has to process the inflection. If children heavily rely on canonical word order, do not notice the inflectional difference and/or do not know the rules for subject verb agreement, there is a chance they will always choose a subject interpretation in the comprehension task, regardless of whether the verb form signals a singular or plural subject.

The children were presented with 36 sentences in total (6 SR-ambiguous, 6 SR-pass, 12 SR-PL and 12 OR-PL) (see Appendix G for the total list of items). The order of the items was randomized. The task started with two simple sentences to familiarize the child with the task. If children chose the wrong picture, feedback was given by simply repeating the sentence and asking the child whether the sentence really matched the picture pointed at. After the practice trials, no feedback on picture choice was provided. Administration of the task took 5-10 minutes. Answers were coded as either wrong (0) or right (1).

7.3.2 Judgment task Relative Clauses (RCJ)

In the RCC-task that was described in the previous section, a number distinction between subject and object was used to test the comprehension of object relative clauses. Object interpretations could be derived via interpretation of the inflection of the verb, which agrees with the subject. Participants thus had to interpret the verb inflection correctly in order to obtain an object interpretation of the sentence. As was already noted in the previous section, failure to pick the picture denoting an object interpretation could therefore mean a problem with object relative structures, but could also reflect problems in subject verb agreement. Because subject verb agreement is a vulnerable area in SLI, as was described in §6.2, the verb inflection cue might not be entirely accessible to participants with SLI. Furthermore, the fact that the cue was at the end of the sentence meant that participants had to process the whole sentence in order to hear the cue. Because children with SLI also have problems in information processing, as became clear in §4.4, this might be an extra difficulty in the RCC-task (these possibilities will be discussed in more detail when we present the results).

Another perception task was therefore constructed in which the object interpretation was forced in a different way and in which comprehension was tested using a judgement design. In the judgement task for relative clauses, participants were shown a picture and were presented auditorily with a sentence. They had to judge whether the sentence could match the picture or not. The restrictive relative clauses ('this is the x that…') were made felicitous by adding arrows to the relativized element (see Figure 7.4).

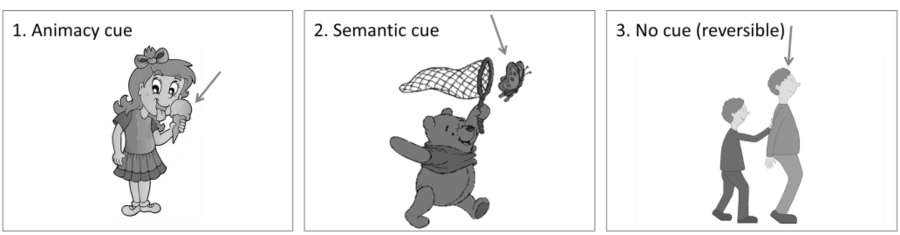


Figure 7.4. Examples of items for the three different types of object relatives in the RCJ-task, see sentences 25-27)[13]

The sentences were either simple sentences (correct or incorrect), subject relatives, or object relatives. Within the object relatives, three types of sentences were included, as is shown in examples 25-27, and depicted in Figure 7.4. The first type of object relative had an inanimate object (example 25, Figure 7.4, picture 1), which

[13] The pictures for the items with an animacy cue and a semantic cue were found on the internet. For the reversible items without a cue, pictures were developed by Sofia Miliopoulos for the COST-project IS0804 (Language Impairment in a Multilingual Society – Linguistic patterns and the road to assessment) (de Jong, 2015).

facilitates an object interpretation since a subject interpretation is not plausible due to *animacy* constraints (if one NP is animate and the other one is inanimate, the inanimate one is likely to be the object) and thematic fit with the verb (ice creams cannot lick).

(25) OR-animacy cue *dit is het ijsje dat het meisje likt*
'this is the ice cream that the girl licks'

In the second type of object relatives, both subject and object NPs were animate, but thematic roles could be deduced by semantic inference, since verbs and nouns were chosen that semantically or pragmatically triggered an expectation concerning thematic role assignment. Example 26 shows such an item with a bear and a butterfly and the verb *'to catch'* (Figure 7.4, picture 2). Although in terms of animacy, bear and butterfly are similar – when it comes to the action of catching, a bear is more likely to catch a butterfly than vice versa.

(26) OR-semantic cue *dit is de vlinder die de beer vangt*
'this is the butterfly that the bear catches'

Finally, some object relatives were fully reversible, as is shown in example 27 (Figure 7.4, picture 3).

(27) OR-no cue (reversible) *dit is de man die de jongen duwt*
'this is the man that the boy pushes'

The relative clause judgement task (*RCJ-task*) specifically aims to test whether participants will accept the object construction in relative clauses if they are forced to make a choice as to whether the sentence can be true or not. If the children do not have the object construction available in their grammar, they will reject the OR. Furthermore, we expect differences between the types of sentences, with the highest acceptance of an object interpretation in items with an animacy cue, followed by items with a semantic cue. Items without any cue for an object interpretation (reversible sentences) are expected to be the hardest to accept.

Piloting with TD children between six and ten years old showed that the reversible object sentences influenced judgement of the items with an animacy cue and semantic cue. Once children had been presented with a reversible item, this seemed to strengthen their reliance on canonical word order. While object sentences with a clear animacy cue or semantic cue were accepted in the first place, they were not after having been presented with a reversible object relative. In the final version of the judgement task, the object relatives with an animacy cue were therefore introduced first, followed by the object relatives with a semantic cue. The reversible object relatives were tested last. In this way, children were prevented from developing a strategy of reliance on word order. The items with an object relative were mixed in with the simple sentences and the subject relatives.

A clear disadvantage of the judgment design is that you can only deduce something about the knowledge of participants when they reject certain structures. If children consistently reject object structures, they have clearly not yet learned this structure. However, if they accept a sentence, it is not clear whether they really have knowledge of the structure or accept the sentence on another basis. They could, for instance, just listen to the lexical items of the sentence and check whether they see these lexical items in the picture, without paying attention to the sentence structure. These drawbacks were circumvented by adding simple sentences with the same lexical items, but with an incorrect ordering of arguments in combination with the picture:

(28)	Incorrect simple sentence	*het huis schildert de man*
		the house paints the man
		'the house is painting the man'

In this way, it was possible to test whether children paid attention to the structure of the sentence. If children did not reject these simple sentences (if they rejected less than four out of six incorrect items), their scores on the judgment task were excluded from the analyses. Eight children from the younger SLI group and two children from the younger TD group had to be excluded for this reason.

In total, children heard 30 sentences (6 simple correct (S), 6 simple incorrect (*S), 9 subject relatives (SR) and 9 object relatives (OR)) (see Appendix H for the total list of items). For each item type, a practice trial was added to familiarize the child with the task. If children accepted the incorrect simple sentence in the practice phase, feedback was given by simply repeating the sentence and asking the child whether the sentence really matched the picture. After the practice trials, no feedback concerning their choice was provided. Administration of the task took 5-10 minutes. Answers were coded as either wrong (0) or right (1).

7.3.3 Elicitation task relative clauses (RCP)

To elicit relative clauses, a relative clause production task (*RCP-task*) was constructed based on an existing task designed by Novogrodsky & Friedmann

(2006). In the original task, designed for Hebrew, right-branching relatives were elicited targeting semantically reversible sentences with two animate noun phrases. The test has been used cross-linguistically and with several TD and language impaired populations (e.g., hearing impaired: Friedmann & Szterman, 2006; Friedmann & Haddad-Hanna, 2014; Szterman & Friedmann, 2014). The task was, however, not suitable for Dutch due to the structural ambiguity of relative clauses in the context of noun phrases with similar characteristics. A pilot with an exact Dutch translation of the original items elicited many ambiguous answers that could not be analysed as SR or OR with certainty (van Hout, personal communication, june 18, 2012). As was the case for the comprehension task, the elicitation test was therefore adapted to Dutch by adding a number distinction between subject and object in half of the items. Four different types of sentences were therefore elicited: 6 subject and 6 object relatives without a number distinction, which are strictly spoken ambiguous (*SR-ambiguous* and *OR-ambiguous*) and 6 subject and 6 object relatives with a number distinction (*SR-nr. distinction* and *OR-nr. distinction*) which disambiguated the sentence.[14]

Relative clauses were elicited by presenting participants with two short stories about two children (boys or girls, depending on the gender of the participant) and asking them to choose which of the two children they preferred to be. By asking them to start their sentence with the words *Ik ben liever…* 'I would rather be…' relative clause structures were primed (although the analyses will show that a number of alternative structures were used by participants). In the items eliciting subject relatives, the two children in the pictures were performing either two different actions or the same action but with a different object. In the items eliciting object relatives, the two children underwent two actions or the same action performed by different subjects. Table 7.1 shows an example of every item type. The SR-ambiguous and the OR-ambiguous have the same target answer (hence the ambiguity), with the only difference being that in OR-ambiguous contexts, a passive is a good alternative target answer.

The examples show variation on the nouns – which were stressed in reading out the story. Within each item type, half of the items showed variation on the noun, and half of the items showed variation on the verb. If the verb varied, the two verbs carried contrastive stress (*een jongen <u>kietelt</u> de moeder en een jongen <u>knuffelt</u> de moeder* 'a boy <u>tickles</u> the mother and a boy <u>hugs</u> the mother'*)*. The noun phrase *een jongen 'a boy'* (or 'girl' for female participants) was pronounced as naturally as

---

[14] The Dutch adaptation was made in collaboration with Naama Friedmann, who kindly provided help during a visit in November 2011, financed with a COST IS0804 grant for short term scientific missions.

possible, which meant it received stress in subject position (in subject relative items) but no stress in object position (in object relative items).

Table 7.1. Examples of items in the elicitation task for relative clauses (RCP-task)

Item type	Story	Target possibilities
SR-ambiguous	There are two boys and a father and a mother. One boy calls the father and one boy calls the mother. Which boy would you rather be?	*ik ben liever de jongen die de vader roept* 'I would rather be the boy that calls the father'
OR-ambiguous	There are two boys and a father and a mother. The father calls a boy and the mother calls a boy. Which boy would you rather be?	*ik ben liever de jongen die de vader roept* 'I would rather be the boy that the father calls' *ik ben liever de jongen die wordt geroepen door de vader* 'I would rather be the boy that is called by the father'
SR-nr. distinction	There are two boys and two footballers and two ballerinas. One boy calls the footballers and one boy calls the ballerinas. Which boy would you rather be?	*Ik ben liever de jongen die de voetballers roept* 'I would rather be the boy that calls the footballers'
OR-nr. distinction	There are two boys and two footballers and two ballerinas. The footballers call a boy and the ballerinas call a boy. Which boy would you rather be?	*Ik ben liever de jongen die de voetballers roepen* 'I would rather be the boy that the footballers call' *Ik ben liever de jongen die door de voetballers geroepen wordt* 'I would rather be the boy that is called by the footballers

The task was introduced with the following description:

> *'Ik ben een lijst aan het maken voor grote mensen, zodat ze beter weten wat kinderen fijn vinden. Wil jij me helpen met de lijst? Ik vertel je korte verhaaltjes. De verhaaltjes gaan over twee jongens/meisjes. Aan het eind van het verhaaltje moet jij me vertellen welke jongen/welk meisje jij wil zijn. Je moet je antwoord altijd beginnen met 'ik ben liever...'*
>
> 'I am making a list for adults with the preferences of children. Can you please help me with this list? I will tell you small stories. The stories are about two boys/girls. When the stories are finished, I will ask you to tell me which boy/girl you would rather be. You should start your answer with 'I would rather be…'

The test consisted of 24 items and the order of the item types was randomized. The task started with a practice trial which triggered a simple subject relative (without an object). If the participant did not start the answer with a relative clause during the practice phase, the item was repeated and the start of the relative clause was provided as an extra cue (*ik ben liever de jongen die…*, 'I would rather be the boy that...'). In such a case, an extra practice trial was then added to make sure the participant understood the task. If a participant immediately produced a correct target response on the first practice trial, the experimenter started the test. During the experimental phase, extra cues were not provided. The experimenter could only help the participant by saying 'I would rather be...', reminding them of the beginning phrase they had to use in their answers and offering them some help to start their answers. The question 'Which boy/girl would you rather be?' was provided structurally in the first items but could be left out if the participant did not need this cue anymore (a list of items can be found in Appendix I).

Administration of the task took around 10 to 15 minutes and was audiorecorded. The recordings were transcribed afterwards and coded for type of response. Dependent on the item type (SR-ambiguous, OR-ambiguous, SR-nr.distinction, OR-nr.distinction) the coding scheme varied since the list of optional answer structures was dependent on the target sentence structure (passives were for instance only expected in object contexts) (see Table 7.1). A list of all codings can be found in Appendix J. For statistical analyses, the codings were categorized into analysable and non-analysable answers. Non-analysable were answers without a relative clause/relative marker (e.g., *ik ben liever opa's* 'I am rather grandfathers) or answers which could for various reasons not be paraphrased into an intended meaning (e.g., *ik ben liever de jongen waarvan degene roept* 'I am rather the boy about whom someone calls/whose someone calls'??). Within the analysable answers, target answers and non-target answers were distinguished. For the object items, target answers were further divided into object relatives and passives, to see whether all groups produced the same amount of passive structures. Two children in the younger SLI group and two children in the younger TD group had to be excluded from the analyses of the elicitation task. These children either did not produce any relatives despite numerous elicitation cues (i.e. the two younger SLI children) or the experimenter did not give enough extra cues to elicit a relative in case of an answer with another structure, or gave away too much information (the two younger TD children).

7.4 Results

The results of the comprehension task will be presented first (§7.4.1), followed by the results of the judgement task (§7.4.2). Afterwards, the outcomes on the elicitation task will be discussed in §7.4.3. In all sections (for all tasks) the difference between subject and object relatives will be examined to see whether effect sizes of this linguistic factor are different in SLI than in TD. In §7.4.4, the effect of linguistic context (subject versus object relative) will be linked to the processing abilities described in Chapter 4.

7.4.1 Results Comprehension task relative clauses (RCC)

Figure 7.5 shows the total percentage of accurate answers on the comprehension task for every group. Again, the boxes and whiskers represent the range of scores, and the bold line in the middle of the boxes represents the median. As expected, adults performed well on this task, having a total accuracy rate of above 95%. The SLI and the TD groups did not perform at ceiling on the comprehension of relative clauses, and a clear difference between SLI and TD in overall accuracy appears. Note that the TD groups show a large variance in scores, and the SLI and the TD groups are partly overlapping.

If the total scores on the relative clause comprehension task are compared statistically, by means of an ANCOVA with total percentage accuracy as dependent variable, Group and Age as independent variables and IQ as a covariate, main effects of Group ($F(1,120) = 46.66$, $p < .001$, $\eta_p^2 = .280$) and Age ($F(1,120) = 4.65$, $p = .033$, $\eta_p^2 = .037$) are found. As hypothesized, Age and Group effects did not interact, indicating that the effect of SLI is not significantly different in the younger

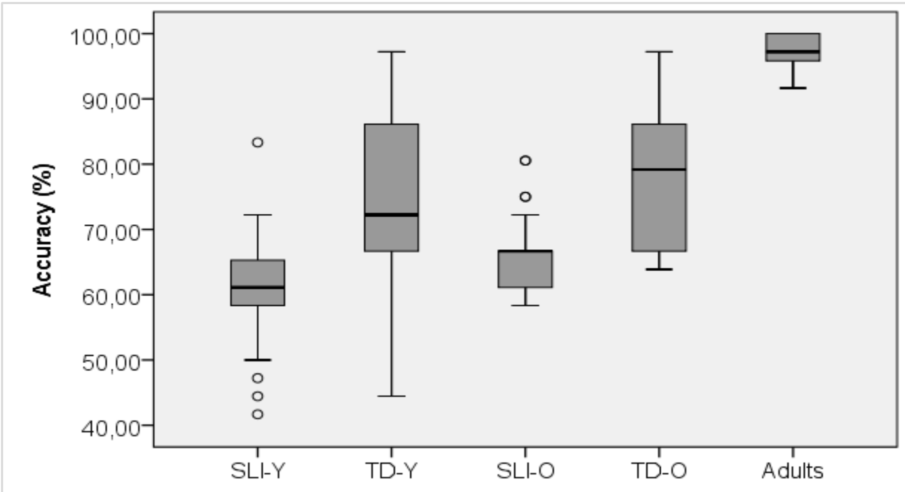


Figure 7.5. Accuracy (%) in the relative clause comprehension task (RCC) across the different groups

groups than in the older groups. Problems in the comprehension of relatives thus seem to be persistent in SLI.

In Table 7.2, the scores on the different item types in the comprehension task are shown (SR with a plural (SR-PL), OR with a plural (OR-PL), SR ambiguous and SR with a passive (SR-pass)). Adults showed minor differences between the different item types, as expected. They were able to use the agreement cue on the verb in the items with a number distinction between subject and object (SR-PL and OR-PL). Some errors were made in the comprehension of object relatives, but performance was at ceiling. In ambiguous sentences, adults preferred a subject interpretation, as expected (see our discussion in §7.1). Passives were comprehended without any problems.

In the SLI and the TD groups, a different picture appeared. In ambiguous contexts, all participants favoured the subject interpretation, as expected, and no significant differences between SLI and TD were found. Significant differences were found in the comprehension of the passive in SLI (a main effect of Group: $F(1,120) = 24.12$, $p < .001$, $\eta_p^2 = .167$ and Age: $F(1,120) = 9.70$, $p < .002$, $\eta_p^2 = .075$ and an interaction between the two ($F(1,120) = 7.26$, $p < .008$, $\eta_p^2 = .057$)). We will come back to these results when we discuss the data of the elicitation task, in which passives were a possible target answer (§7.4.3).

Table 7.2. Accuracy (%) for the different groups in the comprehension of relative clauses (RCC) (total score, SR ambiguous, SR-pass, SR-PL and OR-PL)

	N	Total % (SD)	SR-PL %	OR-PL %	SR-ambiguous %	SR-pass %
SLI-Y	32	60.94 (8.35)	89.84	16.67	86.46	66.15
TD-Y	32	75.18 (13.05)	96.61	39.06	89.58	90.10
SLI-O	31	66.04 (5.77)	94.62	11.82	94.62	88.71
TD-O	30	77.96 (10.92)	95.00	44.45	93.33	95.56
Adults	19	96.93 (2.91)	99.12	94.30	94.74	100.00

In both groups, object relatives with a number distinction were erroneously interpreted as subject relatives in the majority of cases. Object relatives thus seem to be a particularly difficult sentence structure to interpret. The fact that even the TD-O group was not yet scoring at ceiling raises the question when adult levels will be reached in typical development. The large variance in scores in the TD groups also raises the question what determines good or poor performance.

Both in the TD group and in the SLI group, the difference between subject and object relatives was significant. Repeated Measures with SR-PL and OR-PL as levels, with Group and Age as between-subject factors and IQ as covariate showed a significant effect of the linguistic context (OR versus SR) in the whole sample ($F(1,120) = 525.14$, $p < .001$, $\eta_p^2 = .814$) and a significant interaction of this effect with group membership ($F(1,120) = 17.96$, $p < .001$, $\eta_p^2 = .130$). As hypothesized, object relatives were more difficult than subject relatives, and this difference was significantly larger in the SLI groups than in the TD groups.

One might argue here that these differences between TD and SLI are related to children with SLI having problems interpreting the subject verb agreement cue. Such an explanation does not hold for the variability in the TD groups, since TD children from six year onwards no longer have problems with this grammatical aspect, as was confirmed by the ceiling effects in the judgement of subject verb agreement presented in §6.4.1. In children with SLI, on the other hand, the comprehension of subject verb agreement was shown to be persistently different. They had problems rejecting the use of inflections in erroneous contexts, including 3rd person singular and plural markings, which are the relevant inflections here. In the SLI group, problems with the comprehension of object relatives containing a number distinction may therefore be related to problems in interpreting inflections. To investigate this option, we performed correlational analyses between the ability to comprehend object relatives with a number distinction and the ability to judge subject verb agreement (Chapter 6). These analyses revealed a significant correlation in SLI (Spearman's rho: r = -.248, p = .050). The correlation is, however, in the opposite direction of what we would expect. The negative value means that the children with SLI who performed better on the judgement of subject verb agreement scored lower on the comprehension of object relatives with a number distinction and vice versa. Problems in subject verb agreement thus did not seem to play a role in explaining the lower comprehension of object relatives in SLI.

Similarly, one might argue that the difference between SLI and TD that was found in the comprehension task might stem from difficulties to interpret a cue at the end of the sentence. Participants had to listen and process the whole sentence in order to hear the cue and be able to use it to form an object interpretation. Because children with SLI have persistent problems in verbal information processing, as became clear in §4.4, this might imply extra difficulties in the comprehension of object relative clauses with an inflection cue at the end of the sentence. Within the SLI group, no correlations were found between the verbal processing measures and the number of object relatives that were comprehended correctly. This might be due to floor effects and lack of variation in the SLI group in both tasks. When correlations were computed for the pooled dataset (TD and SLI together), they appeared to be significant (r = .440, p < .001 for SR, r = .244, p = .006 for NWR and r = .271, p = .002 for digit span). The ability to comprehend object relatives thus seems to be related to processing abilities. Whether the link is explained by the position of the cue or by the syntactic processes involved remains, however, undetermined. As was discussed in §7.1.1, object relatives are often assumed to require more processing load than subject relatives because the distance between gap and antecedent is larger and *wh*-movement takes place across another syntactic or thematic role. Performance on object relatives in the comprehension task may, therefore, be linked to processing abilities in different ways. This issue will be readdressed when we discuss the correlation between linguistic context factors and processing abilities.

7.4.2 Results Judgement task relative clauses (RCJ)

As was mentioned in §7.3.2, eight children in the younger SLI group and two children in the younger TD group had to be excluded from the analyses of the judgement task because they did not reject the incorrect simple sentences (a score of

0, 1 or 2 out of 6 on the incorrect items led to exclusion).[15] The results of the remaining subjects on the judgement task for relative clauses are shown in Figure 7.6.

In contrast to the results of the comprehension task, adults did not perform at ceiling on the judgement task and showed a large variance in scores. Furthermore, differences between SLI and TD were not as clearly present as they were in the comprehension task. A statistical comparison of the total scores of the different groups in the judgement task, by means of an ANCOVA with total percentage accuracy as dependent variable, Group and Age as fixed factors and IQ as a covariate, revealed no significant main effect for Group (p = .26), no significant main effect for Age (p = .24) and no significant interaction between Group and Age (p = .19). The ability to judge the grammaticality of relatives was thus not different in children with SLI in comparison to TD peers, and did not develop significantly with age (this will be discussed in more detail in §7.5 when we summarize the results).

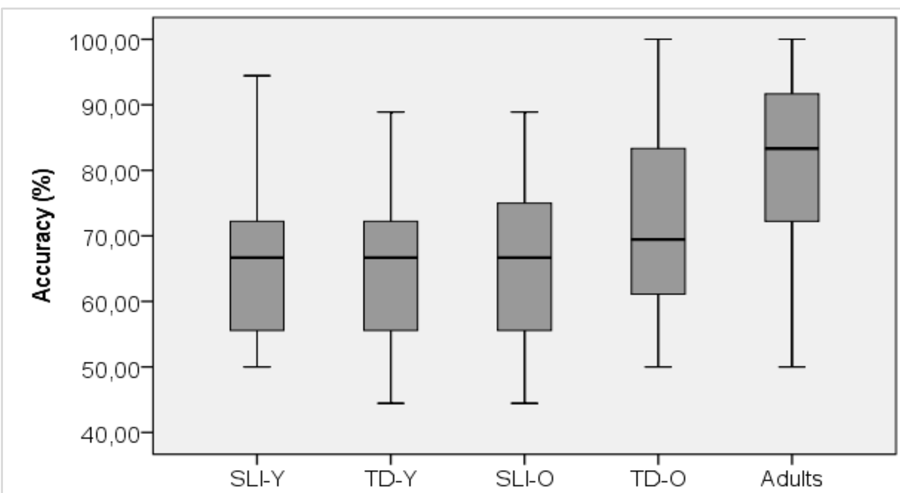


Figure 7.6. Accuracy (%) in the relative clause judgement task (RCJ) per group (SR and OR combined)

Table 7.3 shows the scores on the different item types in the judgement task (correct simple sentences (S), incorrect simple sentences (*S), subject relatives (SR) and object relatives (OR)). The judgement of correct simple sentences was at ceiling in all groups, but incorrect simple sentences were sometimes erroneously accepted by

[15] The reason for this relatively large number is unclear and needs to be explored in future research.

the two SLI groups (note that participants who accepted more than half of these items were already excluded from these numbers). This corroborates our previous findings in the judgement of grammatical gender (§5.4.1) and subject verb agreement (§6.4.1). In those judgement tasks, children with SLI were also shown to have difficulties in the rejection of incorrect utterances.

Table 7.3. Accuracy (%) for the different groups in the judgement of relative clauses (RCJ) (total score, S, *S, SR and OR)

	N	Total % (SD)	S %	S* %	SR %	OR % (SD)
SLI-Y	24	76.39 (6.87)	97.92	86.81	97.22	34.36 (26.40)
TD-Y	30	77.89 (6.10)	99.44	95.00	97.04	32.96 (20.94)
SLI-O	31	78.06 (7.44)	98.39	92.37	97.85	35.13 (24.70)
TD-O	30	83.11 (8.02)	100.00	98.89	98.89	45.46 (28.79)
Adults	19	89.30 (8.50)	100.00	100.00	97.08	67.25 (25.52)

As becomes apparent from Table 7.3, there was a clear difference between SRs and ORs in judgement accuracy in all groups, even in adults. While subject relatives were usually accepted with only a few exceptions, object relatives were rejected in the majority of cases (adults rejected them in one third of the cases). Repeated Measures with SR and OR as levels, with Group and Age as between-subject factors and IQ as covariate showed a significant effect of linguistic context (OR versus SR) in the whole sample ($F(1,110) = 600.78$, $p < .001$, $\eta_p^2 = .845$), but did not show significant interactions of this effect with group membership ($p = .368$), nor with age ($p = .434$). The hypothesis that the asymmetry between the judgement of subject and object relatives would be larger in the SLI groups than in the TD groups was thus not confirmed.

Recall that the object relatives in the judgement task consisted of three types. Object items contained either an animacy cue, a semantic cue, or no cue for object interpretation. Table 7.4 shows the scores for the object relatives, separated into the three different types of cue for object interpretation. In line with our hypotheses, the three types of cue for object interpretations resulted in different scores in all groups. As expected, object relatives with an animacy cue (the most typical object relatives) were accepted most often, followed by object relatives with a semantic cue. In reversible object relatives in which no cue for object interpretation was provided, acceptance rates were the lowest. In all groups, large variances in scores appeared, indicating that some individuals accepted object relatives while others (consistently) rejected them. The type of cue influenced the results significantly ($F(1,110) = 121.45$, $p < .001$, $\eta_p^2 = .525$), but this influence did not significantly interact with Group ($p = .529$) or Age ($p = .572$). The hypothesis that children with SLI profit less from cues for object interpretation than TD children was therefore not supported.

Table 7.4. Accuracy (%) for the different groups in the judgement of object relative clauses (RCJ) (total OR score, OR animacy cue, OR semantic cue, OR no cue)

	N	OR Total % (SD)	OR animacy cue %	OR semantic cue %	OR no cue %
SLI-Y	24	34.36 (26.40)	50.00 (34.05)	38.89 (36.34)	13.89 (30.95)
TD-Y	30	32.96 (20.94)	62.22 (36.86)	27.78 (34.00)	8.89 (23.05)
SLI-O	31	35.13 (24.70)	59.14 (39.17)	35.48 (35.42)	10.75 (20.00)
TD-O	30	45.46 (28.79)	63.33 (35.40)	47.78 (39.81)	25.56 (35.75)
Adults	19	67.25 (25.52)	87.72 (19.91)	70.17 (34.95)	43.86 (45.88)

Two issues deserve some further discussion here: the large variance in scores on object relatives and the difference in outcomes between the different types of object relatives in the two different tasks. In order to discuss these two issues, the data on the different types of object relatives in the two perception tasks (RCC and RCJ) is combined in Figure 7.7 (object relatives with a number distinction in the comprehension task (RCC), object relatives with an animary cue, a semantic cue, or no cue in the judgement task (RCJ)).

As is clearly visible in the graph with the combined data, almost all adults interpreted object relatives with a subject verb agreement cue correctly in the comprehension task. Their scores showed increasing variance in the judgement of object relatives with different type of cues in the judgement tasks. This indicates that some adults consistently accepted object relatives under all conditions, while others only accepted them in certain circumstances. This raises the question what determines the ability to interpret object relatives correctly in all conditions. In the

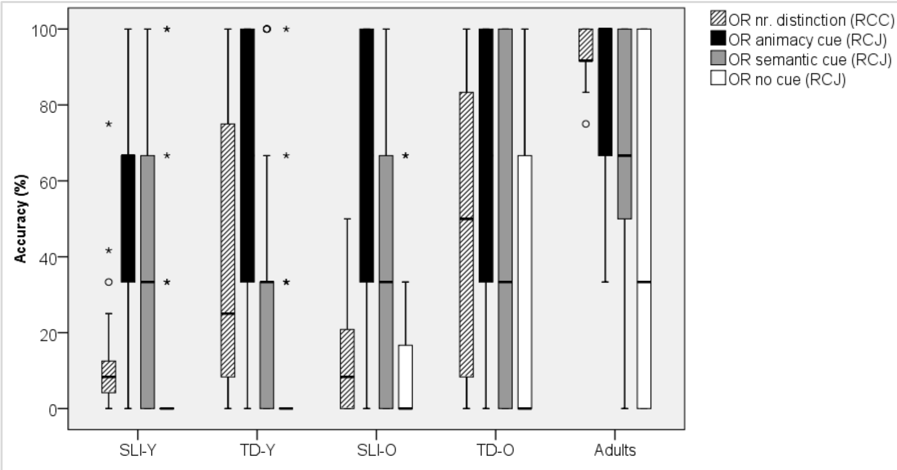


Figure 7.7. Combined plot of the mean accuracy (%) on different types of object relatives in the comprehension task (RCC) and the judgement task (RCJ) for the different groups

child and adolescent groups, variance was even larger. The child and adolescent data showed no significant effects of IQ on the ability to judge object structures in the judgement task (RCJ), or to interpret object relatives accurately in the comprehension task (RCC). The effect of type of cue did not interact with IQ either. IQ therefore thus not seem to explain differences in the ability to interpret object relatives. It remains to be determined which factor explains the patterns in the object relative clause data.

Furthermore, the combined graph shows that the children and the adolescents had a different ordering in their object relative results than the adults. If we compare the different type of object relatives, adults performed best on the object relatives with a subject verb agreement cue, even though these items included two animate arguments, which is not typical for (Dutch) object relatives (Mak et al., 2002; 2006; Kidd et al., 2007). Children and adolescents, on the other hand, had the highest performance rates in object relatives with an animacy cue. From the four types of object relatives, these items had the most typical characteristics because they involved a difference in animacy, with the object being inanimate and the subject being animate (see §7.1.1 and 7.1.2 for a more elaborate discussion on the typical characteristics of (Dutch) object relatives). Children and adolescents therefore seemed to rely more on the typicality of object relatives, while adults could also interpret non-typical object relatives correctly on the basis of a grammatical cue.

Apart from a different ordering between groups, the two perception tasks for relative clauses yielded different conclusions regarding the difference between SLI and TD. Differences were found in the comprehension task while they were not found in the judgement task. Conclusions on difficulties with object relatives in SLI are thus very much dependent on the type of task and the characteristics of the object items.

7.4.3 Results Elicitation task relative clauses (RCP)

As was described in §7.3.3, a detailed coding scheme was constructed for the elicited answers in the elicitation task for relative clauses. Coding depended on the item type since the list of optional answer structures was dependent on the target sentence structure (SR or OR). Passives were, for instance, only expected in object contexts (see Appendix J for the different coding schemes). For analysis, the codings were categorized into analysable and non-analysable answers. As described in §7.3.3, an answer was considered analysable if it included a relative clause and if a paraphrase could be constructed. Within the analysable answers, target answers and non-target answers were distinguished. For the object items, target answers were further divided into object relatives and passives.

Figure 7.8 shows the elicited answers for the different type of relative clauses. The upper two graphs show the elicitation of subject relatives in ambiguous contexts (left) and contexts with a number distinction (right) and the lower two graphs show the elicitation of object relatives in ambiguous contexts (left) and contexts with a number distinction (right). Each graph shows the proportion of target structures (SR in SR-contexts, OR or passives in OR-contexts), non-target structures (OR in SR-contexts and SR in OR-contexts) and non-analysable answers. Ambiguous relatives were analysed as target subject relatives in SR-contexts and as target object relatives in OR contexts, thus giving the children the benefit of the doubt. To give an example: the answer *ik ben liever de jongen die de tante kietelt* 'I would rather be the boy that tickles the aunt/the aunt tickles' was counted as an SR in ambiguous SR contexts and as an OR in ambiguous OR contexts.

The individual graphs in Figure 7.8 show that adults always produced subject relatives in subject contexts, and produced passives in the majority of cases in object contexts (around 95%). Object relatives were produced by adults in only 5% of the cases. In the TD groups, a clear developmental pattern towards adult-like behaviour emerged, both in the production of subject relatives and in the production of object relatives. Younger TD children produced more unanalysable answers than older TD children, and more subject relatives in object contexts. The use of the passive as an alternative structure in object contexts also seemed to develop with age in typical development. The SLI groups clearly produced far more unanalysable answers than the TD groups, both in subject and in object contexts. As was shown before in §6.4.3, the production of subordinate clauses is - in general – difficult for subjects with SLI.[16] However, the older SLI group showed a clear growth of the amount of relative structures used, both in subject and in object contexts. In contrast to the TD groups, they produced a considerable number of role reversals (OR in SR contexts and SR in OR contexts).

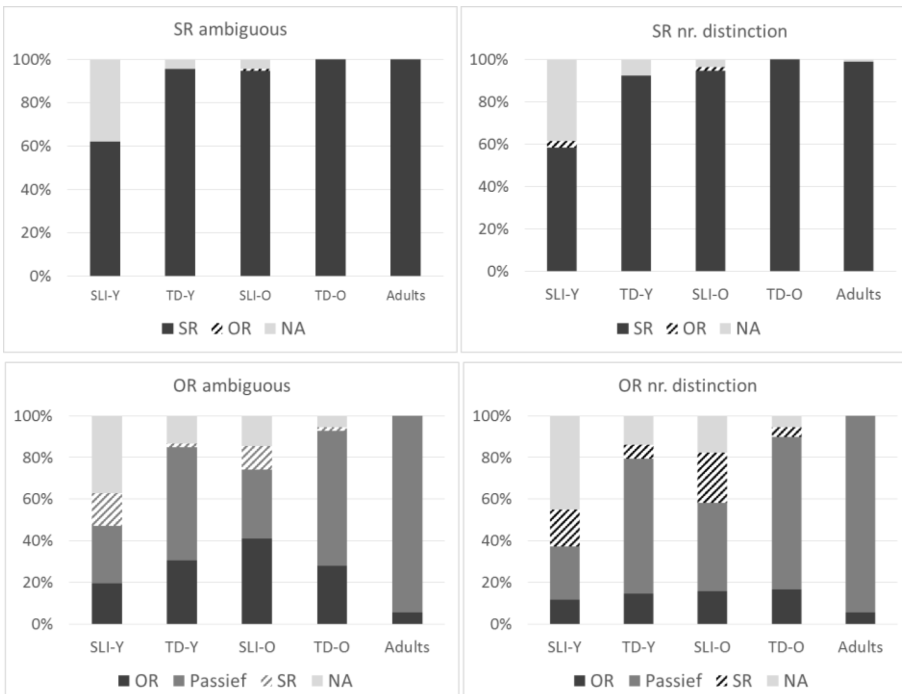


Figure 7.8. Answer types in the production of relative clauses, per type of relative clause and per group

Figure 7.8 seems to suggest that the SLI groups used more object relatives than the TD groups and the adults. The most plausible interpretation is that children with SLI were less able to use avoidance strategies in object contexts and produced more ambiguous structures (which were counted as object relatives in object contexts, to give children the benefit of the doubt, as described above). Almost all of the 'object relatives' produced in ambiguous contexts by the children and adolescents with SLI

[16] In Chapter 6, subject verb agreement was tested in different syntactic contexts. Subordinate clause contexts elicited many unanalysable answers due to word order alternations in the SLI-groups (see Table 6.16).

were ambiguous structures. In contexts with a number distinction, half of the ‘object relatives’ were also ambiguous, due to the fact that children did not repeat the plural subject in their answers. Whether children intended to construct an object relative or meant a subject relative with reversed roles is therefore hard to tell, although the latter seems more plausible. The main conclusion that can be drawn on the use of object relatives in SLI is therefore that these children and adolescents made less use of the alternative strategy of using a passive. As was discussed in §7.4.1, participants with SLI were also shown to make more errors in the interpretation of the passive. The difficulties in the comprehension of passives may indicate that children and adolescents with SLI have less access to strategies to avoid the object structure.

For statistical analyses, the data were split into target answers and non-target answers. Non-analysable answers were excluded and passives and object relatives were combined as target answers in object contexts. Grammatical or structural errors did not influence the accuracy analysis of the relative structure but were coded separately (and will be discussed later on). Thus, relative clauses without a relative word order, with an erroneous relative marker or with a role reversal were included as target answers in the analyses below. Children who had less than eight analysable answers (one third of the total number of items) were excluded from the statistical analyes for reasons of validity. In the younger SLI-group, five children were excluded on this basis, as opposed to one child in the younger TD group. Figure 7.9 shows the total accuracy on the relative clause elicitation task for the different groups.

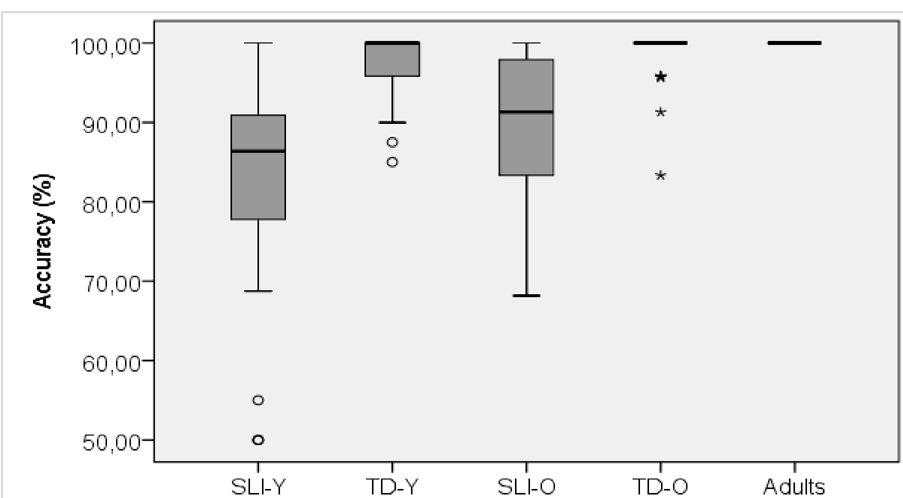


Figure 7.9. Accuracy (%) in the production of relative clauses per group

The results for the production of relative clauses were analysed with an ANCOVA, with total percentage accuracy as dependent variable, Group and Age as independent variables and IQ as covariate. This analysis showed significant main effects for Group ($F(1,111) = 43.733$, $p < .001$, $\eta_p^2 = .283$) and Age ($F(1,111) = 7.749$, $p = .006$, $\eta_p^2 = .065$). No significant interaction between Group and Age was found, although the *p*-value approached significance ($p = .052$). The overall ability to produce relatives was thus significantly different between individuals with and without SLI, but this difference was not dependent on age. Although the SLI group showed development in their production of target relatives, the difference in accuracy persisted.

Table 7.5 shows the accuracy in subject relative and object relative contexts. Since the results looked similar in the ambiguous contexts and the contexts with a number distinction, as we saw in Figure 7.8, the data of the two contexts was collapsed in the statistical analyses for convenience of presentation. In all groups, a significant difference between the production of subject and object relatives was present, as shown by Repeated Measures with the collapsed SR and OR scores as levels, Group and Age as between-subject factors and IQ as a covariate ($F(1,109) = 52.61$, $p < .001$, $\eta_p^2 = .326$). The difference between object relatives and subject relatives was larger in younger groups than in older groups, as indicated by a significant interaction with Age ($F(1,109) = 3.93$, $p = .050$, $\eta_p^2 = .035$). Furthermore, the difference between subject and object relatives significantly interacted with Group ($F(1,109) = 25.43$, $p < .001$, $\eta_p^2 = .189$). The difference between the subject and object relatives was larger in SLI than in TD.

Table 7.5. Accuracy (%) for the different groups in the production of analysable relative clauses (total score, SR and OR)

	N	Total accuracy % (SD)	SR target (SR) % (SD)	OR target (OR/passive) % (SD)
SLI-Y	25	81.21 (15.27)	97.61 (5.81)	63.55 (32.11)
TD-Y	29	97.69 (4.32)	100.00 (0.00)	95.27 (8.79)
SLI-O	31	89.09 (10.15)	98.50 (5.68)	77.99 (22.70)
TD-O	30	98.45 (3.58)	100.00 (0.00)	96.84 (7.23)
Adults	15	100 (0.00)	100.00 (0.00)	100.00 (0.00)

Besides a hypothesized difference in accuracy and effects of linguistic context, a difference between SLI and TD was expected in the use of the passive as a (grammatically correct) avoidance strategy in object contexts. Both object relatives and passive answers were counted as correct in object contexts, but they were coded separately in order to be able to see whether children with SLI and TD children used this ‘strategy’ to the same extent. Table 7.6 shows the number of object relative target answers, divided into answers with a (sometimes ambiguous) object relative and answers with a passive. As can be seen, the target answers of the SLI groups mainly consisted of (ambiguous) object relatives while the TD groups produced more passives in object contexts. An ANCOVA with the percentage of passives as dependent variable, Age and Group as fixed factors and IQ as a covariate showed a significant difference between groups in the use of the passive ($F(1,108) = .211$, $p = .017$, $\eta_p^2 = .052$). TD children behave more like adults, who use the passive construction in the vast majority of cases. Note, however, that the effect size is small and standard deviations in the SLI groups are high, indicating considerable variability in scores.

Table 7.6. Number of OR target answers divided into object relative and passive answers (%)

	N	OR target (OR or passive) N	OR % (SD)	Passive % (SD)
SLI-Y	23	152	55.16 (43.33)	44.84 (43.33)
TD-Y	29	296	31.37 (36.39)	68.63 (36.39)
SLI-O	31	246	52.95 (38.55)	47.05 (38.55)
TD-O	30	329	26.52 (34.80)	73.48 (34.80)
Adults	15	180	5.56 (7.50)	94.44 (7.50)

No significant effect was shown for Age ($p = .146$) and Group and Age did not interact significantly ($p = .647$). The slight development of using the passive strategically in object contexts that can be seen in Table 7.6, both in SLI and in TD, is thus not significant.

Two issues merit more discussion here. First of all, the fact that children with SLI produced so many non-analysable answers – also in subject contexts - raised the question what the unanalysable answers looked like. A more detailed description of the answer strategies of children with SLI in the elicitation task will possibly tell us more about the nature of their problems in the construction of complex clause structures. Table 7.7 provides more detailed information on the unanalysable answers of the SLI groups. The different categories represent different reasons for exclusion of the response (no relative clause, no relative marker, a relative marker + preposition, a relative clause with a resumptive pronoun, a complement clause instead of a relative clause, not paraphrasable, a procedural mistake or other). The table presents percentages of the total amount of unanalysable answers (shown in brackets).

As can be seen in Table 7.7, the number of unanalysable answers in the younger SLI group was roughly similar in subject and object contexts. Furthermore, the table does not show any clear differences in the types of unanalysable constructions that were used in the two contexts. Both in subject and in object contexts, the unanalysable answers in the younger SLI group mainly consisted of children providing answers without a relative clause (despite several attempts of the experimenter to elicit a relative). Another frequently used construction was the complement clause instead of the relative clause. Although children correctly repeated the cue *Ik ben liever...*'I would rather be', they seemed to finish the sentence as if they interpreted the start of the sentence as *Ik heb liever…'*I would rather have' - which requires a complement clause. Those answers rarely occurred in the TD group and seem to indicate a problem with the interpretation of the elicitation cue rather than with the construction of a complex clause.

In the older SLI group, a difference between subject and object relatives was visible, both in the number of unanalysable answers and in the type of unanalysable answers. Some types of unanalysable answers only appeared in object contexts and might be interpreted as an indication of problems with and avoidance of object structures. A clever and grammatical alternative to an object relative is the use of the relative marker *waar* + prepositions *van* or *bij* (this alternative accounted for most of the unanalysable answers in the TD groups). Another - incorrect – alternative is the relative construction with a resumptive pronoun at the trace position. *In sum,* the analysis of the type of unanalysable answers seems to indicate that children with SLI have problems with the construction of a relative in general, while adolescents with SLI mainly try to avoid object relatives (in other ways than using a passive). Caution is, however, required in drawing conclusions because the number of unanalysable answers in the older SLI group was quite small.

Table 7.7. Percentages of different types of unanalysable answers in the SLI groups in SR and OR contexts

Category	SLI-Y		SLI-O		Example (*CUE* +…)
	SR % (N=135)	OR % (N=148)	SR % (N = 15)	OR % (N=60)	*CUE:* *Ik ben liever…* 'I would rather be'
no relative clause	62.96	56.76	20.00	10.00	*de jongen met de vader* 'the boy with the father'
no relative marker	2.96	1.35	6.67	5.00	*de jongen de vader roept* 'the boy the father calls'
rel marker + prep (*waarvan/waarbij*)	0	0.68	0	16.67	*de jongen waarvan de vader (hem) roept* 'the boy whose father calls (him)'
relative clause with resumptive pronoun	0	3.38	0	10.00	*het meisje dat de vader haar roept* 'the girl that the father calls her'
complement clause instead of relative[1]	17.04	18.24	53.33	15.00	*dat de ridders een jongen vangen* 'that the knights catch a boy'
not paraphrasable	0	0	0	8.33	*de jongen die de prinses bellen* 'the boy that the princess call'
other	8.15	9.46	0	28.33	*de jongen met de tante die knijpt* ' the boy with the aunt that pinches'
procedural mistake	8.88	10.14	20	6.67	(no additional cue provided)

[1] The complementizer in Dutch complement clauses is homophonous with the relative marker *dat* 'that'.

The other issue is the grammatical correctness of the relative clauses that were produced by individuals with SLI. So far, this section has focused on the ability to comprehend and produce relative structures, and therefore grammatical and syntactic errors were not taken into account in the analyses of the elicitation data. A more detailed analysis of the error types in SLI may, however, provide insightful information about their struggles with the construction of Dutch relative clauses. The literature suggests that sometimes children with SLI make errors that do not occur in typical development. Error analyses do, however, not always agree across studies. Omission of the relative marker is, for instance, mentioned as a common mistake in (younger) participants with SLI in some studies (e.g., Schuele & Tolbert, 2001), while other studies (with different languages and ages of subjects) do not find these errors (e.g., Hesketh, 2006) (see §7.2 for a more detailed discussion of the cross-linguistic findings on relative clauses in SLI).

Table 7.8 provides an overview of the substitutions and word order alternations in the SLI groups in SR and OR contexts (omissions, substitutions and word order alternations). The overview presents numbers rather than percentages because omissions were part of the unanalysable data while substitutions and word order errors were part of the analysable data (they therefore did not have the same denominator).

Table 7.8. Type of grammatical errors in the SLI groups in SR and OR contexts

Type of error	SLI-Y		SLI-O		Example (*CUE + ...)*
	SR N (223)	OR N (212)	SR N (357)	OR N (312)	*CUE: Ik ben liever…* ‘I would rather be…’
Omission relative marker	4	2	1	3	**de jongen de vader roept* ‘the boy the father calls’
Substitution relative marker (*wie/welke*)	24	8	12	2	**de jongen wie/welke de vader roept* ‘the boy who the father calls’
V2 (word order alternation)	49	22	59	9	**de jongen die roept de vaders* ‘the boy who calls the fathers’

As we saw before in Table 7.7, omission of relative markers did sometimes occur, but not very frequently. Instead, children with SLI more often substituted the relative marker *die* (for common gender heads) or *dat* (for neuter gender heads) with the relative markers *wie* or *welke*.[17] Dutch children with SLI thus seem to substitute relative markers rather than omitting them, at least after age 6. Word order alternations occurred frequently, both in the younger and in the older SLI group. These alternations denote the use of the main clause word order with the verb in second position (V2) (*ik ben liever de jongen die roept de vader*s 'I would rather be the boy who calls the fathers'), while embedding requires the verb to be in final position. The number of word order alternations is higher in the older SLI-group than in the younger SLI-group, but recall that this group also produced far more relative clauses. The fact that the alternation rate is higher in subject relatives than in object relatives is explained by the fact that object relatives were often produced in the passive voice without an object being present (*ik ben liever de jongen die wordt gevangen* 'I would rather be the boy that is being caught'). Without an object being present, word order alternations are not visible. *In sum,* the error analyses show that Dutch children with SLI after 6 years of age substitute relative markers rather than omitting them. Furthermore, difficulties in the construction of a relative clause also revealed themselves in word order alternations.

The occurrence of word order alternations in subordinate clause contexts were also reported in §6.4.3, when the elicitation of subject verb agreement in different sentential positions was discussed. It proved to be difficult to elicit verb inflections in subordinate clause contexts because children used main clause word orders in half of the items. At that point, it was not clear whether children intended to produce the targeted subordinate clause and made a word order error, or whether they started a new main clause and did not interpret the test correctly (a procedural type of error). The production of subordinate clauses in SLI did, however, prove to be difficult in other ways: children made more errors in verb inflections in the subordinate clause and used much more avoidance strategies in subordinate clauses (AUX + INF constructions). The results of the elicitation task for relative clauses seem in line with these results. Not only were relative clauses frequently avoided, the relative clauses that were produced contained many word order alternations. It is still possible that children duly repeated the cue, but started a new main clause afterwards because they did not understand the task procedure. However, the many ways in which subordination proved to be difficult for children with SLI - and the ability of TD children the same age to interpret the elicitation cue correctly -

[17] Substitution of the relative marker *die/dat* by *wie/welke* can be analysed as colloquial Dutch, as Boef (2012) suggests. However, this answer type did not occur at all in TD children and adults.

suggests that main clause word orders in a subordinate clause can be interpreted as showing difficulties in constructing a complex clause.

7.4.4 Correlations between linguistic context effects and processing abilities

As we saw in the previous sections, several effects of task and linguistic context were found in the comprehension and production of relative clauses. Influences of the type of task (differences between SLI and TD were found in the comprehension task but not in the judgement task) could not be tested statistically, because the two tasks also involved different item types. The influence of linguistic context (subject versus object relative) was found in all groups, but proved to be significantly larger in the SLI groups in the RCC-task and the RCP task. To investigate whether the disproportional difficulty with object relatives in SLI in these tasks can be explained by limitations in processing abilities, correlations with the processing measures described in Chapter 4 were computed. For every participant, difference scores were calculated by computing the difference between the accuracy in subject contexts and the accuracy in object contexts Correlations were hypothesized between the larger effects of linguistic context in SLI (in RCC and RCP) and nonword repetition and sentence repetition, since these measures showed persistent limitations in SLI (as was discussed in §4.5). Correlations with other measures of processing were, however, computed. Correlations with the effect of linguistic context in the judgement task were also added, although this effect was not larger in SLI than in TD. Table 7.9 shows the values of Spearman's rho and the corresponding *p*-values.

Table 7.9. Correlations between effects of linguistic context (SR-OR) and the processing measures in the comprehension (RCC), judgement (RCJ) and elicitation of relative clauses

Spearman's rho:	Linguistic context RCC (SR – OR)	Linguistic context RCP (SR - OR)	Linguistic context RCJ (SR - OR)
Sentence rep.	r = -.35 (p < .001)	r = -.46 (p < .001)	n.s. (p = .116)
Nonword rep.	r = -.18 (p < .043)	r = -.46 (p < .001)	n.s. (p = .951)
Digit recall	r = -.22 (p < .013)	r = -.45 (p < .001)	n.s. (p = .713)
Visual recall	n.s. (p = .203)	r = -.32 (p < .001)	n.s. (p = .760)
Inhibition (EF)	n.s. (p = .513)	n.s. (p = .443)	n.s. (p = .544)

As can be seen in Table 7.9, the difference between subject and object relatives in the comprehension task was negatively correlated with measures of verbal working memory but not with visual working memory and inhibition. This suggests that the ability to interpret object relatives is related to working memory capacity (the greater the working memory capacity, the smaller the difference between subject and object relatives and vice versa). However, if we reduce the significance level to .01 because we ran multiple comparisons (Bonferroni-correction, .05/number of tests), only sentence repetition shows a significant correlation with the difference score in the comprehension task. The link between working memory and the interpretation of object relative clauses therefore does not seem to be a strong one. Problems in the comprehension of object relatives in SLI, therefore, cannot be explained by problems in general information processing capacities.

In §7.4.1, we discussed possible influential factors on performance in the comprehension task. Performance on object relatives appeared to be related to processing abilities. We also discussed that there are several possible ways to interpret this correlation in the comprehension data: it can either reflect difficulties to process the cue in sentence final position (subject verb agreement) or may alternatively show difficulties in syntactic movement over longer distances or across another syntactic or thematic role (*wh*-movement). At that point, it was not possible to distinguish between the two rationales. On the basis of the correlations of the judgement task, also presented in Table 7.9, we can conclude that the latter rationale does not hold and the former explanation seems to be more valid. If the relation between processing and performance on object relatives would be related to movement, we would have expected correlations to show up in the judgement task as well. The fact that the correlation only showed up in the comprehension task therefore seems to reflect that children with SLI had difficulties to process the cue in sentence final position.

In the production task, strong negative correlations were found between the effect of linguistic context (SR versus OR) and the verbal and non-verbal measures of working memory (sentence repetition, nonword repetition, digit recall and visual recall). Greater working memory capacity was related to smaller differences between the two different contexts. As expected, no correlations were found with motor inhibition. Processing limitations thus do seem to contribute to explaining problems in object relative clause production in SLI.

7.5 Conclusion and discussion

In this final section, we will summarize the findings from the different tasks for relative clauses and discuss whether the findings are in line with our hypotheses and the literature discussed in §7.1. We expected to find differences between SLI and TD in the comprehension, judgement and production of relative clauses, and we expected these differences to persist into adolescence. Furthermore, we expected to find differences between subject and object relatives in all tasks, and between different types of cue for an object interpretation in the judgement task. The effect sizes of these differences were expected to be larger in SLI. Our final hypothesis was that the differences in scores between linguistic contexts would be related to processing abilities.

The outcomes on the production of relative clauses confirmed the hypotheses and are in line with the literature on relative clause production in SLI. In the production task, persistent differences between SLI and TD were found, both in the amount of target structures used (participants with SLI produced many more unanalysable answers) and in the amount of errors in the target structures. Unanalysable answers mainly consisted of answers without a relative clause or relative marker (despite repeated cueing of the experimenter). Furthermore, in the relative clauses that were produced by the children and adolescents with SLI, relative markers were often substituted and many word order alternations occurred.

Together, this indicates general problems in the construction of relatives in SLI. Within the analyzable answers, the SLI groups produced fewer target answers than the TD groups (recall that target forms included a passive in object contexts, since this construction was preferred by adults. We will discuss this issue in more detail below). Both the TD groups and the SLI groups made more errors (role reversals) in object relative contexts than in subject relative contexts, but this difference was larger in SLI, as hypothesized. This corroborates previous findings of a (larger) subject-object asymmetry in SLI in other languages (Novogrodsky & Friedmann, 2006; Jensen de Lopéz et al., 2012). As hypothesized, the effect of linguistic context was correlated with processing abilities: larger differences between subject and object relatives were found in children with poorer processing skills. This seems to indicate that the disproportional difficulty of children with SLI in the production of object relatives is explained by their poorer processing abilities – as has been suggested before in the literature (e.g., Avrutin, 2000).

While the outcomes on the production task were in line with our expectations and previous findings, the outcomes on the comprehension and judgement of relative clauses were contradictory and require more discussion. In the comprehension task, the expected difference between SLI and TD was shown and this difference appeared to be persistent. However, in the judgement task no significant difference between groups was found. Similarly, the difference between subject and object relatives was larger in SLI in the comprehension task, as hypothesized, but no significant difference in effect size was found in the judgement task. Finally, the hypothesis that larger differences between subject and object relatives in SLI would be related to processing abilities was (partly) confirmed in the comprehension task but not in the judgement task. Conclusions on difficulties with the interpretation of object relative clauses in SLI thus seem to be dependent on the type of task and the characteristics of the items. In principle, children with SLI are able to comprehend object relative structures, but the conditions under which they accept or comprehend them seem to be more constrained.[18]

The comprehension task, which included items where the subject and object were both animate and where the cue for an object interpretation was verb inflection in sentence-final position, seems to have disadvantaged the SLI group in several ways. First of all, object relatives typically have an animate subject and an inanimate object, as we discussed in §7.1.1. The object items in the comprehension task where therefore not very typical – and distributional frequency effects in the input may have played a role, as has been suggested before by Kidd and colleagues (2007). Furthermore, the fact that the cue for an object interpretation was in sentence-final position may also have put the SLI group at disadvantage. As became clear in §4.4, the SLI groups had difficulties in information processing, and a cue in final position means that the whole sentence has to be processed in order to interpret the sentence correctly. Processing abilities appeared to be related to performance on object relatives in the comprehension task. This correlation could in theory also reflect difficulties with syntactic movement in object relatives, but the fact that a correlation between processing and object relative interpretation was absent in the judgement task indicates that this rationale does not hold. Children and adolescents with SLI thus seemed to have more problems in interpreting the cue in final position.[19]

---

[18] Here, we would like to come back to the disadvantages of judgement task designs mentioned earlier in §7.3.2. An important argument in favour of the validity of using judgements in testing the acceptance of object relatives is the clear effect of the type of cue. Because the same syntactic structure was accepted and rejected by the same participants (in the same task), depending on the semantic/pragmatic context, we can conclude something on the syntactic knowledge of participants.

[19] Negative correlations between performance on the comprehension task and the judgement of subject verb agreement (Chapter 6) showed that the task effect in SLI could not be explained by difficulties with the subject verb agreement cue itself.

These same factors that may have caused the differences between the comprehension task and the judgement task in this study (effects of animacy and position of the cue) may also account for differences between our conclusions and conclusions in the literature on relative clauses comprehension in SLI. As was discussed in §7.1, object relative interpretation was previously found to be facilitated by number discrepancies between subject and object in Italian children with SLI (Adani et al., 2010). The fact that this was not confirmed in the current study may be due to typological differences between Italian and Dutch (e.g., differences in the position of the verb or in the richness of the verb inflection paradigm). Furthermore, disproportional difficulties with the comprehension of object relatives in SLI were often reported in studies using comprehension tasks with animate subject and objects in which the thematic roles are semantically reversible (e.g., Friedmann & Novogrodsky, 2004; Jensen de Lopéz, 2012). Although we do not know whether object relatives with animate objects are more typical or frequent in other languages, we assume that this factor might play a role in the reported differences between SLI and TD.

In sum, this study on the comprehension, judgement and production of relative clauses in Dutch participants with SLI leads to the following conclusions. Dutch children and adolescents with SLI have persistent problems with the production of relative clauses in general, although their abilities improve over time. Object contexts are harder than subject contexts, and this difference is larger in SLI. The enhanced effect of linguistic context on the production of relatives is linked to processing abilities. In contrast to what is reported in the cross-linguistic literature on the comprehension of relatives in SLI, children and adolescents with SLI are equally able to interpret object relatives as TD peers if object relatives have typical characteristics. They do, however, show stronger negative effects when object relatives are non-typical or when the cue for interpretation is in sentence-final position.

As was discussed in the introduction of this chapter, relative clauses are constructed differently in different languages and the structure and developmental trajectory of relative clauses vary across languages. The acquisition of the Dutch relative clause had hardly been studied. As well as providing answers to our questions on differences between SLI and TD, this study also revealed some insights into the acquisition of the Dutch relative clause in TD children and gave some new information on comprehension and production patterns in adults.

First of all, an important finding was the 'avoidance' of Dutch object relatives by adults in the elicitation task. Instead of the target object relative, adults produced passives almost without exception. This is in line with previous findings for Italian

(Belletti & Contemori, 2009) and with outcomes on the elicitation of Dutch *wh*-questions (Schouwenaars, 2012), in which the passive was also used in object contexts by adults (*welke prinses wordt door de boerinnen gewassen?* 'which princess is being washed by the farmers?'). The avoidance of an object relative in our elicitation task and in the studies on Italian and Dutch *wh*-questions may be attributed to the fact that the context was not very conducive to the production of (Dutch) object relatives. Object structures were elicited in items involving two animate arguments, while (Dutch) object relatives typically involve an animate subject and an inanimate object, as we discussed in §7.1.1 and §7.1.2 (Traxler et al., 2002; Mak et al., 2002, Kidd et al., 2007). Future research might be able to indicate whether object structures are as hard to elicit in a more typical context (although it will not be easy to come up with a task design to elicit those).

The adult outcomes in the elicitation task had consequences for the analysis of the data of the children and adolescents. Passives were added as target answers in object contexts and, strictly speaking, the task was therefore not testing whether participants could construct an object relative clause, but whether they could construct an easier target answer when they were confronted with a difficult object context. However, the fact that a subject-object asymmetry was found (in terms of errors and unanalyzable answers), indicates that the task was still able to detect difficulties with object relative contexts. Furthermore, the task detected interesting differences in the ability to bypass a difficult construction. The two TD groups showed development towards adult-like use of the passive in object contexts, as shown before in the Italian studies (Belletti & Contemori, 2009) and in Dutch *wh*-questions (Schouwenaars, 2012). The SLI groups did not use this 'avoidance strategy' to the same extent. Passive constructions also appeared to be difficult to interpret in SLI. As a result, children and adolescents with SLI used considerably more (ambiguous) object relatives than TD peers and adults (which presumably were role reversals in most of the cases). Children and adolescents with SLI thus seemed to have fewer linguistic tools to bypass the difficult object construction.

The results on the interpretation of Dutch relative clauses also revealed a number of interesting patterns. Adults were perfectly able to interpret object sentences on the basis of a number distinction between subject and object and an agreement cue on the verb. This was the case despite the fact that both subject and object were animate full noun phrases, which is not typical for object relatives (Kidd et al., 2007; Mak et al., 2006; Fox & Thompson, 1990). In contrast to the findings of previous studies, more typical object relatives, with an inanimate object and an animate subject, were harder to interpret for adults, as indicated by higher levels of variation in scores in the judgement task. This may be partly due to the nature of the judgement task, which requires metalinguistic knowledge to a greater extent. In general, adults made more errors in judgement than in comprehension, also in contexts where errors were not expected (subject relatives). In previous studies, knowledge of Dutch relatives was tested more implicitly using comprehension or reading tasks (Mak et al., 2002; 2006). However, the effect may not be entirely attributable to task differences: for adults, a grammatical cue seemed to be stronger than an animacy cue in the interpretation of object relatives.

Children, on the other hand, seemed to rely more on typicality than on grammatical cues. Their scores were highest on the object items with a difference in animacy, and the interpretation of object relatives with a subject verb agreement cue was less accurate. As was described in §7.1.2, it has been previously found that TD children are not yet able to use a verb inflection cue consistently to construct an object interpretation in *wh*-questions, although performance improved at age 9 (Metz et al., 2010, Schouwenaars et al., 2014). The fact that scores are still far from adult-like in the adolescents between 12 and 16 in this study raises the question at what age TD children will attain adult levels – if at all. Future studies with children older than sixteen should reveal at what point adult-like patterns will appear.

Another issue in the typical data is the large variance in scores. Even adults showed large variances in their interpretation of object relatives, especially when roles were reversible. Apparently, all adults can accept object relatives, but the conditions under which they accept them vary. The fact that the adults in this study were all highly educated suggests that in a more representative sample of the population, variability might be even larger. This raises the question what determines the ability to interpret object relatives in less felicitous conditions. Since IQ was not related to performance, it remains undetermined which factor explains the variability in the adult data. We can, however, conclude that Dutch adults do not all have the same grammatical intuitions when it comes to object relative clauses.

To conclude this chapter, we would like to reflect on whether the findings of this study match the theories on the subject-asymmetry in the comprehension of relative clauses discussed in §7.1. As was mentioned in that section, several theories have been put forward to explain the disproportional difficulty of object relatives. Two explanations were based on the findings that the subject-object asymmetry is dependent on the characteristics of the arguments: the idea that object relatives are harder than subject relatives due to intervention effects (Belletti & Rizzi, 2013), and the idea that the asymmetry is explained by the distributional and discourse regularities of the input (Kidd et al., 2007). The former theory argues that similarity between subject and object defines the difficulties in the interpretation of object relatives, because *wh*-movement in object relatives takes place across an intervening subject. This theory predicts that higher similarity between subject and object in terms of features (e.g., number, animacy) would lead to more difficulties in the assignment of thematic roles and thus in slower reaction times or lower accuracy rates. Kidd and colleagues (2007) argued that the established effect of similarity between intervening element and antecedent was in fact an effect of typicality. Object relatives typically have an inanimate head, and an animate subject, and if object relative have these characteristics, the subject-object asymmetry disappears. If this dissimilarity in animacy is reversed (the subject is inanimate, and the object is animate), object relatives are again harder to interpret (Traxler et al, 2002; Mak et al, 2002), which seems to indicate that it is the typicality of object relatives rather than the (dis)similarity in features that determines whether object relatives are hard to interpret (Kidd et al, 2007).

The adult data of this study did not support the idea that typicality is the strongest predictor of difficulties in object relative clause comprehension. As we discussed above, adults scored higher on the interpretation of atypical object relatives with two animate referents that differed in number than on typical object relatives with a difference in animacy. Although the two types of object relatives were tested in different tasks, the difference cannot fully be attributed to task differences. For adults, a grammatical cue thus seemed to be stronger than an animacy cue in the interpretation of object relatives. The child data, on the other hand, supported the idea that typicality determines difficulties with the interpretation of object relatives. All four child and adolescent groups performed best on the most typical object relatives, and worse on the less typical ones. This effect was larger in individuals with SLI – who are probably relying more heavily on typicality.

One of the explanations for a subject-object asymmetry that has been invalidated in typical development, but might play a role in SLI, is the effect of canonicity. In typical development, the idea that non-canonical structures are harder to interpret does not seem to hold since other non-canonical constructions, like passives, are not problematic in typical processing. The findings on the comprehension and production of relatives with a passive indicate, however, that in the SLI groups passives are also harder to interpret and produce. Although canonicity may not be the only factor explaining the enhanced asymmetry between subject and object relatives in SLI, it may be an influential one.

8 Conclusion & Discussion

In this study the processing abilities and grammatical abilities of children and adolescents with SLI were tested with two research goals in mind. Firstly, we aimed to gain insight into the persistence of the problems that are often regarded as characteristic for SLI in childhood. Secondly, we wanted to test the hypothesis that language problems in SLI can be explained by a processing deficit that affects both the acquisition of grammatical knowledge as well as its implementation, leading to variability in grammatical performance. This was posited by Bishop (1994) as the Vulnerable Markers Hypothesis (VMH). Testing the predictions of the VMH led to the second and the third research questions: whether there would be (larger) variability in grammatical performance in SLI dependent on the (linguistic) context and whether this variability was related to (impaired) processing capacities. The two goals were closely related: the predictions of the VMH could best be tested in adolescent subjects in which grammatical knowledge has stabilized. In these older subjects, problems in the implementation of grammatical knowledge would presumably be more clearly visible.

To investigate these three research questions, a group of children and a group of adolescents with SLI were recruited and matched to two TD groups on age. Three aspects were chosen as grammatical variables on the basis of previous studies on difficulties in SLI: grammatical gender, subject verb agreement and relative clauses. Problems in grammatical gender and subject verb agreement have been attested before in Dutch children with SLI. Problems in relative clauses were assumed to exist on the basis of the international literature on SLI; they had not been studied before in depth in Dutch subjects with SLI. Each grammatical variable was tested both in perception and in production and for each variable, a number of linguistic context factors (phonological characteristics, frequency, syntactic characteristics) were varied to examine their effect on performance.

As well as studying the three grammatical aspects, several tasks for measuring processing capacity were included. These processing abilities were described in Chapter 4 and the outcomes on each of the three grammatical variables were discussed in Chapters 5, 6 and 7. The goal of this final chapter is to bring those findings together and to examine whether the hypotheses postulated in Chapter 2 were supported. In doing so, we will follow the order of the research questions presented in Chapter 1. Firstly we will reflect on the question whether grammatical differences between SLI and TD are persistent into adolescence (§8.1). Then we will discuss whether variability in grammatical performance, dependent on the linguistic context, was larger in SLI than in TD (§8.2). After having discussed the assumption that processing abilities are impaired in SLI, we will turn to the question whether variability in performance was related to these processing abilities (§8.3). Finally, this chapter will reflect on the impact of these findings, and will provide suggestions for further research (§8.4 and §8.5).

8.1 The persistence of grammatical difficulties in SLI

The first question of this study was whether the grammatical differences between SLI and TD that are characteristic in childhood persist into adolescence. Based on a review of the literature on adolescents with SLI, presented in Chapter 2, grammatical difficulties were hypothesized to be persistent. We expected to find differences between SLI and TD in the comprehension and production of grammatical gender marking, subject verb agreement and relative clauses, and we expected these differences to persist in the older groups. Figure 8.1 shows an overview of the outcomes on the different variables with respect to this first hypothesis. For every variable and every task it is indicated whether the expected difference between SLI and TD was present, and whether the difference persisted.

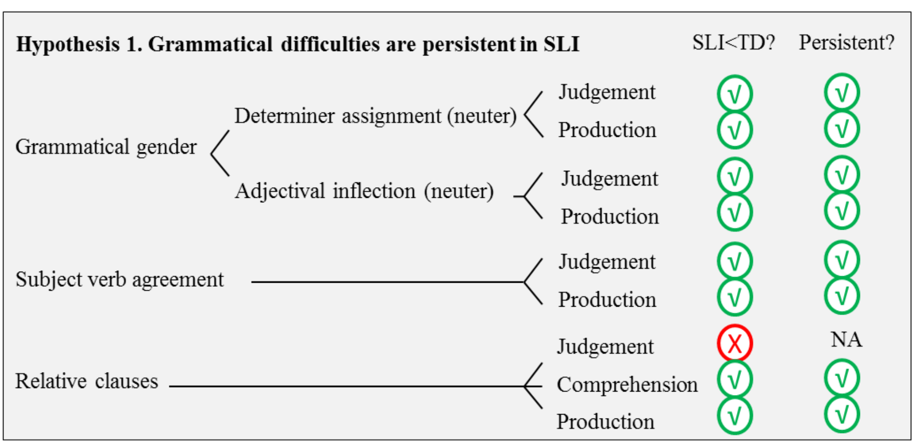


Figure 8.1. Overview of the outcomes on the persistence of grammatical problems

There were significant persistent differences between SLI and TD for all three grammatical aspects and in all tasks but one. The differences were not equally significant, as we will discuss in more detail below.

Large differences were found in the judgement and production of grammatical gender marking, both in determiners and in adjectives, as was described in Chapter 5. For common gender (the default gender in Dutch) persistent differences were not expected, but small differences appeared in the judgement task (this effect is not presented in Figure 8.1 since the neuter context is the relevant context here). For neuter gender, very large effects were found in all tasks. Some children with SLI do seem to discover a gender feature, as evidenced by production of the correct determiner and correct adjectival inflection for the same neuter nouns. However, the majority of children and adolescents with SLI either only produced default forms of the determiner and adjective, or seemed to have stored some nouns as being neuter (or, in Dutch, a *het*-word), but did not produce the adjective correctly. This result indicates a difficulty with the rule for adjectival inflection. When the two SLI groups were compared, no 'improvement' was found between the younger and the older group. Knowledge of grammatical gender seems therefore to fossilize before adolescence in SLI. These findings are in line with the previous literature on the acquisition of grammatical gender in Dutch children with SLI. Grammatical gender was previously found to be problematic in SLI till at least age 12 (Weerman et al., 2011; Keij et al., 2012). This study confirms the persistence of difficulties with grammatical gender for an even older age group. The majority of individuals with SLI will presumably never reach the stage of ultimate attainment in grammatical gender marking.

In subject verb agreement (Chapter 6) persistent differences between SLI and TD were also found, both in the judgement and in the production task. Differences were, however, not that large and became even smaller over time. On average, production accuracy in SLI was around 90%, which is often viewed as an indication that a grammatical aspect has been acquired (following Brown, 1973). The younger SLI group still showed a large variation in scores, indicating that some children had not yet mastered the rules for subject verb agreement. In the older group, variance was less and the majority of adolescents performed between 90 and 100% correct. Group differences in error rates were still found but the small amount of errors indicate that the rules for subject verb agreement have generally been mastered in adolescence in SLI.

Persistent errors in subject verb agreement consisted of omissions and substitutions of inflection markers, as has been found before in studies on Dutch children with SLI (de Jong, 1999; Steenge, 2006; Orgassa, 2009; Weerman et al., 2011). The nature of the errors did, however, alter with age: while younger individuals with SLI showed errors in singular and plural contexts, older subjects only made errors in the singular domain. As well as showing an 'improvement' in SLI in terms of a decrease in the number of errors, older individuals with SLI also used fewer auxiliary constructions. These constructions were often found in the younger SLI group and are often regarded as a strategy to avoid inflection of the lexical verb. The conclusion that subject verb agreement is not persistently problematic in SLI was therefore corroborated by the decrease in avoidance strategies in adolescence.

In the production of relative clauses, described in Chapter 7, differences between SLI and TD also appeared to be large and persistent. In the elicitation of these complex structures, the SLI groups produced more unanalysable answers, fewer target answers and their complex structures involved more grammatical errors. The older participants with SLI performed better than the younger ones, but differences with the TD group persisted, especially in object relative clauses. Furthermore, the strategy to produce a passive in object relative contexts – which is a perfectly legitimate alternative for object relatives in Dutch – was not developing as quickly in SLI as in typical development. In the production of relative clauses and alternative structures, persistent differences were thus clearly present.

In the interpretation of object relative clauses, on the other hand, differences between SLI and TD were found to be dependent on the task and the type of object relative. Differences were found in the comprehension task, but not in the judgement task. The two tasks differed in the type and typicality of their items. While the comprehension task included non-typical items with two animate arguments and a subject verb agreement cue in sentence-final position, the judgement task included more typical items with an animate subject and an inanimate object. Individuals with SLI performed equally well on the more typical items in the judgement task, but showed poorer performance in the comprehension task. They either rely more on the typicality of items, or were disadvantaged by the final position of the cue for correct interpretation. This indicates that the difficulties in the comprehension of (object) relatives that are often reported in the literature on SLI may be partly explained by other factors like typicality or effects of the position of the cue. Children with SLI are able to interpret object relatives to some extent, but are (persistently) more constrained by the linguistic context than their TD peers.

In sum, large and persistent differences between SLI and TD were found for grammatical gender marking and the production of (object) relative clauses. These aspects were persistently different in SLI, and seemed to indicate genuine problems in the grammatical domain; they may be suitable for distinguishing between children with SLI and TD peers. For subject verb agreement, differences between SLI and TD were also persistently found, but scores were generally quite high in both SLI groups (and seemed to develop over time). Although subject verb agreement has been suggested as a clinical marker in younger Dutch children with SLI (Blom et al., 2013a), it clearly loses distinctive power with age. Grammatical aspects thus remain vulnerable in SLI, but the specific grammatical profile changes over time.

8.2 Variability in grammatical performance in SLI and the influence of linguistic factors

The second and third research questions aimed at testing a theoretical claim about cognitive processes underlying SLI. The second question was whether children and adolescents with SLI show (larger) variability in grammatical performance, dependent on the complexity of the linguistic context in which grammatical aspects are tested. In §2.3, we discussed how the concept 'complexity' was defined and operationalized in this thesis. Complexity factors were very broadly defined as 'factors affecting processing accuracy or processing rate', and the linguistic factors that were implemented in the tasks were chosen on the basis of previously found effects. Various previous studies have indicated inconsistencies in the performance of children with SLI, which suggests grammatical knowledge might be present but is not always implemented in performance. We therefore expected to find variability in grammatical performance in SLI, dependent on the linguistic context. Variability was also expected in the TD groups, but larger effects were predicted in the SLI groups. Figure 8.2 shows the outcomes on the different variables with respect to this variability in grammatical performance. For every linguistic context factor, it shows whether these factors had an influence on performance in general and whether this effect was larger in SLI than in TD. All but one of the linguistic context factors had a significant effect on performance but this effect was not always larger in SLI than in TD.

In the tasks that were designed for grammatical gender, the influence of noun frequency and distance between determiner and noun were measured in the assignment of determiners. In adjectival inflection, performance was compared between items in which one adjective had to be produced and items in which two adjectives were required. Although noun frequency had a significant influence on the results, the effect was not larger in SLI than in TD. The other factors showed the hypothesized effect on performance: while scores for determiner assignment dropped significantly in the SLI group when the distance between determiner and noun was increased, no effect of this factor was found in TD individuals.

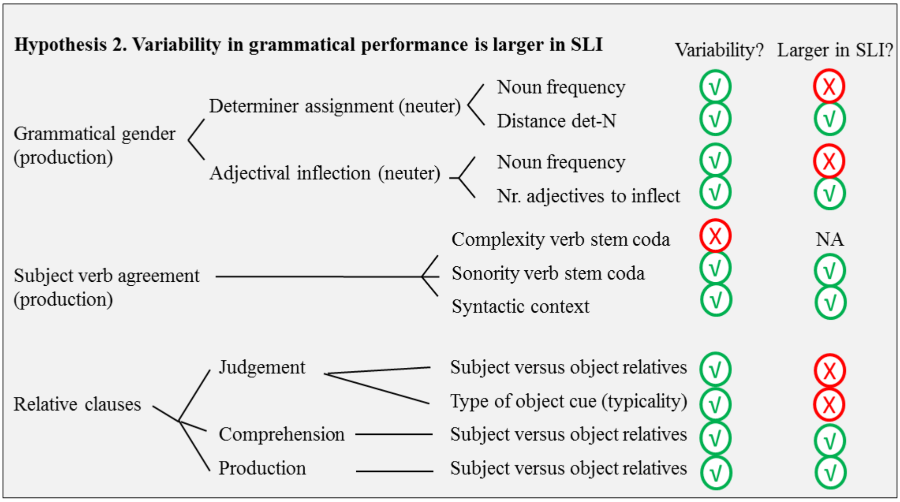


Figure 8.2. Overview of the outcomes on variability

Similarly, the number of adjectives had a significant effect on performance in the SLI groups, but not in the TD groups. To illustrate this with examples: the correct and incorrect determiners and adjectives in examples 1-4 were produced by the same child with SLI (C refers to common and N refers to neuter).

(1) *het net*
the.N fishing net.N
'the fishing net'

(2) *een blauw net*
a blue.N fishing net.N
'a blue fishing net

(3) **de rode, kleine net*
the.C red, small fishing net.N
'the red, small fishing net'

(4) **een kleine, blauwe net*
a small.C, blue.C fishing net.N
'a small, blue fishing net'

Although performance on grammatical gender was low in general in individuals with SLI, scores dropped significantly in more complex contexts. Previous studies that had noted effects of distance or proximity between a noun and its dependent elements found these effects in processing or perception tasks (e.g., Keating, 2009). The current study confirmed that such factors also influence production, and that individuals with SLI show larger effects of the linguistic context on their grammatical performance. It furthermore showed that linguistic complexity factors have a higher impact than frequency in SLI.

As well as effects of the linguistic context in which grammatical gender was tested, children with SLI also showed an effect of task: the older SLI group performed better than the younger SLI group in the judgement of determiners, but no significant ‘improvement’ was found in the production of determiners (while both tasks included the same nouns). Such ‘task effects’ were absent in the TD group (although they might be present in a younger TD group who do not yet perform at ceiling). Correlations between the two tasks showed a stronger correlation in the TD group than in the SLI group, indicating that individuals with SLI had a larger discrepancy in scores between the two tasks. In line with previous findings in the literature on SLI (Miller et al., 2008; Keij, 2009), variability in grammatical performance thus also seems to depend on the type of task in which performance is tested.

In measuring subject verb agreement two phonological factors and a syntactic factor were studied. Verb inflections were elicited in verbs that differed in the phonological complexity of the verb stem (inflections followed either a simple coda of one consonant or a complex coda of two consonants) and in the sonority of the verb stem coda (inflections followed a sonorant, a fricative or a plosive). The complexity of the verb stem did not seem to influence performance in general, in contrast to previous findings in other languages (e.g., Marshall & van der Lely, 2007; Song et al., 2009). Phonological reduction processes may account for these differences, since some of the complex verb stem codas in this study can be easily (and almost unnoticeably) altered to simple ones in production. It is therefore questionable whether the effect of complexity was adequately tested here. Sonority did have a significant influence on subject verb agreement accuracy, as had been previously found by Blom and colleagues (2014). Suffixation was more difficult when the verb stem coda ended in a plosive than when it ended in a sonorant. This caused more variability in the SLI groups than in the TD groups.

The syntactic factor that was implemented in the subject verb agreement task was a comparison between performance in main clauses and in subordinate clauses. These syntactic contexts had previously been found to affect performance in Dutch participants with SLI (Weerman et al., 2011; de Jong et al., 2013), with better performance in subordinate clauses than in main clauses (fewer auxiliary constructions and fewer inflection errors). In previous studies, these differences were interpreted as an effect of differences in syntactic movement: (dummy) auxiliary constructions were interpreted as avoidance of movement of the finite lexical verb (de Jong et al., 2013) and suffixation was argued to be more difficult in main clauses due to syntactic movement operations (Weerman et al., 2011). These explanations were challenged by the results of this study. In contrast to previous findings, we found that auxiliary constructions were more often used in subordinate clause contexts, and subordinate clauses more often contained inflection errors than main clauses. These effects were larger in SLI than in TD, and were mainly present in the younger SLI group.

The contrasting findings can presumably be explained by differences in design. In comparison to previous test designs, in which the subject of the subordinate (relative) clause was provided in the elicitation cue, participants in the current study had to construct the subordinate (complement) clause as a whole (see §6.5 for some examples and more details on the differences in task design). The differences in the amount of syntactic structure that had to be constructed by the participants might have led to the contrasting findings. The facilitative effect of subordinate word order is apparently overshadowed by the effort it takes to construct a complex clause. This challenges the idea that it was the difference in syntactic movement between the two clause types that triggered the greater use of dummy auxiliaries and the larger amount of inflection errors in main clauses in previous studies. For both phenomena, another explanation is conceivable: auxiliary constructions are not only less costly in terms of movement, but also avoid the inflection of the lexical verb. Similarly, inflection in main clauses may not only be harder due to movement operations, but also because of time constraints (e.g., Song et al., 2009; Blom & Baayen, 2013). Whatever factor may explain the contrasting findings between studies, the effect of syntactic contexts and the larger effect size in SLI than in TD confirm the idea that children with SLI show variability in grammatical performance, dependent on the (complexity of the) linguistic context.

For the third grammatical variable, that is relative clauses, effects of the role of the relative marker were expected (subject versus object relatives). In most languages, relatives are easier to comprehend or produce when the relative marker takes a subject role than when it takes an object role – although the asymmetry is dependent on the typology of the language (Laka, 2013) and the characteristics of the arguments (Kidd et al., 2007; Mak et al., 2002; 2006). In the production of relatives, both the TD groups and the SLI groups made more errors in object relative contexts than in subject relative contexts and this difference was larger in SLI. In the interpretation of relatives, the type of relative was only found to have an effect in the comprehension task, but not in the judgement task. As was discussed in §8.1 (and more elaborately in Chapter 7), the comprehension and the judgement task differed in the type and typicality of the object items, which influenced whether differences between SLI and TD were found. The type of object relative only affected performance in SLI in non-typical items. In more typical object items, the asymmetry was not larger in SLI than in TD.

In sum, variability in grammatical performance in SLI was found in all grammatical variables with all linguistic factors having an effect, except for the complexity of the verb stem coda. Performance appeared to be lower in contexts that were assumed to be more complex, and effects of these complexity factors were generally found in all groups but were often larger in SLI than in TD, as hypothesized. Because processing complexity was operationalized differently in the three grammatical variables (necessarily so since they differed in nature), we can now list a range of factors that more strongly influence grammatical performance in Dutch SLI. The distance between elements in a dependency relation seems to affect grammatical performance in SLI, as does the phonological context in which grammatical morphemes have to be placed. Furthermore, the use of correct grammatical morphemes is dependent on the syntactic context and the number of elements that have to be inflected. Lastly, syntactic performance in SLI is dependent on the characteristics of the arguments (canonicity, typicality). From a linguistic point of view, it is hard to come up with a common denominator for all these different effects of the linguistic context on grammatical performance. They cannot be exclusively explained by one specific linguistic notion such as canonicity, dependency, hierarchy, intervention or movement. All factors were, however, selected on the basis of their attested or hypothesized effect on (linguistic) processing. We will now verify whether the effects were indeed related to processing abilities.

8.3 Is variability in performance explained by (poor) processing abilities?

The third and final question of this thesis was whether the variability in performance that was found to be larger in SLI could be related to (impaired) processing abilities. To answer this question, the assumption that children with SLI also have persistent problems in information processing had to be first checked. We tested different aspects of processing ability using tasks that previously had been shown to discriminate between SLI and TD. Motor inhibition and four different recall tasks with pictures, digits, non-words and sentences (testing short term memory or working memory with different amounts of verbal load) were administered. Based on previous studies on processing problems in adolescents with SLI, processing difficulties were hypothesized to be persistent (see Chapter 4). This study shows that the assumption that processing abilities are persistently limited in SLI needs some refinement. Figure 8.3 summarizes the differences found between SLI and TD for the various processing measures, and indicates whether differences persisted. Whether difficulties persisted seemed to be linked to the amount of verbal load in a processing task. Motor inhibition did not reveal any differences at all and in the recall tasks differences between SLI and TD increased with the amount of verbal load involved (sentence repetition > NWR > digit recall > visual recall). Differences only seemed to persist in the two tasks with the highest amount of verbal load: the nonword repetition task and the sentence repetition task. Caution is required in concluding that the processing skills that do not show persistent differences (digit recall and visual recall) showed development with age: differences between SLI and TD mainly disappeared because the older TD group performed worse than the younger TD group. This might be an effect of better matching in the older groups, in terms of educational level, IQ and SES (see §3.2). Although we tried to correct for the large differences on these scales in the younger groups by entering IQ as a covariate, we cannot be sure that the differences were entirely accounted for (we will come back to this issue in §8.5).

Although impaired processing abilities have often been interpreted as an indication that the impairment in SLI is not specific to language, the outcomes of this study indicate that a general impairment in processing abilities is not an accurate description of the impairment in SLI either. We will discuss this point in more detail in §8.4.

Testing an assumption: Are processing abilities persistently impaired in SLI

	Sentence rep.	NWR	Digit recall	Visual recall	Inhibition (EF)
Impaired in SLI-Y?	√ $\eta_p^2 = .740$	√ $\eta_p^2 = .565$	√ $\eta_p^2 = .441$	√ $\eta_p^2 = .137$	X $\eta_p^2 = .014$
Persistently impaired?	√ $\eta_p^2 = .498$	√ $\eta_p^2 = .380$	X $\eta_p^2 = .048$	X $\eta_p^2 = .001$	NA ($\eta_p^2 = .020$)

Figure 8.3. Overview of the outcomes on the different processing measures

In the investigation of the third research question on the relationship between variability in grammatical performance in SLI and processing abilities, we only expected to find (negative) correlations with nonword repetition and sentence repetition because these processing measures showed persistent differences between SLI and TD. Correlations for the other measures were, however, computed in order to be able to link the results of this study to relations between grammar and processing measures that had been previously reported. In Figure 8.4 the outcomes of the correlational analyses are summarized.

Hypothesis 3. Variability in grammatical performance is related to (impaired) processing abilities

Negative correlation with:	Sentence rep.	NWR	Digit recall	Visual recall	Inhibition (EF)
Distance det-N?	√***	√***	√***	X	X
Nr. of adjectives to inflect?	X	√**	X	√**	X
Sonority verb stem coda?	√**	√***	X	X	X
Syntactic context SVA?	X	X	X	X	X
Syntactic context SVA? (*Younger groups*)	√***	√***	√***	√*	X
SR versus OR - RCC?	√***	√*	√*	X	X
SR versus OR - RCP?	√***	√***	√***	√***	X

Figure 8.4. Overview of the correlational analysis between variability factors and processing abilities. A (√) signals a significant negative correlation (* = $p < .05$, ** = $p < .01$, *** = $p < .001$)

For most of the linguistic context factors that resulted in larger variability in SLI than in TD, the expected correlations with sentence repetition and nonword repetition were found. Participants who showed a larger decrease in scores in more complex linguistic contexts often had weaker verbal processing abilities, and vice versa. For the number of adjectives that had to be inflected, the expected negative correlation was only found with nonword repetition, but not with sentence repetition. In contrast, variability dependent on this context factor was related to visual recall. This seems to indicate that it was visual processing that partly triggered variability. The pictures that were used for the complex contexts in adjectival inflection (see §5.3.3 for some examples) were visually quite complex: four different objects had to be distinguished on the basis of size and colour. However, if visual processing load triggered variability, similar relations would be expected in the determiner assignment task since this task made use of the same pictures. Visual complexity possibly only affects performance in grammatical aspects that are linguistically more complex (adjectival inflection rather than determiner assignment because it requires gender assignment in combination with implementation of a grammatical rule).

The expected negative correlations were also not found for the influence of syntactic context on subject verb agreement. As was discussed in §6.4.3, this factor mainly affected the younger SLI group and did not affect the scores in the older SLI group. When correlations are computed in both younger groups separately, they are highly significant. The effect of syntactic context on subject verb agreement accuracy is therefore related to processing abilities.

In sum, we can therefore conclude that variability in performance in SLI seems to be related to impaired processing abilities. Bishop's idea that 'that errors would occur when the speech production system is stressed by the need to produce output that makes heavy demands on its processing capacity' (Bishop, 1994, page 528) is therefore supported. We do not claim that SLI is *explained* by a deficit in processing capacities, nor do we claim that the relationship with processing that we found in this study has a clear causality. The results of this study do, however, indicate that errors in language production in SLI do not necessarily reflect an impaired grammatical system but might sometimes stem from difficulties to implement knowledge in production when the processing load is increased. In the next sections, we will discuss the theoretical and clinical implications of these findings.

8.4 Theoretical and clinical implications

The outcomes of this study add to the theoretical debate on the origin of SLI in several ways. The inclusion of a younger and an older group of individuals with SLI provided us with the opportunity to test whether the linguistic deficit in SLI is affected by age, as some theories of SLI claim. Furthermore, the inclusion of different grammatical variables and the manipulation of the linguistic context in which grammatical markers had to be produced provided the opportunity to evaluate whether an explanation for the deficit in SLI should be sought at the representational level or in more general information processing abilities.

The outcomes do not support the idea that SLI can be explained by an impairment in the representation or maturation of specific grammatical aspects. As was discussed in Chapter 1, the range of grammatical difficulties that are evidenced in SLI can only be explained as a deficit in linguistic knowledge if such a deficit is not too specific. An explanation in terms of an impairment in tense/agreement relations or of an extension of the optional infinitival stage is too narrow to explain the range of symptoms in SLI. Some representational theories, like the Computational Grammatical Complexity Hypothesis (CGCH) (van der Lely, 2005) or the Representational Deficit for Dependent Relations (RDDR) (van der Lely & Battell, 2003), do assume the representational deficit to be broader. They either assume a deficit in hierarchical relations (CGCH) or in dependency relations (RDDR) in SLI. The RDDR even leaves room for variability in grammatical performance by assuming an impairment in the linguistic principles that drive movement. This causes optionality of movement operations, leading to random performance. Although the CGCH and the RDDR capture a range of grammatical symptoms in individuals with SLI and also explain why performance can be variable, they cannot explain why this variability is influenced by linguistic context factors. As was mentioned in §8.2, it is hard to come up with a common denominator for the linguistic context factors that were found to affect grammatical performance. The greater effect of these linguistic factors in SLI cannot exclusively be captured by assuming a deficit in a linguistic notion such as hierarchy or dependency.

The outcomes do, however, support the idea of the VMH that grammatical errors in the language production of individuals with SLI may also stem from problems in implementing this grammatical knowledge in performance (Bishop, 1994, see Figure 8.5). The VMH was based on the observation of phonological and syntactic patterns in the utterances in SLI that contained a grammatical error (phonological complexity, number of syllables, number of phrases, sentential position of the grammatical marker). These patterns were observed in spontaneous data and Bishop described the need for experimental studies to complement the naturalistic data and to investigate the conditions which influence performance. She hypothesized that enhanced variability in grammatical performance in SLI would point to impaired processing abilities and that implementation of grammatical knowledge would be more vulnerable in more complex contexts that involve greater processing load (Bishop, 1994).

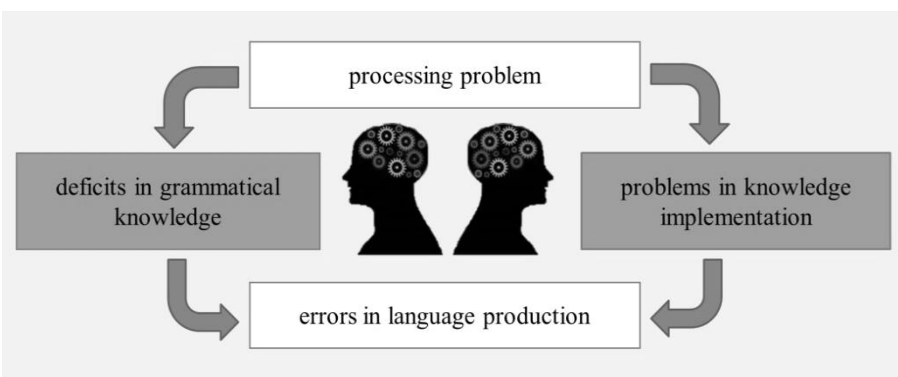


Figure 8.5. Schematic illustration of the Vulnerable Markers Hypothesis (VMH) (based on Bishop, 1994)

This study confirms the vulnerability of the implementation of grammatical knowledge in SLI and confirms the influence of linguistic context factors on grammatical performance. In addition to the factors studied by Bishop, a number of other linguistic context factors were shown to affect grammatical performance (distance between elements in a dependency relation, phonological context in which grammatical morphemes have to be placed, the number of elements that have to be inflected, canonicity/typicality of arguments). Effects were more clearly detected than would have been possible in an analysis of spontaneous language. Furthermore, this study also confirmed a link between the effects of linguistic context on grammatical performance and information processing. As Bishop hypothesized, problems in implementing grammatical knowledge in performance were related to the processing measures that (persistently) differentiated between SLI and TD. However, it cannot be concluded that SLI is caused by a processing deficit, as will be explained in more detail now.

In Chapter 1, we showed how processing theories on SLI differ in the way the processing deficit is described: some theories claim a general deficit in the processing of linguistic and non-linguistic information (e.g., Kail, 1994; Weismer et al., 2005), while others assume the processing deficit to be specific to linguistic information processing such as phonological processing or speech perception (Chiat, 2001, Joanisse & Seidenberg, 1998). The results of this study do not support the idea of a general limitation in information processing in SLI. As was described in Chapter 4 and summarized in §8.3, differences between SLI and TD were largest (and most persistent) in the processing tasks with a high verbal load (sentence repetition and nonword repetition). At the same time, the findings do not support the idea of a processing problem in one particular linguistic domain (phonological processing) or one particular modality (perception) either. It is possible that the processing tasks that were chosen to measure non-linguistic information processing (visual working memory, inhibition) were not complex enough to detect difficulties in SLI. Previous studies with older subjects with SLI reported that individuals with SLI have persistent problems particularly in more complex processing tasks or conditions (Weismer, 1996; Fazio, 1998; Hoffman & Gillam, 2004; Hick et al., 2005; Archibald & Gathercole, 2007; Spaulding et al., 2008). However, the fact that effect sizes of the difference between SLI and TD clearly increased with the amount of verbal load (while task complexity did not necessarily differ) seems to indicate that the processing problem in SLI is more 'language specific' than is often assumed.

Furthermore, the problems in verbal processing do not automatically imply a causal relation. As plausible as it sounds that poor verbal processing abilities cause difficulties in the acquisition of grammatical knowledge, the reversed causal relation might also hold: poor grammatical knowledge influences verbal processing, especially when the verbal processing task involves the repetition of sentences, which include phonological, morphological, syntactic and semantic aspects. For nonword repetition, an influence of grammatical abilities on performance may seem less intuitive, but nonword repetition accuracy has previously been shown to be linked to lexical abilities such as vocabulary size (Munson et al., 2005) and children with SLI often have smaller vocabularies (Hick et al., 2005).

Any conclusion regarding a causal relationship between processing difficulties and language impairments should therefore be drawn with caution. Henry and colleagues (2012) proposed a method to find out whether the relationship is causal and in which direction the causality runs. They computed whether differences in processing between SLI and TD remained after correction for differences in language ability. Because that was the case, they argued that the processing difficulties in children with SLI could not be attributed to their language impairment. Their conclusion was that children with SLI have a 'general difficulty with more complex forms of cognition, regardless of modality' (Henry et al., 2012, p. 43). If we apply their method to our data, we find that the differences in sentence repetition and nonword repetition remain after correction for general language ability in the older groups (although language ability accounts for a major part of the variance). However, in the younger groups the differences are entirely eliminated. These findings do not correspond with the idea that processing difficulties underlie the language problems in SLI.

In terms of the theoretical debate on the underlying cause of SLI both the existing representational theories on SLI and the existing processing theories on SLI seem to fall short in explaining the difficulties displayed by the population of this study. In order to provide an adequate explanation for the disorder, a theory needs to be able to explain the differences in information processing, the difficulties in the acquisition of grammatical knowledge and the enhanced effects of linguistic context on the implementation of grammatical knowledge in performance. It should also be able to explain why some grammatical aspects are more problematic than others to acquire. What is clear is the fact that difficulties in SLI are mainly found when linguistic, processing or task demands are increased. Children with SLI have problems in information processing when the tasks are more complex or when the verbal load is increased. Problems are more severe in grammatical aspects that are less transparent or syntactically more complex. Furthermore, grammatical performance is lower in linguistically complex contexts. Bishop's description of a limited capacity system that fails when it has to handle several operations in parallel might be an adequate description of the difficulties displayed in SLI. Leonard and colleagues (1992) had a similar idea when they hypothesized that morphological markers that are not very salient in the input may place more perceptual and production demands, which in turn may limit the resources available for additional operations of "deducing the grammatical function of a form and placing it in a morphological paradigm" (Leonard et al., 1992, p.153). It seems as if children with SLI mainly show problems when their linguistic abilities or their processing abilities are severely taxed.

Bishop (1994), in her description of the VMH, linked the idea of the loss of grammatical markers to the speech production model by Levelt (1989). In Levelt's model, the speech production process involves different steps and processors. Firstly, the 'conceptualizer' produces a preverbal message. Secondly, the 'formulator' transforms this preverbal message into a surface structure by retrieving content words and grammatical markings from the lexicon (grammatical encoding) and transforms the surface structure to a phonetic plan by retrieving the phonological forms connected to the words and grammatical markers and by adding prosody (phonological encoding). Thirdly and finally, the 'articulator' produces the phonetic plan. The three different processors operate in parallel rather than sequentially, because the speech production process is incremental: while (part of) an utterance is articulated, the next utterance is already being conceptualized or formulated (Levelt, 1989). In order to operate in parallel, automaticity is required.

According to Bishop, grammatical markers may be lost at different points in this speech production process: it may be that children with SLI do not achieve automaticity – leading to interference between conceptualizer and formulator, which potentially causes loss of grammatical marking when the intended message is complex (more errors when the message is semantically more complex). Another option is that grammatical encoding in the formulator is too slow, leading to the formulation of incomplete surface structures when the formulator has to generate a great deal of material. A third option Bishop mentions is that grammatical markers are formulated in the surface structure, but are lost in the phonological encoding of this surface structure into a phonetic plan when a considerable amount of material has to be processed since uninflected forms are easier to retrieve. It is hard to tease apart the second and the third option since they make similar predictions (more errors occur when the message requires more words or when the words contain more grammatical markers). A final potential locus of loss of grammatical markers in language production is the conversion of the phonetic plan into overt speech. Speech errors in the production of content words indicate that the speech production process is sometimes disrupted in this peripheral output stage. Bishop discusses the possibility that this might be due to the fact that the formulator and the articulator share resources in children with SLI due to poor phonological abilities. A phonologically encoded element may therefore be lost in articulation if the phonetic context is complex (more errors in phonetically complex contexts).

The effects of the different linguistic context factors on grammatical performance that were found in this study indicate that the problems in the implementation of grammatical markers are not restricted to one particular locus in the speech production model. The effects of the distance between elements in a dependency relation, or of the number of elements that have to be inflected correspond to the hypothesized problems in the formulator, while the effect of typicality/canonicity seems to indicate interference between conceptualizer and formulator. Furthermore, the influence of phonological/phonetic context may indicate a problem in phonological encoding, or in the process of converting the phonetic plan into overt speech. Children with SLI therefore seem to be impeded in their application of grammatical knowledge for different reasons.

At this point, it is opportune to make some remarks on the individual patterns in the SLI data. So far, we have based our conclusions on the group results, but many of our variables showed a large individual variability in performance. This raises the question whether specific profiles are hidden in the group data on processing abilities and grammatical abilities in SLI. Were the same children scoring high or low on all measures, or did different children have different 'difficulties'? This question is particularly relevant since we recruited children with SLI who had an official diagnosis of SLI, but whose language disorder did not necessarily involve a grammatical deficit (the inclusion criteria for a diagnosis SLI also involved other linguistic domains, as was discussed in §3.1). One way to address this question is to look at the correlations between the different measures within the SLI group. If the same children score high or low on all measures, we would expect to find high correlations between the different tasks.

Calculations of Spearman's rho between the different processing measures indicate that sentence repetition was correlated with digit span (r = -.328, p = .009) and nonword repetition was correlated with visual recall (r = -.386, p = .002). All other correlations were not significant. This indicates that different profiles in processing abilities exist in SLI.

The different profiles in processing abilities in SLI are also visible when we convert the processing scores into z-scores and plot a graph of the outcomes per individual case. Figure 8.6 shows the z-scores on the different processing measures per individual in the younger SLI group (upper graph) and the older SLI group (lower graph). The *length of the lines* in between the z-scores indicates whether the scores of the individual were similar on all tasks or whether different results were obtained in different tasks. The *position of the symbols* indicates whether children performed below or above average (the 0-line). Furthermore, each task is indicated with a different symbol.

As can be seen in Figure 8.6, most children in the younger SLI group performed below average on all processing measures. Some children performed above average on some processing tasks but none of the children performed well on all measures. In the older SLI group, there were a few 'good processors' who performed on or above average on all processing tasks and a few 'bad processors' who performed poorly on all tasks. The majority of adolescents with SLI did, however, show quite different scores on the different processing tasks. Which processing measures were difficult differed between children, as is reflected by different orderings of the symbols.

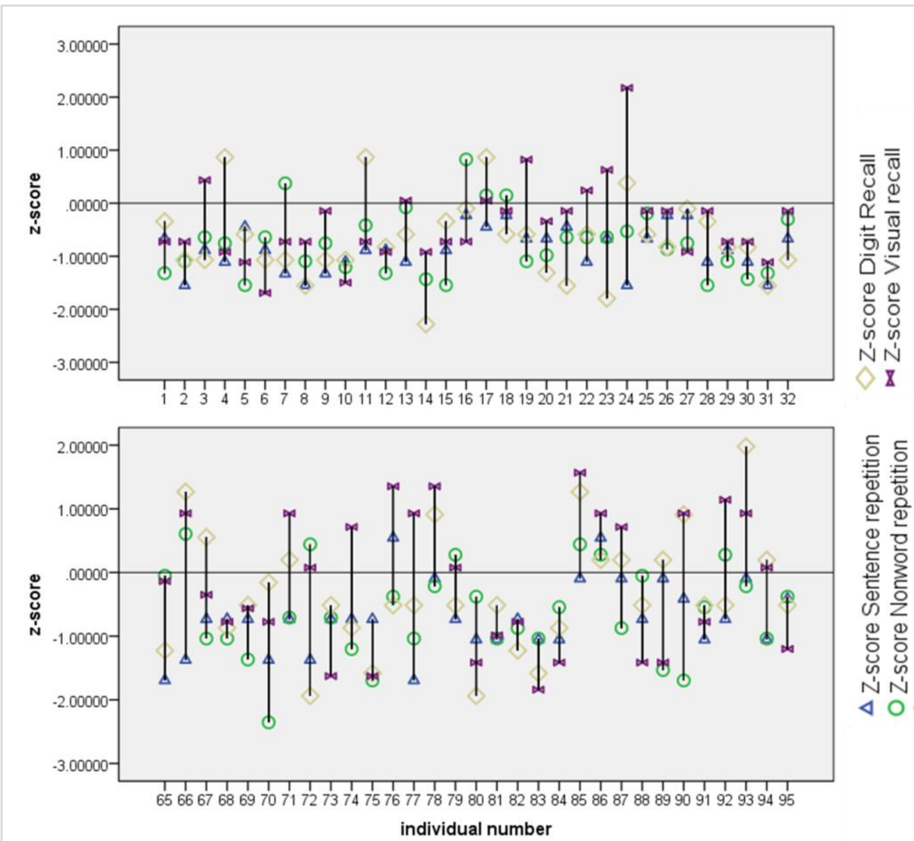


Figure 8.6. Individual patterns in the different processing measures (z-scores) in the younger SLI group (upper graph) and the older SLI group (lower graph)

In sum, almost all individuals with SLI therefore had difficulties in aspects of information processing, but which aspects were difficult differed per child.

The question whether different profiles could be detected in the individual scores of children with SLI also applies to the grammatical variables. Also within grammatical performance, an analysis of the individual scores reveals considerable variability in performance within individuals with SLI. Figure 8.7. shows the z-scores on the different grammatical measures in the younger SLI group (upper graph) and the older SLI group (lower graph). Again, the length of the lines in between the z-scores indicates whether the scores of the individual were similar on all tasks or whether different results were obtained in different tasks. The position of the symbols indicates whether children performed below or above average (the 0-line) and each task is indicated with a different symbol. As can be seen in Figure 8.7, z-scores are often quite far apart, indicating that individuals performed poorly on some grammatical aspects, but relatively well on others. None of the individuals with SLI performed on or above average on all grammatical aspects.

Correlational analyses between the different grammatical aspects within the SLI-group reveal that not all grammatical aspects are related. Interestingly, neuter determiner assignment is only weakly correlated with adjectival inflection in neuter contexts (this supports the conclusion we drew in §8.1 that as well as having a problem with the assignment of gender, some children with SLI have separate problems with the rule for adjectival inflection). Adjectival inflection does not show any correlations with the other grammatical measures. The only other significant correlations that appear are the correlation between subject verb agreement and determiner assignment and between subject verb agreement and relative clauses. This corroborates the discussion of the individual patterns presented in Figure 8.7, which showed very few consistencies within grammatical performance.

Table 8.1. Correlations between the different grammatical production tasks in SLI

Production of:	Adjectival inflection	Subject verb agreement	Relative clauses
Determiner assignment (N)	r = .252 (p = *.047)*	r = .481 (p < *.001)*	n.s. (p = *.070)*
Adjectival inflection (N)	1	n.s. (p = *.631)*	n.s. (p = *.537)*
Subject verb agreement		1	r = .358 (p = *.008)*
Relative clause production			1

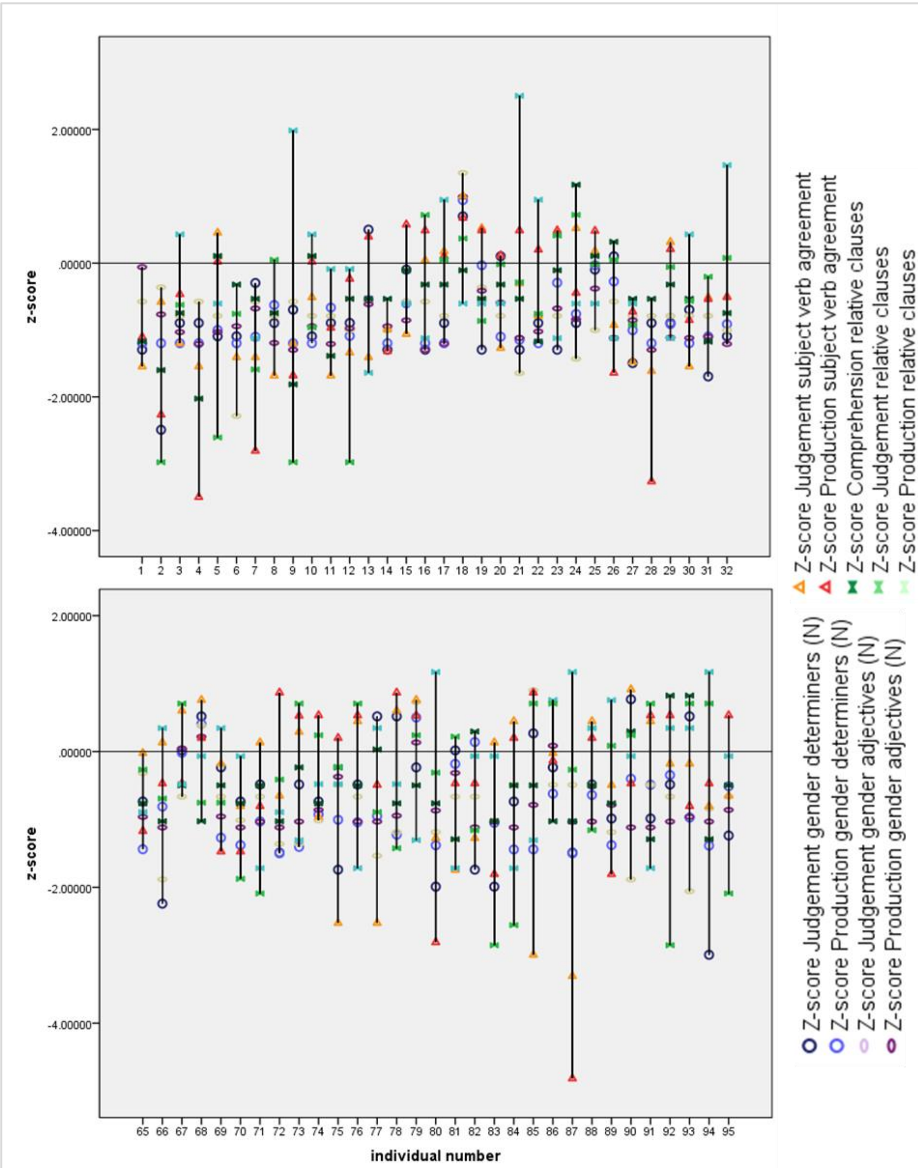


Figure 8.7. Individual patterns in the different grammatical measures (z-scores) in the younger SLI group (upper graph) and the older SLI group (lower graph)

The analysis of the individual results therefore did not support the idea that only some of the children and adolescents with SLI have problems in processing abilities or have severe impairments in grammatical aspects, while others are good processors or do not have a grammatical deficit. Although the variance in scores was large, most children and adolescents with SLI performed poorly on at least some aspects of processing. Similarly, all children and adolescents with SLI had difficulties with some grammatical aspects, despite the fact that they were not recruited strictly on the basis of problems in this domain. The results confirm the statement made in Chapter 1 that grammatical deficits are a core characteristic in children and adolescents with SLI. However, both in processing ability and in grammatical ability different profiles seem to exist, as was also indicated by significant correlations between some measures. Future research with larger sample sizes should indicate whether clear subgroups can be distinguished in the population of individuals with SLI.

In sum, this study adds to the theoretical debate on SLI by providing insight into the shortcomings of existing representational and processing accounts of SLI. That variability in grammatical performance is dependent on the linguistic context does not fit with the idea that the difficulties in SLI are purely explained by a deficit in linguistic knowledge. At the same time, the deficit in SLI seems to be more 'language specific' than is often assumed by processing theories. A common denominator in all the phenomena attested in SLI (problems in processing abilities, grammatical abilities and variability in the implementation of grammatical knowledge) seems to be that performance is mainly impeded when demands exceed. An adequate account of SLI should be able to account for the range of phenomena displayed in SLI and should presumably involve a description of the ‘limited capacities in SLI that lead to failure when demands exceed’. The individual results indicate that different individuals with SLI might have different difficulties in processing or in grammatical performance, although almost all seem to be persistently impaired in both processing and grammatical abilities. Future studies with larger groups should indicate whether these patterns point to different subgroups in the population of individuals with SLI. Such results might well then generate a different explanation.

The results of this study also contribute significantly to the clinical practice of diagnosing, treating, educating and raising children and adolescents with SLI. As we described in §1.1, the diagnosis of language impairments in adolescence is made more difficult by the fact that language milestones are less clearly identifiable because individual differences become larger and standardized tests are scarce

(Nippold, 1995; Reed, 2005). The results of this study showed that some of the grammatical markers that have been suggested as clinical markers for SLI in childhood, such as subject verb agreement, lose discriminatory power over time. Other grammatical markers like adjectival inflection in neuter contexts, on the other hand, seem to have considerable potential to discriminate between SLI and TD: the scores of the TD groups and the SLI groups did not show any overlap.[1]

For the treatment and education of individuals with SLI, the results of this study indicate that many children with SLI need long-term provision of special care. An important clinical issue is whether the persistent differences found between adolescents with and without SLI constitute a clinical problem. To what extent do the differences in grammatical performance imply a difficulty in daily life that would be relevant to remediate? Problems with grammatical aspects like grammatical gender or subject verb agreement do not automatically prevent the transmission of the message, but errors are immediately obvious to listeners since almost every sentence contains determiners and verb inflections. Similarly, complex relative clauses can be avoided in production, but problems in this domain can cause misunderstandings in oral or written communication, and can impede an individual in conveying a complex message. Apart from effects on communication, these differences may also cause social and emotional difficulties in adolescents. Listeners will often make evaluative judgements on the basis of errors in speech production since social awareness of SLI as a condition is limited.[2] It may therefore be relevant to help individuals with SLI in the acquisition of grammatical aspects. Not all aspects were, however, equally difficult in SLI. Subject verb agreement was, for instance, problematic in childhood but seemed to be acquired by adolescence. The most persistent difficulties were found in grammatical gender marking and relative clauses – these aspects should be taken into account when treating or educating adolescents with SLI.

Finally, the results of this study may help children and adolescents with SLI and their parents to accept the persistent nature of the disorder SLI and to make

---

[1] This only holds for monolingual individuals with SLI. Orgassa (2009) showed that adjectival inflection does not have discriminatory power in bilinguals.

[2] Awareness of language impairments and its symptoms is, however, growing, partly due to existing campaigns like the 'Raising Awareness of Language Learning Impairments' (RALLI) in England. This campaign was launched in 2012 and consists of several informative online videos (Conti-Ramsden, Bishop, Clark, Norbury, & Snowling, 2014). In the Netherlands, a Dutch organization for SLI called Spraaksaam has produced several products that promote awareness (a movie, a book) and organizes (informative) events.

realistic plans for the future with appropriate support. This will hopefully prevent them from experiencing the social, emotional, psychological and behavioral difficulties that are often reported as outcomes of the language disorder in the literature on SLI.

8.5 Suggestions for further research

Conducting scientific research is never without challenges and always raises new questions. This section offers the opportunity to reflect on those challenges and to list suggestions for future research. Some of those ideas have already been mentioned in previous sections.

In §8.3 we mentioned that differences in processing ability between SLI and TD are most clearly visible and most persistent in more complex processing tasks or conditions. The non-linguistic tasks that we used in our study may not have been complex enough to detect differences between SLI and TD in adolescence. Future studies on the relationship between language and processing abilities in SLI should therefore include processing measures that involve a more advanced level of complexity, especially when older subjects are studied. Furthermore, it is good to realize that it is not always clear what basic aspects are being measured by processing tasks. The visual-spatial task that we used in this study (the odd-one-out task, Henry, 2001), for instance, allowed subjects to remember the visual patterns using verbal cues. It was therefore not clear whether the task purely tested visual-spatial span or also tapped verbal memory.

In §8.4 we suggested that within processing and within grammatical abilities, different profiles seem to exist in SLI. These different profiles might be more clearly distinguishable when larger samples of SLI are considered. The correlations and individual patterns in this study might form the basis for the formation of hypotheses on profiles within processing or grammar. If such a study could be carried out, it would be useful to select a more homogeneous group of children with SLI. As was described in Chapter 3, the children in this study were all diagnosed as having SLI, but the diagnosis could have been based on different language domains. For an investigation of different profiles within grammar in SLI, it could be useful to select children for whom it is already known that they have problems in the grammatical domain (at the same time, the fact that all children in this study showed problems in some aspects of grammar while they were not selected on this basis indicates that grammatical problems are present in all children with SLI, also when their diagnosis is based on other linguistic symptoms). Instead of testing all different grammatical variables separately, which would be time-consuming if a large group of children has to be examined, a more parsimonious task design should be found.

Another issue related to the recruitment of subjects is the matching between SLI and TD. As was discussed in §3.2, the older TD group in this study was recruited from similar educational levels as the SLI group and differences in non-verbal IQ and SES between the groups were small. This was a major advantage, because we could be confident that the persistent differences were genuinely related to the language impairment status of the adolescents with SLI, and not to their non-verbal IQ, SES or educational level. The younger SLI and TD groups in this study, however, showed large differences on those variables. We have tried to correct for those differences by entering IQ as a covariate in the analyses, but future studies should think of matching the groups more closely on non-verbal IQ and SES. Since educational level is not yet determined in the Netherlands before the end of primary school (11 years of age), it may require testing on non-verbal IQ and SES in a large number of children in order to select two groups that are well-matched. Because children with SLI are often hard to recruit, it will be most feasible to test the SLI group first and search for good matches in a larger TD pool afterwards.

The differences in non-verbal IQ and SES were strongly related and for this reason we did not correct for both measures in the statistical analyses, but only corrected for differences in non-verbal IQ (this measure could also be more easily corrected for, since the variable was continuous rather than ordinal). However, the lower SES of children with language impairments merits some discussion. SES was defined in terms of the educational level of the mother (low, middle, high) since this has been shown to be a reliable measure of socioeconomic status. The fact that SES was lower in families from children and adolescents with SLI than in families of TD children suggests a links between SES and language ability in monolinguals. This link has often been attested in bilinguals, but little research exists on the relationship between the two in monolinguals (although Elbro, Dalby and Maarbjerg (2011) do report a link between language impairments and SES in adults with SLI). Future studies on SLI should therefore take SES into consideration, and will thus provide more insight into this relationship (especially when the groups are well-matched on non-verbal IQ).

A last issue to consider with respect to the recruitment of subjects is the selection of adult controls. As was previously discussed in Chapter 7, the adults in this thesis were all highly educated, which raises the question whether they were representative of the total population of adults. This question may especially be relevant for grammatical variables that are acquired late in typical development, such as the ability to interpret a Dutch object relative in certain contexts. The adults in this study all interpreted object relatives with a subject verb agreement cue correctly but showed a large variance in scores in other type of object relatives. Similarly, typically developing children and adolescents showed a large variance in scores and had not reached adult levels. This raised the question whether all TD individuals will eventually attain adult levels and whether a more representative group of adults would not also show variability in scores in object relatives with a subject verb agreement cue. Non-verbal IQ did not explain the differences between TD children in their accuracy on object relative items, but there may be another factor such as educational level, language ability or SES that can explain the individual patterns. This might be interesting to investigate in future studies.

In conclusion, we want to point out that variability in performance in SLI can be tested in ways other than those used in this thesis. We tested variability in individuals by looking at dips in performance in relation to linguistic factors that were predicted to impede the implementation of knowledge. Instead of looking at larger dips in performance in SLI, future studies could also look at effects of priming and thus detect peaks in performance. This idea was put forward by Kas and Lukács (2014). Seeing how much facilitation children with SLI can use should provide greater insight into their abilities and also possibly help discriminate children with SLI from TD peers.

Bibliography

Adani, F., Forgiarini, M., Guasti, M.T., & van der Lely, H.K. (2014). Number dissimilarities facilitate the comprehension of relative clauses in children with (grammatical) specific language impairment. *Journal of Child Language, 41*(04), 811-841.

Adani, F., van der Lely, H.K., Forgiarini, M., & Guasti, M.T. (2010). Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua*, *120*(9), 2148-2166.

de Ajuriaguerra, J., Jaeggi, A., Guignard, F., Kocher, F., Maquard, M., Roth, S., & Schmid, E. (1976). The development and prognosis of dysphasia in children. In D. Morehead, & A. Morehead (Eds.), *Normal and deficient child language* (pp. 345-386). Baltimore: University Park Press.

Alain, C., Woods, D.L., & Knight, R.T. (1998). A distributed cortical network for auditory sensory memory in humans. *Brain Research*, *812*(1), 23-37.

Aram, D.M., Ekelman, B.L., & Nation, J.E. (1984). Preschoolers with language disorders: 10 years later. *Journal of Speech and Hearing Research*, *27*(2), 232-244.

Archibald, L.M., & Gathercole, S.E. (2006). Visuospatial immediate memory in specific language impairment. *Journal of Speech, Language, and Hearing Research*, *49*(2), 265-277.

Archibald, L.M., & Gathercole, S.E. (2007). The complexities of complex memory span: Storage and processing deficits in specific language impairment. *Journal of Memory and Language*, *57*(2), 177-194.

Arkkila, E., Räsänen, P., Roine, R.P., Sintonen, H., Saar, V., & Vilkman, E. (2009). Health-related quality of life of adolescents with childhood diagnosis of specific language impairment. *International Journal of Pediatric Otorhinolaryngology*, *73*(9), 1288-1296.

Armon-Lotem, S., Haman, E., Jensen de López, K., Smoczynska, M., Yatsushiro, K., Szczerbinski, M., ..., & Kamandulytė-Merfeldienė, L. (2016). A large-scale cross-linguistic investigation of the acquisition of passive. *Language Acquisition*, *23*(1), 27-56.

Audring, J. (2006). Pronominal Gender in Spoken Dutch, *Journal of Germanic Linguistics*, *18*(2), 85-116.

Avrutin, S. (2000). Comprehension of *Wh*-questions by children and Broca's aphasics. In: Y. Grodzinsky, L.P. Shapiro, & D.A. Swinney (Eds.), *Language and the Brain: Representation and Processing* (pp. 295-312). San Diego: Academic Press.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189-208.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Baddeley, A.D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47-89.

Baker, L., & Cantwell, D.P. (1987). A prospective psychiatric follow-up of children with speech/language disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *26*(4), 546-553.

Barker, J., & Nicol, J. (2000). Word frequency effects on the processing of subject–verb number agreement. *Journal of Psycholinguistic Research*, *29*(1), 99-106.

Barry, J.G., Yasin, I., & Bishop, D.V.M. (2007). Heritable risk factors associated with language impairments. *Genes, Brain and Behavior*, *6*(1), 66-76.

Bastiaanse, R., Bol, G., van Mol, S., & Zuckerman, S. (2002). Verb movement and finiteness in language impairment and language development. *Clinical Linguistics: Theory and Applications in Speech Pathology and Therapy*, *227*, 119.

Bavin, E.L., Wilson, P.H., Maruff, P., & Sleeman, F. (2005). Spatio-visual memory of children with specific language impairment: Evidence for generalized processing problems. *International Journal of Language & Communication Disorders*, *40*(3), 319-332.

Beitchman, J.H., Wilson, B., Brownlie, E.B., Walters, H., Inglis, A., & Lancee, W. (1996a). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(6), 815-825.

Beitchman J.H., Wilson, B., Brownlie, E.B., Walters, H., & Lancee, W. (1996b). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(6), 804-814.

Beitchman, J.H., Wilson, B., Johnson, C., Atkinson, L., Young, A., Adlaf, E., …, & Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(1), 75-82.

Belletti, A., & Contemori, C. (2009). Intervention and attraction. On the production of subject and object relatives by Italian (young) children and adults. *Language Acquisition and Development, Proceedings of Generative Approaches to Language Acquisition (GALA)* (pp. 39-52).

Belletti, A., & Rizzi, L. (2013). Ways of avoiding intervention: Some thoughts on the development of object relatives, passive and control (pp. 115-126). In M. Piatelli-Palmarini, & R.C. Berwick (Eds.), *Rich languages from poor inputs,* Oxford: Oxford University Press.

Kas, B., & Lukács, Á. (2014). Structural priming of morphosyntactic features in Hungarian children with SLI. *Paper presented at the Experimental Psycholinguistics Conference (ERP), Madrid, Spain.*

van Berkum, J.J. (1996). *The psycholinguistics of grammatical gender: Studies in language comprehension and production* (Doctoral dissertation, Radboud University Nijmegen).

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

Bishop, D.V.M. (2004). Specific language impairment: Diagnostic dilemmas. In L. Verhoeven, & H. van Balkom (Eds.), *Classification of developmental language* *disorders: Theoretical issues and clinical implications* (pp. 309-326). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bishop, D.V.M. (2008). Specific language impairment, dyslexia and autism: Using genetics to unravel their relationship. In C.F. Norbury, J.B. Tomblin, & D.V.M. Bishop *(*Eds.), *Understanding Developmental Language Disorders* (pp. 67-78), Sussex: Psychology Press.

Bishop, D.V.M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders*, *49*(4), 381-415.

Bishop, D.V.M., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders*, *52*(2), 156-173.

Bishop, D.V.M., & Norbury, C.F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology I: Generativity. *Autism*, *9*(1), 7-27.

Bishop, D.V.M., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental Medicine & Child Neurology*, *37*(1), 56-71.

Bishop, D.V.M., & Snowling, M.J. (2004). Developmental dyslexia and specific language impairment: Same or different?. *Psychological Bulletin*, *130*(6), 858.

Blankenstijn, C., & Scheper, A. (2004). *Language development in children with psychiatric Impairment.* Utrecht: LOT (Doctoral dissertation, University of Amsterdam).

Blom, E., & Baayen, H.R. (2013). The impact of verb form, sentence position, home language, and second language proficiency on subject–verb agreement in child second language Dutch. *Applied Psycholinguistics*, *34(*4), 777-811.

Blom, E., de Jong, J., Orgassa, A., Baker, A., & Weerman, F. (2013a). Verb inflection in monolingual Dutch and sequential bilingual Turkish–Dutch children with and without SLI. *International Journal of Language & Communication Disorders*, *48*(4), 382-393.

Blom, E., Orgassa, A., & Polišenská, D. (2008a). Flexie in Taalproductie (FlexiT): Elicitatiemateriaal voor de inventarisatie van de mondelinge taalvaardigheid van werkwoordsvervoeging, werkwoordsplaatsing, vervoeging van het bijvoedlijk naamwoord en lidwoordtoekenning bij diverse taalleerders van het Nederlands. [Inflection in language production: Elicitation material for the investigation of speech production data on verb inflection, verb placement, attributive adjectival inflection and definite article assignment in various learner populations of Dutch]. *Amsterdam: University of Amsterdam*.

Blom E., Polišenská D., & Weerman, F. (2008b). Articles, adjectives and age of onset: the acquisition of Dutch grammatical gender, *Second Language Research* 24(3), 297-331**.**

Blom, E., van de Craats, I., & Verhagen, J. (2013b). Introduction. In E. Blom, I. van de Craats, & J. Verhagen (Eds.), *Dummy auxiliaries in first and second language acquisition* (pp. 1-8). Boston/Berlin: Walter de Gruyter, Inc.

Blom, E., & Vasić, N. (2011). The production and processing of determiner–noun agreement in child L2 Dutch. *Linguistic Approaches to Bilingualism*, *1*(3) 265-290.

Blom, E., Vasić, N., & de Jong, J. (2014). Production and processing of subject–verb agreement in monolingual Dutch children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *57*(3), 952-965.

Boef, E.P.M. (2008). Variatie in Nederlandse lange-afstandsrelatieven, *Taal & Tongval*, *60*, 90-120.

Boef, E.P.M. (2012). *Doubling in relative clauses. Aspects of morphosyntactic microvariation in Dutch*. Utrecht: LOT (Doctoral Dissertation, Utrecht University).

Bol, G.W., & Kuiken, F. (1988). *Grammaticale analyse van taalontwikkelingsstoornissen.* Utrecht: Elinkwijk (Doctoral Dissertation, University of Amsterdam).

Booij, G. (1995). *The phonology of Dutch* (The phonology of the world's languages). Oxford: Clarendon.

Booij, G. (2002). *The morphology of Dutch.* Oxford: Oxford University Press.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry*, *46*(3), 317-326.

Botting, N. (2006). ADHD symptoms in children with a history of SLI: relationship with language skills, *Posterpresented at Child Language Seminar, Newcastle-upon-Tyne.*

Botting, N. (2014). Inaugural lecture on *Atypical language development in children,* City University London, England.

Botting, N., Faragher, B., Simkin, Z., Knox, E., & Conti-Ramsden, G. (2001). Predicting pathways of specific language impairment: what differentiates good and poor outcome? *Journal of Child Psychology and Psychiatry*, *42*(8), 1013-1020.

Botwinik-Rotem, I., & Friedmann, N. (2008). Linguistic bases of child language disorders. In R.G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 143-173). New York: Psychology Press.

Brosnahan, L.F. (1959). The affricates of the high German consonant shift. *Neophilologus*, *42*(1), 112-123.

Brown, H.D. (1971). Children's comprehension of relativized English sentences. *Child Language*, *42*(6), 1923-1936.

Brown, R. (1973). *A first language: The early stages.* Harvard: Harvard University Press.

Burger, E.A., & Rijpma, G. (1998). Taalvaardigheidsonderzoek bij taalgestoorde leerlingen in het voortgezet onderwijs. *Toegepaste Taalwetenschap in Artikelen*, *60*(3), 29-38.

Carlisle, R.S. (2001). Syllable structure universals and second language acquisition. *International Journal of English Studies*, *1*(1), 1-19.

Case, R., Kurland, D.M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386-404.

Catts, H.W., Adlof, S.M., Hogan, T.P., & Weismer, S.E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research*, *48*(6), 1378-1396.

CBS (Centraal Bureau Statistiek) (2016). *Speciale scholen; leerlingen, schooltype, leeftijd* [Data file]. Retrieved from: http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=37746SOL&D1=0&D2=045&D3=a&D4=7,l&HD=1409291541&HDR=T,G3,G2&STB=G1 (date of retrieval 25-5-2016).

Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes*, *16*(2-3), 113-142.

Chiat, S., & Roy, P. (2007). The preschool repetition test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research*, *50*(2), 429-443.

Chomsky, N. (1977). On wh-movement. In P.W. Culicover, T. Wasow, & A. Akmajian (Eds.). *Formal Syntax* (pp. 71-132), New York: Academic Press.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics*, *10*(2), 151-171.

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders? A follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry*, *46*(2), 128-149.

Cohen, M., Campbell, R., & Yaghmai, F. (1989). Neuropathological abnormalities in developmental dysphasia. *Annals of Neurology*, *25*(6), 567-570.

Contemori, C., & Garraffa, M. (2010). Comparison of modalities in SLI syntax: A study on the comprehension and production of non-canonical sentences. *Lingua*, *120*(8), 1940-1955.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *46*(5), 1029-1037.

Conti-Ramsden, G., Bishop, D.V., Clark, B., Norbury, C.F., & Snowling, M.J. (2014). Specific language impairment (SLI): The internet RALLI campaign to raise awareness of SLI. *Psychology of Language and Communication*, *18*(2), 143-148.

Conti-Ramsden, G., & Botting, N. (2004). Social difficulties and victimization in children with SLI at 11 years of age. *Journal of Speech, Language, and Hearing Research*, *47*(1), 145-161.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001b). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*(6), 741-748.

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001a). Follow-up of children attending infant language units: outcomes at 11 years of age. *International Journal of Language & Communication Disorders*, *36*(2), 207-219.

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research*, *40*(4), 765-777.

Conti-Ramsden, G., & Durkin, K. (2008). Language and independence in adolescents with and without a history of specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *51*(1), 70-83.

Conti-Ramsden, G., & Durkin, K. (2012). Postschool educational and employment experiences of young people with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *43*(4), 507-520.

Conti-Ramsden, G., Durkin, K., & Walker, A.J. (2012). The messages they send: E-mail use by adolescents with and without a history of specific language impairment (SLI). *International Journal of Language & Communication Disorders*, *47*(2), 217-228.

Conti-Ramsden, G., Knox, E., Botting, N., & Simkin, Z. (2002). Focus on practice: Educational placements and National Curriculum Key Stage 2 test outcomes of children with a history of specific language impairment. *British Journal of Special Education*, *29(*2), 76-82.

Conti-Ramsden, G., Mok, P.L., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities*, *34*(11), 4161-4169.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Cornips, L. (2008). Loosing grammatical gender in Dutch: The result of bilingual acquisition and/or an act of identity? *International Journal of Bilingualism*, *12*(1-2), 105-124.

Cowan, N. (1999). An embedded-processes model of working memory (pp. 62-101). In A. Miyake, & P. Shah (Eds.). *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.

Crain, S., McKee, C., & Emiliani, M. (1990). Visiting relatives in Italy. In L. Frazier, & J. de Villiers (Eds.) *Language Processing and Language Acquisition* (pp. 335-356), Dordrecht: Kluwer Academic Publishers.

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

Davison, F.M., & Howlin, P. (1997). A follow-up study of children attending a primary-age language unit. *International Journal of Language and Communication Disorders*, *32(*1), 19-36.

Delage, H., Monjauze, C., Hamann, C., & Tuller, L. (2007). Relative clauses in atypical acquisition of French. *Language Acquisition and Development, Proceedings of Generative Approaches to Language Acquisition (GALA)* (pp. 166-176).

Deevy, P., & Leonard, L.B. (2004). The comprehension of wh-questions in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *47*(4), 802-815.

Dispaldro, M., Leonard, L.B., Corradi, N., Ruffino, M., Bronte, T., & Facoetti, A. (2013). Visual attentional engagement deficits in children with specific language impairment and their role in real-time language processing. *Cortex*, *49*(8), 2126-2139.

Dromi, E., Leonard, L.B., & Shteiman, M. (1993). The grammatical morphology of Hebrew-speaking children with specific language impairment: Some competing hypotheses. *Journal of Speech, Language, and Hearing Research*, *36*(4), 760-771.

Duinmeijer, I. (2007). *Comorbiditeit tussen SLI en ADHD*. (Unpublished bachelor thesis, University of Amsterdam).

Duinmeijer, I. (2013). Persistent problems in SLI: which grammatical problems remain when children grow older? *Linguistics in Amsterdam*, *6*, 28-48.

Duinmeijer, I. (2011). [Pilot of a Dutch translation of the preference task for the elicitation of relative clauses]. Unpublished raw data.

Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders*, *47*(5), 542-555.

Ebbels, S. (2014). Introducing the SLI debate. *International Journal of Language & Communication Disorders*, *49*(4), 377-380.

Elbro, C., Dalby, M., & Maarbjerg, S. (2011). Language-learning impairments: A 30-year follow-up of language-impaired children with and without psychiatric, neurological and cognitive difficulties. *International Journal of Language & Communication Disorders*, *46*(4), 437-448.

Ellis, N.C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(2), 143-188.

van Ewijk, L., & Avrutin, S. (2010). Article omission in Dutch children with SLI: A processing approach. *Entropy*, *12(*4), 798-817.

Eyer, J.A., & Leonard, L.B. (1995) Functional categories and specific language impairment: A case study, *Language Acquisition, 4*(3), 177-203.

Fazio, B.B. (1998). The effect of presentation rate on serial memory in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(6), 1375-1383.

Fee, E.J. (1995). The phonological system of a specifically language-impaired population. *Clinical Linguistics & Phonetics*, *9*(3), 189-209.

Finneran, D.A., Francis, A.L., & Leonard, L.B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *52*(4), 915-929.

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

Fox, B., & Thompson, S. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language*, *66*, 856-870.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory*, *5*, 519-559.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*(1), 67-88.

Friedmann, N., & Haddad-Hanna, M. (2014). The comprehension of sentences derived by syntactic movement in Palestinian Arabic speakers with hearing impairment. *Applied Psycholinguistics*, *35*(3), 473-513.

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*(3), 661-681.

Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? Brain and Language, *101*, 50-63.

Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua, 121,* 367-382.

Friedmann, N., & Szterman, R. (2006). Syntactic movement in orally-trained children with hearing impairment. *Journal of Deaf Studies and Deaf Education, 11,* 56-75.

Gallinat, E., & Spaulding, T.J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1363-1382.

Gathercole, S.E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513.

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*(3), 336-360.

Gillam, R.B., & Hoffman, L.M. (2004). Information processing in children with specific language impairment. In L. Verhoeven, & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp.137-157). Mahwah, New Jersey & London: Lawrence Erlbaum Associates Publishers.

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

de Glopper, K., Damhuis, R., de Boers, M., & Kienstra, M. (1992). *Woordenlijst voor 4- tot 6-jarigen: een streeflijst voor kleuters.* Rotterdam: Projectbureau OVB Rotterdam.

Goodluck, H., & Tavakolian, S. (1982). Competence and processing in children's grammar of relative clauses. *Cognition*, *11*(1), 1-27.

Gopnik, M. (1990). Feature blindness: A case study. *Language Acquisition*, *1*(2), 139-164.

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

Graf Estes, K.G., Evans, J.L., & Else-Quest, N.M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *50*(1), 177-195.

Gray, S. (2003). Diagnostic accuracy and test–retest reliability of nonword repetition and digit span tasks administered to preschool children with specific language impairment. *Journal of Communication Disorders*, *36*(2), 129-151.

Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *49*(5), 955-969.

Gregg, K.R. (1990). The variable competence model of second language acquisition, and why it isn't. *Applied linguistics*, *11*(4), 364-383.

Grela, B.G., & Leonard, L.B. (2000). The influence of argument-structure complexity on the use of auxiliary verbs by children with SLI. *Journal of Speech, Language, and Hearing Research*, 43(5), 1115-1125.

van Groningen, R. (2010). *Complexity and correctness in older children with SLI: Related or not*? (MA thesis, University of Amsterdam).

de Haan, A. (1996). *De verwerving van morfologische finietheid in het Nederlands*. (MA Thesis, Groningen University).

Hackman, D.A., Gallop, R., Evans, G.W., & Farah, M.J. (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science*, *18*(5), 686-702.

Haendler, Y., Kliegl, R., & Adani, F. (2015). Discourse accessibility constraints in children's processing of object relative clauses. *Frontiers in Psychology*, *6*, 1-17

Haeseryn, W., K. Romijn, G. Geerts, J. de Rooij, & M.C. van den Toorn (1997). *Algemene Nederlandse Spraakkunst (ANS*). Groningen/Deurne: Martinus Nijhoff.

Håkansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, *27*(2), 313-333.

Hall, P.K., & Tomblin, J.B. (1978). A Follow-up study of children with articulation and language disorders. *Journal of Speech and Hearing Disorders*, 43(2), 227-241.

Hamann, C. (2006). Speculations about early syntax: The production of wh-questions by normally developing French children and French children with SLI. *Catalan Journal of Linguistics*, *5*(1), 143-189.

Hamburger, H., & Crain, S. (1982). Relative acquisition. In S. Kuczaj (Ed.), *Language development.* (Volume 1: Syntax and Semantics) (248-274). Hillsdale, NJ: Erlbaum.

Hasson, N., Dodd, B., & Botting, N. (2012). Dynamic assessment of sentence structure (DASS): Design and evaluation of a novel procedure for the assessment of syntax in children with language impairments. *International Journal of Language & Communication Disorders*, *47*(3), 285-299.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge University Press.

Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across grammars. *Language*, 244-285.

Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. *Language Acquisition and Language Disorders*, *32*, 175-206.

Henry, L.A. (2001). How does the severity of a learning disability affect working memory performance? *Memory*, *9*(4-6), 233-247.

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.

Hesketh, A. (2006). The use of relative clauses by children with language impairment. *Clinical Linguistics and Phonetics*, *20*(7-8), 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*(5), 443-456.

Hick, R.F., Botting, N., & Conti-Ramsden, G. (2005). Short-term memory and vocabulary development in children with Down syndrome and children with specific language impairment. *Developmental Medicine & Child Neurology*, *47*(8), 532-538.

Hitch, G.J., & McAuley, E. (1991). Working memory in children with specific arithmetical learning difficulties. *British Journal of Psychology*, *82*(3), 375-386.

Hockett, C. F. (1958). A course in modern linguistics. *Language Learning*, *8*(3-4), 73-75.

Hoek, P. (2010). *Older Dutch children with specific language impairment: Constructing an error profile*. (MA thesis, University of Amsterdam)

Hoffman, L.M., & Gillam, R.B. (2004). Verbal and spatial information processing constraints in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *47*(1), 114-125.

de Hoop, H., & Lamers, M. (2006). Incremental distinguishability of subject and object. In L. Kulikov, A. Malchukov, & P. de Swart (Eds.), *Case, valency and transitivity* (pp. 269-287). Amsterdam/Philadelphia: John Benjamins Publishing.

Horohov, J.E., & Oetting, J.B. (2004). Effects of input manipulations on the word learning abilities of children with and without specific language impairment. *Applied Psycholinguistics*, *25*(01), 43-65.

Hsu, H.J., & Bishop, D.V.M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science*, *17*(3), 352-365.

Hughes, D.M., Turkstra, L.S., & Wulfeck, B.B. (2009). Parent and self-ratings of executive function in adolescents with specific language impairment. *International Journal of Language & Communication Disorders*, *44*(6), 901-916.

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development*, *77*(6), 1822-1841.

Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 french. *Lingua*, *121*(3), 339-351.

Jaeger, T.F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 323-335.

Janssen, B.E. (2016). *The acquisition of gender and case in Polish and Russion. A Study of monolingual and bilingual children*. Amsterdam: Pegasus (Doctoral Dissertation, University of Amsterdam).

Jensen de Lopéz, K.J., & Baker, A.E. (2015). Executive functions in the assessment of bilingual children with language impairment. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 277-298). Bristol: Multilingual Matters.

Jensen de Lopéz, K.J., 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*(1), 51-83.

Jensen de López, K., & Søndergaard Knudsen, H.B. (2014). Developmental profiles of Danish SLI children's verbally loaded & non-verbal loaded working memory. Poster presented at *IASCL 2014*, Amsterdam, the Netherlands.

Jernigan, T.L., Hesselink, J.R., Sowell, E., & Tallal, P.A. (1991). Cerebral structure on magnetic resonance imaging in language-and learning-impaired children. *Archives of Neurology*, *48*(5), 539-545.

Joanisse, M.F., & Seidenberg, M.S. (1998). Specific language impairment: A deficit in grammar or processing?. *Trends in Cognitive Sciences*, *2*(7), 240-247.

Johnson, C.J., Beitchman, J.H., Young, A., Escobar, M., Atkinson, L., Wilson, B., …, & Wang, M. (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*(3), 744-760.

Johnson, B.W., & Morris, S.R. (2007). Clinical implications of the effects of lexical aspect and phonology on children's production of the regular past tense. *Child Language Teaching and Therapy*, *23*(3), 287-306.

de Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure*. Groningen: Grodil (Doctoral Dissertation, University of Groningen).

de Jong, J. (2002). Specific language impairment and linguistic explanation. In Y. Levy, & J. Schaeffer (Eds.), *Language competence across populations* (pp. 151-170). Mahwah, NJ: Erlbaum.

de Jong, J. (2015). Elicitation task for subject-verb agreement. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 25-37). Bristol/Buffalo/Toronto: Multilingual matters.

de Jong, J. (to appear). The changing grammatical profile of specific language impairment. To appear in: A. Bar-On, & D. Ravid (Eds.), *Handbook of communication disorders: Theoretical, empirical, and applied linguistic perspectives.* Berlin: Mouton de Gruyter.

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.). Boston/Berlin: Walter de Gruyter Inc.

Jónsdóttir S. (2006). *ADHD and its relationship to comorbidity and gender* (Doctoral Dissertation, University of Groningen).

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122.

Kaan, E. (1997). *Processing subject-object ambiguities in Dutch*. (Doctoral dissertation, University of Groningen).

Kaan, E. (2002). Investigating the effects of distance and number interference in processing subject-verb dependencies: An ERP study. *Journal of Psycholinguistic Research*, *31*(2), 165-193.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *37*(2), 418-421.

van Kampen, J. (1997). *First steps in WH movement* (Doctoral Dissertation, Utrecht University).

Kapa, L.L., & Plante, E. (2015). Executive function in SLI: recent advances and future directions. *Current Developmental Disorders Reports*, *2*(3), 245-252.

Keating, G.D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*, 503–535.

Keenan, E. (1975). Variation in universal grammar. In R.W. Fasold, & R.W. Shuy (Eds.), *Analysing variation in language* (pp. 136-148), Washington DC: Georgetown University Press.

Keenan, E.L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*, 63–99.

Keij, B. (2009). *De verwerving van grammaticaal geslacht in het Nederlands* (MA Thesis, Radboud University Nijmegen).

Keij, B., Cornips, L.M., 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*(4), 379-403.

Kemény, F., & Lukács, Á. (2010). Impaired procedural learning in language impairment: Results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology*, *32*(3), 249-258.

Kester, E.P. (1996). Adjectival inflection and the licensing of empty categories in DP. *Journal of Linguistics*, *32*(1), 57-78.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860-897.

Klein, R. (1974). *Word Order: Dutch Children and their mothers*. Publication No. 9. Amsterdam: Institute for General Linguistics, University of Amsterdam.

Kort, W., Schittekatte, M., & Compaan, E. (2008). *CELF-4-NL: Clinical evaluation of language fundamentals*. Amsterdam: Pearson.

Kort, W., Schittekatte, M., Dekker, P.H., Verhaeghe, P., Compaan, E.L., Bosmans, M., & Vermeir, G. (2005). *WISC-III-NL: Wechsler intelligence scale for children, 3rd edition, Dutch adaptation*. Amsterdam: Harcourt Test Publishers/Dutch Institute for Psychologists.

Koster, J. (1975). Dutch as an SOV language. *Linguistic Analysis*, *1*, 111-136.

Krassowski, E., & Plante, E. (1997). IQ variability in children with SLI: Implications for use of cognitive referencing in determining SLI. *Journal of Communication Disorders*, *30*(1), 1-9.

Laka, I. (2013). Merging from the temporal input: On subject-object asymmetries and ergative language. In M. Piattelli-Palmarini, & R.C. Berwick (Eds.), *Rich languages from poor inputs* (pp. 127-145). Oxford: Oxford University Press.

Larson, V.L., & McKinley N.L. (1995). *Language disorders in older students: Preadolescents and adolescents.* Eau Claire: Thinking Publications.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (1998). Screening for speech and language delay: a systematic review of the literature. *Health Technology Assessment*, *2*(9), 1–185.

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: Literacy, mental health, and employment outcomes. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1401-1416.

van der Lely, H.K.J. (1998). SLI in children: Movement, economy, and deficits in the computational-syntactic system. *Language acquisition*, *7*(2-4), 161-192.

van der Lely, H.K.J. (2005). Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in Cognitive Sciences*, 9(2), 53-59.

van der Lely, H.K.J., & Battell, J. (2003). Wh-movement in children with grammatical SLI: A test of the RDDR hypothesis. *Language*, *79(*1), 153-181.

van der Lely, H.K.J., & Ullman, M.T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, *16*(2-3), 177-217.

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, Massachusetts, London: The MIT Press.

Leonard, L.B. (2000). Specific language impairment across languages. In D.V.M. Bishop (Ed.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp., 115-129). East Sussex: Psychology Press.

Leonard, L.B. (2009). Cross-linguistic studies of child language disorders. In R.G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 308-324). New York: Psychology Press.

Leonard, L.B. (2014). *Children with specific language impairment*. Cambridge, Massachusetts, London: The MIT Press.

Leonard, L.B., Bortolini, U., Caselli, M.C., McGregor, K.K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, *2*(2), 151-179.

Leppänen P.H.T., Lyytinen H., Choudhury N., & Benasich A.A. (2005). Neuroimaging measures in the study of specific language impairment. In L. Verhoeven, & H. van Balkom (Eds.), *Classification of developmental language disorders: theoretical issues and clinical implications* (pp. 99-136). Mahwah, New Jersey: Lawrence Erlbaum Associates.

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

Lewis, R.L., Vasishth, S., & Van Dyke, J.A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447-454.

Love, T., & Swinney, D. (1996). Coreference processing and levels of analysis in object-relative constructions; demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research*, *25*(1), 5-24.

Love, A.J., & Thompson M.G. (1988). Language disorders and attention deficit disorders in young children referred for psychiatric services: analysis of prevalence and a conceptual synthesis, *American Journal of Orthopsychiatry, 58*(1), 52-64.

Lukács, Á., Ladányi, E., Fazekas, K., & Kemény, F. (2016). Executive functions and the contribution of short-term memory span in children with specific language impairment. *Neuropsychology, 30*(3), 296-303.

Lum, J.A., Conti-Ramsden, G., Morgan, A.T., & Ullman, M.T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex, 51*, 1-10.

Lum, J. A., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders*, *45*(1), 96-107.

Luria, A.R., Pribram, K.H., & Homskaya, E.D. (1964). An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). *Neuropsychologia*, *2*(4), 257-280.

MacDonald, M.C., & Christiansen, M.H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996), *Psychological Review, 109*(1), 35-54.

Mak, W.M. (2001). *Processing relative clauses: Effects of pragmatic, semantic, and syntactic variables* (Doctoral dissertation, Radboud University Nijmegen).

Mak, W.M., Vonk W., & Schriefers, H. (2002). The influence of animacy on relative clause processing, *Journal of Memory and Language, 47*, 50-68.

Mak, W.M., Vonk W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush, *Journal of Memory and Language, 43*, 466-490.

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*(2), 202-214.

Marchman, V.A., Wulfeck, B., & Weismer, S.E. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research*, *42*(1) 206-219.

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*(6), 517-533.

Marinis, T., Blom, E., Chondrogianni, V., & Vasić, N. (2010). Production of gender and processing of gender mismatch in Dutch and Greek sequential bilingual children. Paper presented at the colloquium *Child Bilingualism: Bilingual first language acquisition or child second language acquisition?* Meertens Institute (November), Amsterdam, The Netherlands.

Marinis, T., Blom, E., & Unsworth, S. (2010). Using on-line processing methods in language acquisition research. In E. Blom, & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 139-162). Amsterdam/ Philadelphia: John Benjamins.

Marshall, C.R., & van der Lely, H.K.J. (2007). The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Advances in Speech Language Pathology*, *9*(3), 191-203.

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Disorders*, *43*(2), 181-200.

Marton, K., Kelmenson, L., & Pinkhasova, M. (2007). Inhibition control and working memory capacity in children with SLI. *Psychologia*, *50*(2), 110–121.

Marton, K., & Schwartz, R.G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *46*(5), 1138-1153.

Masterson, J.J., & Kamhi, A.G. (1992). Linguistic trade-offs in school-age children with and without language disorders. *Journal of Speech, Language, and Hearing Research*, *35*(5), 1064-1075.

Matthews, P.H. (2007). *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.

McGregor, K.K., Newman, R.M., Reilly, R.M., & Capone, N.C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *45*(5), 998-1014.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A.D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*, 477-494.

van der Meer, M., van Atteveldt, W., Coopmans, P., & Philip, W. (2001). Subject-object asymmetry in Dutch children's comprehension of wie-questions. *Linguistics in the Netherlands*, *18*(1), 167-176.

Meir, N. (to appear). Effects of specific language impairment (SLI) and bilingualism on verbal short-term memory. *Linguistic Approaches to Bilingualism*.

Menezes, C.G.D.L., Takiuchi, N., & Befi-Lopes, D.M. (2007). Visual short-term memory in children with specific language impairment. *Pró-Fono Revista de Atualização Científica*, *19*(4), 363-369.

Metz, M., van Hout, A., & van der Lely, H.K.J. (2012). Understanding *who* and *which* questions in five to nine-year-old Dutch children: The role of number. *GAGL*, *51*, 27-41.

Miller, C.A., Kail, R., Leonard, L.B., & Tomblin, J.B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*(2), 416-433.

Miller, C.A., Leonard, L.B., & Finneran, D. (2008). Grammaticality judgments in adolescents with and without language impairment. *International Journal of Language and Communication Disorders*, *43*(3), 346-360.

Mills, A.E. (1985). The acquisition of German. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 141–254). Hillsdale, NJ: Lawrence Erlbaum.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Mok, P.L., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2014). Longitudinal trajectories of peer relations in children with specific language impairment. *Journal of Child Psychology and Psychiatry*, *55*(5), 516-527.

Munson, B., Kurtz, B.A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*(5), 1033-1047.

Nickisch, A., & von Kries, R. (2009). Short-term memory (STM) constraints in children with specific language impairment (SLI): are there differences between receptive and expressive SLI?. *Journal of Speech, Language, and Hearing Research*, *52*(3), 578-595.

Nippold, M.A. (1995). Language norms in school-age children and adolescents: An introduction. *Language, Speech, and Hearing Services in Schools*, *26*(4), 307-308.

Nippold, M.A., Mansfield, T.C., Billow, J.L., & Tomblin, J.B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development. *American Journal of Speech-Language Pathology*, *17*(4), 356-366.

Nippold, M.A., & Schwarz, I.E. (2002). Do children recover from specific language impairment? *Advances in Speech Language Pathology*, *4*(1), 41-49.

Norbury, C.F., Bishop, D.V., & Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *Journal of Speech, Language, and Hearing Research*, *44*(1), 165-178.

Noterdaeme, M., Amorosa, H., Mildenberger, K., Sitter, S., & Minow, F. (2001). Evaluation of attention problems in children with autism and children with a specific language disorder. *European Child & Adolescent Psychiatry*, *10*(1), 58-66.

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

Oetting, J.B., & Horohov, J.E. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, *40*, 62-74.

O'Grady, W., Lee, M., & Choo, M. (2003). A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition*, *25*, 433-448.

Ooi, C.C.W., & Wong, A.M.Y. (2014). Phonological short-term memory and relative clause comprehension in Cantonese-speaking children with SLI. *Paper presented at Experimental Psycholinguistics Conference (ERP), Madrid, Spain.*

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

Orgassa, A., & Weerman, F. (2008). Dutch gender in specific language impairment and second language acquisition. *Second Language Research*, *24*(3), 333-364.

Parigger, E.M. (2012). *Language and executive functioning in children with ADHD*. Vianen: BOXPress (Doctoral Dissertation, University of Amsterdam).

Parigger E.M., & Baker A. (2005). Taal en aandachtstekortstoornissen met hyperactiviteit, *Logopedie en Foniatrie*, *77*(6), 204-211.

Pennington, B.F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(1), 51-87.

Plante, E. (1991). MRI findings in the parents and siblings of specifically language-impaired boys. *Brain and Language*, *41*(1), 67-80.

Polišenská, D. (2010). *Dutch children's acquisition of verbal and adjectival inflection.* Utrecht: LOT (Doctoral Dissertation, University of Amsterdam).

Poll, G.H., Betz, S.K., & Miller, C.A. (2010). Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language, and Hearing Research*, *53*(2), 414-429.

Poll, G.H., Miller, C.A., & van Hell, J.G. (2015). Evidence of compensatory processing in adults with developmental language impairment: Testing the predictions of the procedural deficit hypothesis. *Journal of Communication Disorders*, *53*, 84-102.

del Prado Martín, F.M., Kostić, A., & Baayen, R.H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition, 94*(1), 1-18.

Rapin, I. (1996). Practitioner review: Developmental language disorders: A clinical update. *Journal of Child Psychology and Psychiatry*, *37*(6), 643-655.

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

Raven, J., Raven, J.C., & Court, J.H. (2003). *Manual for Raven's progressive matrices and vocabulary scales*. San Antonia: Harcourt Assessment.

Reed, V.A. (2005). *An introduction to children with language disorders*. Boston: Pearson/Allyn and Bacon, 168-219.

Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F.K., Morgan, A., …, & Wake, M. (2014). Specific language impairment: a convenient label for whom? *International Journal of Language & Communication Disorders*, *49*(4), 416-451.

Rescorla, L. (2000). Do late-talking toddlers turn out to have reading difficulties a decade later? *Annals of Dyslexia*, *50*(1),85-102.

Resing, W.C.M., Evers. A., Koomen, H.M.Y., Pameijer, N.K., & Bleichrodt, N. (2005). *Indicatiestelling speciaal onderwijs en leerlinggebonden financiering: Condities en instrumentarium.* Amsterdam: Boom Test Uitgevers.

Rice, M.L., Hoffman, L., & Wexler, K. (2009). Judgments of omitted BE and DO in questions as extended finiteness clinical markers of specific language impairment (SLI) to 15 years: A study of growth and asymptote. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1417-1433.

Rice, M.L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, *39*(6), 1239-1257.

Richardson, K., & Klecan-Aker, J. S. (2000). Teaching pragmatics to language-learning disabled children: A treatment outcome study. *Child Language Teaching and Therapy*, *16*(1), 23-42.

Rispens, J., & Baker, A. (2012). Nonword repetition: The relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research*, *55*(3), 683-694.

Rispens, J., Baker, A., & Duinmeijer, I. (2015). Word recognition and nonword repetition in children with language disorders: The effects of neighborhood density, lexical frequency, and phonotactic probability. *Journal of Speech, Language, and Hearing Research*, *58*(1), 78-92.

Rispens, J., & Been, P. (2007). Subject–verb agreement and phonological processing in developmental dyslexia and specific language impairment (SLI): A closer look. *International Journal of Language & Communication Disorders*, *42*(3), 293-305.

Roberts, J., Rescorla, L., Giroux, J., & Stevens, L. (1998). Phonological skills of children with specific expressive language impairment (SLI-E): Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, *41*(2), 374-384.

Roodenburg, J., & Hulk, A. (2008). Puzzles on grammatical gender. *Lingue e linguaggio*, *7*(1), 67-92.

Rothweiler, M. (1993). *Der Erwerb von Nebensätzen im Deutschen. Eine Pilotstudie*. Tübingen: Max Niemeyer.

Roulet-Amiot, L., & Jakubowicz, C. (2006). Production and perception of gender agreement in French SLI. *International Journal of Speech-Language Pathology*, *8*(4) 335-346.

Sabourin, L., Stowe, L.A., & De Haan, G.J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *22*(1), 1-29.

Sagarra, N., & Herschensohn, J. (2013). Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism*, *17*(5), 607-627.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*(8), 2022-2039.

Sauerland, U., Grohmann, K., Guasti, M.T., Andjelkovic, D., Argus, R., Armon-Lotem, S., … , & Yatsushiro K. (2015). How do 5 year olds understand questions? Differences in languages across Europe. *First Language*, *36*(3), 169-202.

Schouwenaars, A. (2012). *Word order overrules number agreement: Dutch children's interpretation and production of which-questions* (MA thesis, University of Groningen).

Schouwenaars, A., van Hout, A., & Hendriks, P. (2014). Word order overrules number agreement: Dutch children's interpretation and production of which-questions. *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition-North America*, Somerville: Cascadilla Proceedings Project.

Schraw, G., & McCrudden, M. (2013). Information processing theory. *Information Processing Theory*, 420-433.

Schuele, M., & Nicholls, L.M. (2000). Relative clauses: evidence of continued linguistic vulnerability in children with specific language impairment, *Clinical Linguistics & Phonetics, 14*(8), 563-585.

Schuele, C.M., & Tolbert, L. (2001). Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics & Phonetics*, *15*(4), 257-274.

Schul, R., Stiles, J., Wulfeck, B., & Townsend, J. (2004). How 'generalized' is the 'slowed processing' in SLI? The case of visuospatial attentional orienting. *Neuropsychologia*, *42*(5), 661-671.

Schwartz, R.G. (2009). *Handbook of child language disorders*, New York: Psychology Press.

Scott, C.M., & Stokes, S.L. (1995). Measures of syntax in school-age children and adolescents. *Language, Speech, and Hearing Services in Schools*, *26*(4), 309-319.

Seidel, C.M. (2013). *Creating an SLI performance profile with load* (Doctoral dissertation, Louisiana State University).

Seidl, A., Hollich, G., & Jusczyk, P.W. (2003). Early understanding of subject and object wh-questions. *Infancy*, *4*(3), 423-436.

Semel, E.M., Wiig, E.H., & Secord, W. (2006). *CELF 4: clinical evaluation of language Fundamentals*. San Antonio, TX: The Psychological Corporation.

Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake, & P. Shah (Eds.). *Models of working memory: mechanisms of active maintenance and executive control* (pp. 1-27). Cambridge: Cambridge University Press.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English, *Journal of Verbal Learning and Verbal Behavior, 13,* 272-281.

Snowling, M.J., Bishop, D.V.M., & Stothard, S.E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry*, *41*(5), 587-600.

Snowling, M.J., Bishop, D.V.M., Stothard, S.E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry*, *47*(8), 759-765.

Song, J.Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular–s. *Journal of Speech, Language, and Hearing Research*, *52*(3), 623-642.

Souman, A., & Gillis, S. (2007). A longitudinal study of the acquisition of diminutives in Dutch. In I. Savickienė, & W.U. Dressler (Eds.), *The acquisition of diminutives: A cross-linguistic perspective* (*Language Acquisition & Language Disorders series).* Amsterdam: John Benjamins Publishing Company.

Spaulding, T.J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research*, *51*(1), 16-34.

Stark, R., Bernstein, L., Condino, R., Bender, M., Tallal, P., & Catts, H. (1984). Four-year follow-up study of language impaired children. *Annals of Dyslexia*, *34*(1), 49-68.

Stark, R. E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders*, *46*(2), 114-122.

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language*, *77*(3), 419-431.

Stavrakaki, S. (2006). Developmental perspectives on specific language impairment: Evidence from the production of wh-questions by Greek SLI children over time. *Advances in Speech Language Pathology*, *8*(4), 384-396.

Steenge, J. (2006). *Bilingual children with specific language impairment: additionally disadvantaged?*. Nijmegen: EAC, Research Centre on Atypical Communication (Doctoral Dissertation, Radboud University Nijmegen).

Stewart, J., & Sinclair, H. (1975). Comprehension of questions by children between 5 and 9. *Linguistics, 151,* 17-26.

Stothard, S.E., Snowling, M.J., 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*(2), 407-418.

Strangmann, I., Slomp A., & van Hout, A. (2014) Development of Dutch children's comprehension of subject and object wh-questions: the role of topicality. *Linguistics in the Netherlands*, *31*(1), 129-144.

Szterman, R., & Friedmann, N. (2014). Relative clause reading in hearing impairment: different profiles of syntactic impairment. *Frontiers in Psychology*, *5*, 1-25.

Tallal, P., & Piercy, M. (1973). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, *11*(4), 389-398.

Tannock R., & Schachar R**.** (1996). Executive dysfunction as an underlying mechanism of behavior and language problems in attention deficit hyperactivity disorder. In J.H. Beitchman, N. Cohen, M.M. Konstantareas, & R. Tannock *(Eds.), Language, learning and behaviour disorders: developmental, biological and clinical perspectives* (pp. 128-155). Cambridge University Press: Cambridge.

Thiel, A., Sanfelici, E., Koch, C., & Schulz, P. (2014). Children’s production of subject vs. object relative clauses: A developmental study on German, *Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft*, 36, Universität Marburg.

Tomblin, J.B. (2009). Genetics of child language disorders. In R.G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 232-256), New York: Psychology Press.

Tomblin, J.B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development*, *3*(4), 269-293.

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*(6), 1245-1260.

Towse, J.N., Hitch, G.J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language*, *39*(2), 195-217.

Traxler, M.J., Morris, R.K., & Seely, R.E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*(1), 69-90.

Tropper, B., & Schwartz, R.G. (2009). Neurobiology of child language disorders. In R.G. Schwarz (Ed.), *Handbook of child language disorders* (pp. 174-200), New York: Psychology Press.

Tsimpli, I.M., & Hulk, A. (2013). Grammatical gender and the notion of default: Insights from language acquisition. *Lingua*, *137*, 128-144.

Tuller, L., Henry, C., Sizaret, E., & Barthez, M. A. (2012). Specific language impairment at adolescence: Avoiding complexity. *Applied Psycholinguistics*, *33*(1), 161-184.

Tyack, D., & Ingram, D. (1977). Children's production and comprehension of questions. *Journal of Child Language*, *4*(2), 211-224.

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

Ullman, M.T., & Pierpont, E.I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*(3), 399-433.

Unsworth, S. (2008). Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research*, *24*(3), 365-395.

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., & Tsimpli, I. (2014). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics*, *35*(04), 765-805.

Vasishth, S., & Lewis, R.L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, *82*(4), 767-794.

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*(3), 1186-1193.

Verrips, M. (1996). *Potatoes must peel. The acquisition of the Dutch passive* (Doctoral dissertation, University of Amsterdam).

de Vincenzi, M., Arduino, L.S., Ciccarelli, L., & Job, R. (1999). Parsing strategies in children comprehension of interrogative sentences. In *Proceedings of European Conference on Cognitive Science* (pp. 301-308). Istituto di Psicologia del CNR, Rome.

Vugs, B., Hendriks, M., Cuperus, J., & Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Research in Developmental Disabilities*, *35*(1), 62-74.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79-112.

van Weerdenburg, M., Verhoeven, L., & Van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry*, *47*(2), 176-189.

Weerman, F.P., Duinmeijer, I., & Orgassa, A. (2011). Effecten van SLI op Nederlandse congruentie. *Nederlandse Taalkunde*, *16*(1), 30-55.

Weerman, F.P., Punt, L., & Bisschop, J. (2006). L1 and L2 acquisition of Dutch adjectival inflection. *ACLC Working Papers*, *1*(1), 5-36.

Weiner, P.S. (1974). A language-delayed child at adolescence. *Journal of Speech and Hearing Disorders*, *39*(2), 202-212.

Weismer, S.E. (1996). Capacity limitations in working memory: The impact on lexical and morphological learning by children with language impairment. *Topics in Language Disorders*, *17*(1), 33-44.

Weismer, S.E., Evans, J., & Hesketh, L.J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1249-1260.

Weismer, S.E., Plante, E., Jones, M., & Tomblin, J.B. (2005). A functional magnetic resonance imaging investigation of verbal working memory in adolescents with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*(2), 405-425.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*(1), 23-79.

Wexler, K., Schaeffer, J., & Bol, G. (2004). Verbal syntax and morphology in typically developing Dutch children and children with SLI: How developmental data can play an important role in morphological theory. *Syntax*, *7*(2), 148-198.

Wijnen, F. (2000). Input, intake and sequence in syntactic development. *From sound to sentence*, 163-186.

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.

Wilsenach, A.C. (2006). *Syntactic processing in developmental dyslexia and in specific language impairment: A study on the acquisition of the past participle construction in Dutch.* Utrecht: LOT (Doctoral Dissertation, Utrecht University).

Windsor, J., & Hwang, M. (1999). Testing the generalized slowing hypothesis in specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1205-1218.

Windsor, J., Kohnert, K., Loxtercamp, A. L., & Kan, P. (2008). Performance on nonlinguistic visual tasks by children with language impairment. *Applied Psycholinguistics*, *29*(2), 237.

Wittke, K., Spaulding, T.J., & Schechtman, C.J. (2013). Specific language impairment and executive functioning: Parent and teacher ratings of behavior. *American Journal of Speech-Language Pathology, 22*, 161–172.

Yoon, J., Campanelli, L., Goral, M., Marton, K., Eichorn, N., & Obler, L.K. (2015). The effect of plausibility on sentence comprehension among older adults and its relation to cognitive functions. *Experimental Aging Research*, *41*(3), 272-302.

Young, A.R., Beitchman, J.H., Johnson, C., Douglas, L., Atkinson, L., Escobar, M., & Wilson, B. (2002). Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *Journal of Child Psychology and Psychiatry*, *43*(5), 635-645.

Zuckerman, S. (2001). *The acquisition of "optional" movement* (Doctoral dissertation, University of Groningen).

Zuckerman, S. (2013). Dummy auxiliaries in Dutch first language acquisition. In E. Blom, I. van de Craats, & J. Verhagen (Eds.), *Dummy auxiliaries in first and second language acquisition* (pp. 39-74). Boston/Berlin: Walter de Gruyter, Inc.

Zwitserlood, R. (2014). *Language growth in Dutch school-age children with specific language impairment*. Utrecht: LOT (Doctoral Dissertation, Utrecht University).

Zwitserlood, R., Wijnen, F., van Weerdenburg, M., & Verhoeven, L. (2015a). 'MetaTaal': enhancing complex syntax in children with specific language impairment - a metalinguistic and multimodal approach. *International Journal of Language & Communication Disorders*, *50*(3), 273-297.

Zwitserlood, R., van Weerdenburg, M., Verhoeven, L., & Wijnen, F. (2015b). Development of morphosyntactic accuracy and grammatical complexity in Dutch school-age children with SLI. *Journal of Speech, Language, and Hearing Research*, *58*(3), 891-905.

Appendices

A. **List of items judgement grammatical gender**

B. **List of items elicitation grammatical gender**

C. **List of items judgement subject verb agreement**

D. **List of items elicitation game subject verb agreement**

E. **Pictures elicitation game subject verb agreement**

F. **List of items elicitation task subject verb agreement with different syntactic contexts**

G. **List of items comprehension relative clauses**

H. **List of items judgement relative clause**

I. **List of items elicitation relative clauses**

J. **List of codings for the elicitation of relative clauses**

A. List of items judgement task grammatical gender

DET	=	determiner	ADJ	=	adjective
N	=	neuter	C	=	common
H	=	high frequency	L	=	low frequency

1	*een groot eiland*	ADJ	N
2	*het hert*	DET	N
3	*een rode raket*	ADJ	C
4	*de bijl*	DET	C
5	*een rood fiets*	ADJ	C
6	*een blauw helm*	ADJ	C
7	*een groene raam*	ADJ	N
8	*de net*	DET	N
9	*een rood huis*	ADJ	N
10	*een rode net*	ADJ	N
11	*de zadel*	DET	N
12	*de schommel*	DET	C
13	*het bijl*	DET	C
14	*de potlood*	DET	N
15	*een groen bijl*	ADJ	C
16	*het eiland*	DET	N
17	*de hert*	DET	N
18	*een rood net*	ADJ	N
19	*het schommel*	DET	C
20	*het cadeau*	DET	N
21	*een witte cadeau*	ADJ	N
22	*een groene bijl*	ADJ	C
23	*het gitaar*	DET	C
24	*een groene boom*	ADJ	C
25	*een groen zadel*	ADJ	N
26	*het net*	DET	N
27	*een wit cadeau*	ADJ	N
28	*een bruin zwaard*	ADJ	N
29	*een groen boom*	ADJ	C
30	*het draak*	DET	C
31	*de huis*	DET	N
32	*het huis*	DET	N
33	*een groot gitaar*	ADJ	C
34	*het raket*	DET	C
35	*het boom*	DET	C
36	*de ballon*	DET	C
37	*een grote schommel*	ADJ	C
38	*een grote eiland*	ADJ	N
39	*een groene zadel*	ADJ	N
40	*het ballon*	DET	C
41	*de draak*	DET	C
42	*de cadeau*	DET	N
43	*het fiets*	DET	C
44	*een groot poes*	ADJ	C
45	*het helm*	DET	C
46	*een witte ballon*	ADJ	C
47	*het zwaard*	DET	N
48	*een rode huis*	ADJ	N
49	*de paard*	DET	N
50	*het zadel*	DET	N
51	*het potlood*	DET	N
52	*een bruine draak*	ADJ	C
53	*de fiets*	DET	C
54	*een bruine zwaard*	ADJ	N
55	*een blauwe helm*	ADJ	C
56	*de raket*	DET	C
57	*een bruin draak*	ADJ	C
58	*een groot hert*	ADJ	N
59	*een blauwe potlood*	ADJ	N
60	*een grote poes*	ADJ	C
61	*het raam*	DET	N
62	*de gitaar*	DET	C
63	*de poes*	DET	C
64	*de raam*	DET	N
65	*een grote hert*	ADJ	N
66	*een rode fiets*	ADJ	C
67	*een blauw potlood*	ADJ	N
68	*een wit ballon*	ADJ	C
69	*de zwaard*	DET	N
70	*een rood raket*	ADJ	C
71	*een grote gitaar*	ADJ	C
72	*de boom*	DET	C
73	*de eiland*	DET	N
74	*het poes*	DET	C
75	*een grote paard*	ADJ	N
76	*de helm*	DET	C
77	*het paard*	DET	N
78	*een groot schommel*	ADJ	C
79	*een groen raam*	ADJ	N
80	*een groot paard*	ADJ	N

B. List of items grammatical gender judgement task

DET	=	determiner
ADJ	=	adjective
N	=	neuter
C	=	common
H	=	high frequency
L	=	low frequency
0	=	adjacent determiner and noun
1	=	1 adjective between determiner and noun /1 adjective to inflect
2	=	2 adjectives between determiner and noun /2 adjectives to inflect

nr.	***Elicitation cue***	**Target**				
1	*Kijk, de ster en een fiets. De ster staat voor….*	de fiets	DET	C	H	0
2	*Kijk, een hand en een zadel. De hand pakt….*	het zadel	DET	N	L	0
3	*Kijk, een baby en een ballon. De baby hangt aan…*	de ballon	DET	C	H	0
4	*De ster staat op….*	het zwaard	DET	N	L	0
5	*Kijk, een jongen en een fiets. De jongen zit op…*	de fiets	DET	C	H	0
6	*Twee fietsen. dit is een…..*	blauwe fiets	ADJ	C	H	1
7	*Dit is een…..*	rode fiets	ADJ	C	H	1
8	*De kip staat voor….*	de blauwe fiets	DET	C	H	1
9	*De vinger wijst naar….*	de rode fiets	DET	C	H	1
10	*Kijk, de ster en een zadel. De ster hangt boven….*	het zadel	DET	N	L	0
11	*Kijk, de ster met een potlood. De ster leunt tegen...*	het potlood	DET	N	H	0
12	*Kijk, de ster en een ballon. De ster staat op…..*	de ballon	DET	C	H	0
13	*Kijk, een bal en een bijl. De bal ligt naast….*	de bijl	DET	C	L	0
14	*Kijk, de ster en een draak. De ster staat boven….*	de draak	DET	C	L	0
15	*Dit is een….*	groen zadel	ADJ	N	L	1
16	*Dit is een….*	bruin zadel	ADJ	N	L	1
17	*De bloem staat op…..*	het bruine zadel	DET	N	L	1
18	*De aap hangt aan…..*	het groene zadel	DET	N	L	1
19	*Kijk, de ster en een schommel. De ster staat naast…*	de schommel	DET	C	H	0
20	*Kijk, een taart en een zwaard. De taart staat naast…*	het zwaard	DET	N	L	0
21	*Kijk, een bloem en een potlood. De bloem staat naast…*	het potlood	DET	N	H	0
22	*Dit is een….*	gele ballon	ADJ	C	H	1
23	*Dit is een….*	witte ballon	ADJ	C	H	1
24	*De hand pakt...*	de witte ballon	DET	C	H	1
25	*De vinger prikt in…*	de gele ballon	DET	C	H	1
26	*Kijk, de baby en een schommel. De baby kruipt onder….*	de schommel	DET	C	H	0
27	*Kijk, een kip en een draak. De kip kijkt naar….*	de draak	DET	C	L	0
28	*Twee zwaarden! Dit is een…*	groen zwaard	ADJ	N	L	1
29	*Dit is een….*	bruin zwaard	ADJ	N	L	1
30	*De eend kijkt naar…*	het bruine zwaard	DET	N	L	1
31	*De dokter wijst naar….*	het groene zwaard	DET	N	L	1
32	*Kijk, de ster en een bijl. De ster staat onder….*	de bijl	DET	C	L	0
33	*Twee schommels! Dit is een….*	blauwe schommel	ADJ	C	H	1

34	*Dit is een….*	rode schommel	ADJ	C	H	1
35	*De aap hangt aan….*	de rode schommel	DET	C	H	1
36	*De dokter wijst naar….*	de blauwe schommel	DET	C	H	1
37	*Kijk, een vinger en een net. De vinger wijst naar...*	het net	DET	N	L	0
38	*Twee draken! Dit is een….*	bruine draak	ADJ	C	L	1
39	*Dit is een….*	groene draak	ADJ	C	L	1
40	*De vogel vliegt over….*	de bruine draak	DET	C	L	1
41	*De eend kijkt naar….*	de groene draak	DET	C	L	1
42	*Kijk, een zuster en een paard. De zuster staat naast..*	het paard	DET	N	H	0
43	*Twee potloden! Dit is een….*	rood potlood	ADJ	N	H	1
44	*Dit is een….*	blauw potlood	ADJ	N	H	1
45	*De kip staat voor….*	het rode potlood	DET	N	H	1
46	*De dokter wijst naar….*	het blauwe potlood	DET	N	H	1
47	*Kijk, de ster met een net. De ster zit in….*	het net	DET	N	L	0
48	*Kijk, de ster en een raket. De ster hangt aan….*	de raket	DET	C	L	0
49	*Twee bijlen! Dit is een….*	bruine bijl	ADJ	C	L	1
50	*Dit is een….*	groene bijl	ADJ	C	L	1
51	*De dokter wijst naar….*	de bruine bijl	DET	C	L	1
52	*De kip staat op….*	de groene bijl	DET	C	L	1
53	*Kijk, een eend en een raket. De eend kijkt naar….*	de raket	DET	C	L	0
54	*Kijk, de ster en een paard. De ster staat onder….*	het paard	DET	N	H	0
55	*Vier ballonnen! Dit is een…..*	grote gele ballon	ADJ	C	H	2
56	*Dit is een……*	kleine witte ballon	ADJ	C	H	2
57	*De eend kijkt naar….*	de grote witte ballon	DET	C	H	2
58	*De kip kijkt naar….*	de kleine gele ballon	DET	C	H	2
59	*Kijk, de ster en een hert. De ster staat op…..*	het hert	DET	N	L	0
60	*Twee netten! Dit is een…*	rood net	ADJ	N	L	1
61	*Dit is een…*	blauw net	ADJ	N	L	1
62	*De hand pakt….*	het rode net	DET	N	L	1
63	*De bal ligt onder….*	het blauwe net	DET	N	L	1
64	*Vier zadels! Dit is een….*	groen groot zadel	ADJ	N	L	2
65	*Dit is een….*	bruin klein zadel	ADJ	N	L	2
66	*De muis zit op….*	het grote bruine zadel	DET	N	L	2
67	*De kip staat voor….*	het kleine groene zadel	DET	N	L	2
68	*Kijk, een vrouw en een hert. De vrouw kijkt naar…*	het hert	DET	N	L	0
69	*Twee raketten! Dit is een….*	blauwe raket	ADJ	C	L	1
70	*Dit is een….*	rode raket	ADJ	C	L	1
71	*De jongen kijkt naar….*	de blauwe raket	DET	C	L	1
72	*De vrouw kijkt naar….*	de rode raket	DET	C	L	1
73	*Kijk, een man en een gitaar. De man speelt op…*	de gitaar	DET	C	L	0
74	*Twee herten! Dit is een….*	groot hert	ADJ	N	L	1
75	*Dit is een….*	klein hert	ADJ	N	L	1
76	*De jongen kijkt naar….*	het grote hert	DET	N	L	1
77	*De auto staat voor….*	het kleine hert	DET	N	L	1
78	*Kijk, de ster met een gitaar. De ster staat onder….*	de gitaar	DET	C	L	0
79	*Vier netten! Dit is een….*	groot rood net	ADJ	N	L	2
80	*Dit is een…*	klein blauw net	ADJ	N	L	2

81	*De wurm kruipt naast…..*	het kleine rode net	DET	N	L	2
82	*De vis zwemt bij….*	het grote blauwe net	DET	N	L	2
83	*Kijk, de ster en een cadeau. De ster staat op…*	het cadeau	DET	N	H	0
84	*Kijk, een vogel en een huis. De vogel vliegt over…*	het huis	DET	N	H	0
85	*Twee gitaren! Dit is een….*	grote gitaar	ADJ	C	L	1
86	*Dit is een….*	kleine gitaar	ADJ	C	L	1
87	*De jongen kijkt naar….*	de kleine gitaar	DET	C	L	1
88	*De hand pakt….*	de grote gitaar	DET	C	L	1
89	*Kijk, een ster en een boom. De ster hangt boven….*	de boom	DET	C	H	0
90	*Vier potloden! Dit is een….*	rood groot potlood	ADJ	N	H	2
91	*Dit is een…*	klein blauw potlood	ADJ	N	H	2
92	*De bal ligt achter…..*	het kleine rode potlood	DET	N	H	2
93	*De muis zit voor….*	het grote blauwe potlood	DET	N	H	2
94	*Kijk, een eend en een cadeau. De eend kijkt naar…*	het cadeau	DET	N	H	0
95	*Daar is de ster en een poes. De ster staat naast……*	de poes	DET	C	H	0
96	*Twee paarden! Dit is een….*	klein paard	ADJ	N	H	1
97	*Dit is een….*	groot paard	ADJ	N	H	1
98	*De koe staat achter….*	het kleine paard	DET	N	H	1
99	*De vrouw kijkt naar….*	het grote paard	DET	N	H	1
100	*Kijk, de ster en een eiland. De ster hangt boven….*	het eiland	ADJ	N	H	2
101	*Kijk, de ster en een huis. De ster hangt boven…….*	het huis	DET	N	H	0
102	*Vier raketten! Dit is een……*	kleine rode raket	ADJ	C	L	2
103	*Dit is een…..*	grote blauwe raket	ADJ	C	L	2
104	*De vrouw kijkt naar….*	de kleine blauwe raket	DET	C	L	2
105	*De vinger wijst naar….*	de grote rode raket	DET	C	L	2
106	*Kijk, een aap en een boom. De aap hangt aan…*	de boom	DET	C	H	0
107	*Kijk, de ster en een raam. De ster staat achter…*	het raam	DET	N	H	0
108	*Twee huizen! Dit is een….*	blauw huis	ADJ	N	H	1
109	*Dit is een….*	rood huis	ADJ	N	H	1
110	*De bloem groeit voor….*	het rode huis	DET	N	H	1
111	*De eend staat naast….*	het blauwe huis	DET	N	H	1
112	*Hee, een koe en een raam. De koe staat voor….*	het raam	DET	N	H	0
113	*Kijk, een hond en een poes. De hond kijkt naar…*	de poes	DET	C	H	0
114	*Twee bomen! Dit is een….*	bruine boom	ADJ	C	H	1
115	*Dit is een….*	groene boom	ADJ	C	H	1
116	*De koe staat naast….*	de groene boom	DET	C	H	1
117	*De auto staat naast….*	de bruine boom	DET	C	H	1
118	*Kijk, een appel en een helm. De appel ligt onder…*	de helm	DET	C	L	0
119	*Kijk, een kip en een eiland. De kip staat op….*	het eiland	ADJ	N	H	2
120	*Twee ramen! Dit is een….*	bruin raam	ADJ	N	H	1
121	*Dit is een….*	groen raam	ADJ	N	H	1
122	*De taart staat voor….*	het groene raam	DET	N	H	1
123	*De bal ligt onder….*	het bruine raam	DET	N	H	1
124	*Twee poezen! Dit is een….*	kleine poes	ADJ	C	H	1
125	*Dit is een….*	grote poes	ADJ	C	H	1
126	*De bal ligt voor….*	de grote poes	DET	C	H	1
127	*De ster hangt boven….*	de kleine poes	DET	C	H	1

128	*Twee cadeaus! Dit is een….*	geel cadeau	ADJ	N	H	1
129	*Dit is een….*	wit cadeau	ADJ	N	H	1
130	*De vinger wijst naar….*	het gele cadeau	DET	N	H	1
131	*De hand pakt….*	het witte cadeau	DET	N	H	1
132	*Vier bomen! Dit is een….*	kleine bruine boom	ADJ	C	H	2
133	*Dit is een….*	kleine groene boom	ADJ	C	H	2
134	*De muis zit naast…..*	de grote bruine boom	DET	C	H	2
135	*De bal ligt onder…..*	de grote groene boom	DET	C	H	2
136	*Kijk, daar is de ster, met een helm. De ster zit op…*	*de helm*	DET	C	L	0
137	*Twee eilanden! Dit is een….*	klein eiland	ADJ	N	L	1
138	*Dit is een….*	groot eiland	ADJ	N	L	1
139	*De ster hangt boven…..*	het kleine eiland	DET	N	L	0
140	*De vogel vliegt boven….*	het grote eiland	DET	N	L	0
141	*Twee helmen! Dit is een….*	witte helm	ADJ	C	L	1
142	*Dit is een….*	blauwe helm	ADJ	C	L	1
143	*De eend kijkt naar….*	de blauwe helm	DET	C	L	1
144	*De bloem staat onder….*	de witte helm	DET	C	L	1
145	*Vier huizen! Dit is een….*	rood groot huis	ADJ	N	H	1
146	*Dit is een….*	klein blauw huis	ADJ	N	H	1
147	*De vinger wijst naar….*	het grote blauwe huis	DET	N	H	2
148	*De hand pakt….*	het kleine rode huis	DET	N	H	2
149	*Vier helmen! Dit is een….*	grote blauwe helm	ADJ	C	L	2
150	*Dit is een….*	kleine witte helm	ADJ	C	L	2
151	*De wurm kruipt onder…..*	de grote witte helm	DET	C	L	2
152	*De fles staat naast….*	de kleine blauwe helm	DET	C	L	2

C. List of items judgement task subject verb agreement

1SG	=	first person singular	1 PL	=	first person plural
2 SG	=	second person singular	2 PL	=	second person plural
3 SG	=	third person singular	3PL	=	third person plural

1	*jullie poets je tanden*	2PL
2	*jij poets je tanden*	2SG
3	*ik kam een poes*	1SG
4	*jij drinkt melk*	2SG
5	*jullie leest een boek*	2PL
6	*jullie drinkt melk*	2PL
7	*hij kamt een hond*	3SG
8	*wij lezen een boek*	1PL
9	*jullie lees een krant*	2PL
10	*wij drinken melk*	1PL
11	*jij poetst je schoen*	2SG
12	*hij leest een krant*	3SG
13	*wij filmt een bloem*	1PL
14	*ik poetst mijn tanden*	1SG
15	*wij lees een krant*	1PL
16	*jullie bakt een ei*	2PL
17	*ik bak een ei*	1SG
18	*wij bak een pannekoek*	1PL
19	*ik lees een krant*	1SG
20	*jullie filmen een lamp*	2PL
21	*hij poetst zijn tanden*	3SG
22	*wij bakt een ei*	1PL
23	*jullie poetst je schoenen*	2PL
24	*wij poets onze tanden*	1PL
25	*wij poetsen onze schoenen*	1PL
26	*wij filmen een bloem*	1PL
27	*jullie bakken een pannekoek*	2PL
28	*hij poets zijn schoen*	3SG
29	*hij lees een boek*	3SG
30	*jullie kamt een poes*	2PL
31	*jij filmt een bloem*	2SG
32	*jullie bak een pannekoek*	2PL
33	*hij film een bloem*	3SG
34	*jullie filmt een bloem*	2PL
35	*wij kammen een poes*	1PL
36	*jij kamt een poes*	2SG

37	*jullie poetsen je tanden*	2PL
38	*ik lees een boek*	1SG
39	*hij kam een poes*	3SG
40	*wij bakken een ei*	1PL
41	*jullie lezen een krant*	2PL
42	*ik bakt een pannekoek*	1SG
43	*ik film een bloem*	1SG
44	*wij leest een boek*	1PL
45	*hij bakt een pannekoek*	3SG
46	*jij bak een pannekoek*	2SG
47	*jullie drinken water*	2PL
48	*ik kamt een hond*	1SG
49	*wij drink water*	1PL
50	*ik filmt een lamp*	1SG
51	*jullie kammen een hond*	2PL
52	*ik drink melk*	1SG
53	*jij leest een boek*	2SG
54	*wij film een lamp*	1PL
55	*jij kam een hond*	2SG
56	*hij filmt een lamp*	3SG
57	*wij kam een hond*	1PL
58	*hij drinkt water*	3SG
59	*jij bakt een ei*	2SG
60	*jullie film een lamp*	2PL
61	*jullie drink water*	2PL
62	*ik poets mijn schoen*	1SG
63	*wij kamt een poes*	1PL
64	*hij drink melk*	3SG
65	*hij bak een ei*	3SG
66	*wij drinkt melk*	1PL
67	*wij poetst onze schoenen*	1PL
68	*jij lees een krant*	2SG
69	*jij drink water*	2SG
70	*jij film een lamp*	2SG
71	*jullie kam een hond*	2PL
72	*ik drinkt water*	1SG

D. List of items elicitation game subject verb agreement

1SG	=	first person singular	1PL	=	first person plural
2SG	=	second person singular	2PL	=	second person plural
3SG	=	third person singular			

nr.	**Target answers**							
1	*wij bakken een ei*	1PL					*jij kamt een hond*	2SG
2	*wij filmen een lamp*	1PL		*Kim kamt een poes*	3SG			
3	*wij lezen een boek*	1PL		*Kim poetst haar tanden*	3SG			
4	*ik lees een boek*	1SG		*jullie poetsen je tanden*	2PL			
5	*ik poets mijn tanden*	1SG		*jullie bakken een ei*	2PL			
6	*ik kam een hond*	1SG		*jullie filmen een lamp*	2PL			
7	*ik poets mijn tanden*	1SG		*Kim bakt een ei*	3SG		*jij poetst je schoen*	2SG
8	*wij poetsen onze schoen*	1PL		*Kim filmt een bloem*	3SG			
9	*wij kammen een hond*	1PL		*Kim filmt een lamp*	3SG			
10	*ik poets mijn schoen*	1SG		*jullie kammen een hond*	2PL			
11	*ik kam de poes*	1SG		*jullie poetsen je schoen*	2PL			
12	*ik kam een hond*	1SG		*Kim bakt een pannekoek*	3SG		*jij drinkt melk*	2 SG
13	*wij drinken water*	1PL					*jij filmt een lamp*	2SG
14	*ik drink water*	1SG		*jullie lezen een boek*	2PL			
15	*ik bak een ei*	1SG		*Kim poetst haar schoen*	3SG		*jij bakt een pannekoek*	2SG
16	*ik film een bloem*	1SG		*Kim kamt een hond*	3SG		*jij leest de krant*	2SG
17	*wij drinken melk*	1PL					*jij filmt een bloem*	2SG
18	*wij lezen de krant*	1PL					*jij drinkt water*	2SG
19	*ik kam een poes*	1SG		*Kim leest een boek*	3SG		*jij poetst je tanden*	2SG
20	*wij kammen een poes*	1PL					*jij leest een boek*	2SG
21	*ik bak een pannekoek*	1SG		*jullie filmen een bloem*	2PL			
22	*ik film een lamp*	1SG		*jullie kammen een poes*	2PL			
23	*ik lees een boek*	1SG		*jullie bakken een pannekoek*	2PL			
24	*wij bakken een pannekoek*	1PL		*Kim drinkt melk*	3SG			
25	*wij poetsen onze tanden*	1PL		*Kim leest een krant*	3SG			
26	*ik drink melk*	1SG		*jullie lezen een krant*	2PL			
27	*ik bak een ei*	1SG		*Kim drinkt water*	3SG		*jij kamt een poes*	2SG
28	*wij filmen een bloem*	1PL					*jij bakt een ei*	2SG
29	*ik film een bloem*	1SG		*jullie drinken melk*	2PL			
30	*ik lees een krant*	1SG		*jullie drinken water*	2PL			

E. Pictures elicitation game subject verb agreement

Aaien – 'to pet' practice item			
Kammen – 'to comb' sonorant, simple		***Filmen*** – 'to film' sonorant, complex	
Lezen – 'to read' fricative, simple		***Poetsen*** – 'to polish' fricative, complex	
Bakken – 'to bake' plosive, simple		***Drinken*** – 'to drink' plosive, complex	

F. List of items elicitation task subject verb agreement with different syntactic contexts

3SG	=	third person singular
3PL	=	third person plural
SON	=	sonorant final consonant of the verb stem
PLOS	=	plosive final consonant of the verb stem
FRIC	=	fricative final consonant of the verb stem
S	=	simple verb stem coda (1 consonant)
C	=	complex verb stem coda (consonant cluster)
MC	=	main clause
V2	=	inversion
RC	=	relative clause

Nr.	***Elicitation cue***	**Target**				
1	*De vrouw…*	kamt een poes	3SG	SON	S	MC
2	*En de man…*	kamt een hond	3SG	SON	S	MC
3	*Op deze plaatjes…*	drinken de vrouwen water	3PL	PLOS	C	V2
4	*En op deze plaatjes…*	drinken de mannen melk	3PL	PLOS	C	V2
5	*De vrouwen…*	poetsen hun tanden	3PL	FRIC	C	MC
6	*En de mannen…*	poetsen hun schoen	3PL	FRIC	C	MC
7	*Het meisje ziet dat*	de vrouw een ei bakt	3SG	PLOS	S	SC
8	*En ze ziet dat….*	de man een pannekoek bakt	3SG	PLOS	S	SC
9	*Op deze plaatjes…*	poetsen de mannen hun tanden	3PL	FRIC	C	V2
10	*Op deze plaatjes….*	poetsen de vrouwen hun schoen	3PL	FRIC	C	V2
11	*De jongen ziet dat….*	de vrouw een lamp filmt	3SG	SON	C	SC
12	*En hij ziet dat….*	de man een bloem filmt	3SG	SON	C	SC
13	*Op dit plaatje…..*	bakt de vrouw een pannekoek	3SG	PLOS	S	V2
14	*En op dit plaatje….*	bakt de man een ei	3SG	PLOS	S	V2
15	*Op dit plaatje…..*	leest de man een krant	3SG	FRIC	S	V2
16	*En op dit plaatje…..*	leest de vrouw een boek	3SG	FRIC	S	V2
17	*Het meisje ziet dat…..*	de vrouw haar schoen poetst	3SG	FRIC	C	SC
18	*en ze ziet dat….*	de man zijn tanden poetst	3SG	FRIC	C	SC
19	*De mannen….*	bakken een ei	3PL	PLOS	S	MC
20	*En de vrouwen….*	bakken een pannekoek	3PL	PLOS	S	MC
21	*op dit plaatje….*	filmt de vrouw een lamp	3SG	SON	C	V2
22	*en op dit plaatje….*	filmt de man een bloem	3SG	SON	C	V2
23	*Het meisje ziet dat…..*	de vrouwen een boek lezen	3PL	FRIC	S	SC
24	*En ze ziet dat….*	de mannen een krant lezen	3PL	FRIC	S	SC
25	*Op dit plaatje….*	kamt de man een poes	3SG	SON	S	V2
26	*En op dit plaatje….*	kamt de vrouw een hond	3SG	SON	S	V2
27	*De vrouwen….*	filmen een lamp	3PL	SON	C	MC
28	*En de mannen….*	filmen een bloem	3PL	SON	C	MC
29	*Het meisje ziet dat…..*	de vrouw water drinkt	3SG	PLOS	C	SC
30	*En zij ziet dat….*	de man melk drinkt	3SG	PLOS	C	SC
31	*Op deze plaatjes…..*	lezen de vrouwen een krant	3PL	FRIC	S	V2
32	*Op deze plaatjes…..*	lezen de mannen een boek	3PL	FRIC	S	V2

33	*Op deze plaatjes….* kammen de mannen een hond	3PL	SON	S	V2
34	*En op deze plaatjes…*kammen de vrouwen een poes	3PL	SON	S	V2
35	*Het meisje ziet dat…..* de vrouwen een lamp filmen	3PL	SON	C	SC
36	*En zij ziet dat…..* de mannen een bloem filmen	3PL	SON	C	SC
37	*Op deze plaatjes…* bakken de vrouwen een ei	3PL	PLOS	S	V2
38	*En op deze plaatjes...* bakken de mannen een pannekoek	3PL	PLOS	S	V2
39	*De man….* poetst zijn tanden	3SG	FRIC	C	MC
40	*De vrouw….* poetst haar schoen	3SG	FRIC	C	MC
41	*De jongen ziet dat…..* de mannen een ei bakken	3PL	PLOS	S	SC
42	*En hij ziet dat….* de vrouwen een pannekoek bakken	3PL	PLOS	S	SC
43	*De jongen ziet dat…..* de vrouwen melk drinken	3PL	PLOS	C	SC
44	*En hij ziet dat….* de mannen water drinken	3PL	PLOS	C	SC
45	*Op dit plaatje…..* poetst de vrouw haar schoen	3SG	FRIC	C	V2
46	*En op dit plaatje….* poetst de man zijn tanden	3SG	FRIC	C	V2
47	*Op dit plaatje…..* drinkt de vrouw melk	3SG	PLOS	C	V2
48	*En op dit plaatje….* drinkt de man water	3SG	PLOS	C	V2
49	*Het meisje ziet dat…..* de mannen een poes kammen	3PL	SON	S	SC
50	*En zij ziet dat….* de vrouwen een hond kammen	3PL	SON	S	SC
51	*De vrouw….* filmt een bloem	3SG	SON	C	MC
52	*En de man….* filmt een lamp	3SG	SON	C	MC
53	*De jongen ziet dat…..* de man een boek leest	3SG	FRIC	S	SC
54	*En hij ziet dat…..* de vrouw een krant leest	3SG	FRIC	S	SC
55	*De mannen…..* drinken water	3PL	PLOS	C	MC
56	*En de vrouwen….* drinken melk	3PL	PLOS	C	MC
57	*De vrouwen….* kammen een hond	3PL	SON	S	MC
58	*En de mannen….* kammen een poes	3PL	SON	S	MC
59	*De mannen….* lezen een krant	3PL	FRIC	S	MC
60	*En de vrouwen….* lezen een boek	3PL	FRIC	S	MC
61	*Op dit plaatje…..* filmen de vrouwen een bloem	3PL	SON	C	V2
62	*En op dit plaatje…* filmen de mannen een lamp	3PL	SON	C	V2
63	*De jongen ziet dat…..* de vrouw een poes kamt	3SG	SON	S	SC
64	*En hij ziet dat….* de man een hond kamt	3SG	SON	S	SC
65	*De man…..* drinkt water	3SG	PLOS	C	MC
66	*En de vrouw…..* drinkt melk	3SG	PLOS	C	MC
67	*De jongen ziet dat..*de mannen hun schoenen poetsen	3PL	FRIC	C	SC
68	*En hij ziet dat…* de vrouwen hun tanden poetsen	3PL	FRIC	C	SC
69	*De vrouw…..* leest een krant	3SG	FRIC	S	MC
70	*En de man…..* leest een boek	3SG	FRIC	S	MC
71	*De man….* bakt een pannekoek	3SG	PLOS	S	MC
72	*En de vrouw….* bakt een ei	3SG	PLOS	S	MC

G. List of items comprehension relative clauses

SR = subject relative (ambiguous)
SR-pass = subject relative with a passive
SR-PL = subject relative with a difference in number between subject and object
OR-PL = object relative with a difference in number between subject and object

Nr.	Item	Context
1	*dit is de dokter die de mannen bellen*	OR-PL
2	*dit is de clown die de olifant trekt*	SR
3	*dit is de man die de jongens duwen*	OR-PL
4	*dit is de jongen die door de man gekieteld wordt*	SR-pass
5	*dit is de man die de dokters bellen*	OR-PL
6	*dit is de prinses die de ridders vangt*	SR-PL
7	*dit is de man die de vrouwen tekenen*	OR-PL
8	*dit is de man die de jongen knijpt*	SR
9	*dit is de piraat die de clowns slaan*	OR-PL
10	*dit is de prins die de prinsessen ziet*	SR-PL
11	*dit is de vrouw die door de man gekust wordt*	SR-pass
12	*dit is de prinses die de prinsen zien*	OR-PL
13	*dit is de man die de vrouwen tekent*	SR-PL
14	*dit is de ridder die de prinsessen vangen*	OR-PL
15	*dit is de man die door de vrouw gekust wordt*	SR-pass
16	*dit is de olifant die de clown trekt*	SR
17	*dit is de jongen die de mannen duwt*	SR-PL
18	*dit is de jongen die de mannen duwen*	OR-PL
19	*dit is de vrouw die de mannen tekent*	SR-PL
20	*dit is de man die door de jongen geknuffeld wordt*	SR-pass
21	*dit is de clown die de piraten slaat*	SR-PL
22	*dit is de man die de jongens duwt*	SR-PL
23	*dit is de clown die de piraten slaan*	OR-PL
24	*dit is de prinses die de prinsen ziet*	SR-PL
25	*dit is de jongen die de man knijpt*	SR
26	*dit is de prins die de prinsessen zien*	OR-PL
27	*dit is de dokter die de mannen bel*	SR-PL
28	*dit is de man die door de jongen gekieteld wordt*	SR-pass
29	*dit is de man die de jongen trekt*	SR
30	*dit is de prinses die de ridders vangen*	OR-PL
31	*dit is de ridder die de prinsessen vangt*	SR-PL
32	*dit is de jongen die door de man geknuffeld wordt*	SR-pass
33	*dit is de jongen die de man trekt*	SR
34	*dit is de man die de dokters belt*	SR-PL
35	*dit is de vrouw die de mannen tekenen*	OR-PL
36	*dit is de piraat die de clowns slaat*	SR-PL

H. List of items judgement relative clauses

S	=	simple sentence
*S	=	incorrect simple sentence
SR-inan	=	subject relative with an inanimate object
SR-sem	=	subject relative with a semantic cue for thematic role assignment
SR-rev	=	subject relative with a reversible subject and object
OR-inan	=	object relative with an inanimate object (animacy cue)
OR-sem	=	object relative with a semantic cue for thematic role assignment (sem cue)
OR-rev	=	object relative with reversible subject and object (no cue)

nr.	**Item**	**Contex**
1	*de jongen schrijft de brief*	S
2	*dit is het ijsje dat het meisje eet*	OR-inan
3	*de schat steelt de piraat*	*S
4	*dit is de vrouw die de cake maakt*	SR-inan
5	*dit is de auto die de man wast*	OR-inan
6	*de hond laat de man uit*	*S
7	*dit is de prinses die de kroon draagt*	SR-inan
8	*dit is de boterham die de jongen smeert*	OR-inan
9	*de kat pakt de muis*	S
10	*dit is het meisje dat de tekening maakt*	SR-inan
11	*de bal slaat de man*	*S
12	*dit is de vrouw die het paard voert*	SR-sem
13	*dit is de man die de jongen kietelt*	SR-rev
14	*dit is de dief die de politieman oppakt*	OR-sem
15	*dit is de vrouw die de man kust*	SR-rev
16	*de patient prikt de dokter*	*S
17	*dit is de jongen die de man knijpt*	OR-rev
18	*het huis schildert de man*	*S
19	*dit is de vlinder die de beer vangt*	OR-sem
20	*dit is de piraat die de papagaai draagt*	SR-sem
21	*de jongen aait de hond*	S
22	*dit is de man die de jongen duwt*	OR-rev
23	*dit is de jongen die de vader trekt*	SR-rev
24	*de hond voert de jongen*	*S
25	*dit is de poes die de brandweerman redt*	OR-sem
26	*dit is de clown die de hond beloont*	SR-sem
27	*de vrouw breidt de trui*	S
28	*de boer melkt de koe*	S
29	*dit is de man die de jongen knuffelt*	OR-rev
30	*de jongen schopt de bal*	S

I. List of items elicitation relative clauses

SR-SG = subject relative with two singular arguments
SR-PL = subject relative with a difference in number between arguments
OR-SG = object relative with two singular arguments
OR-PL = object relative with a difference in number between arguments
Irrev = irreversible
Rev = reversible

	Target	Revers	Change	Item
1	SR-SG	Irrev	noun	*Er zijn twee jongens. Één jongen drinkt melk en één jongen drinkt water.* TARGET : de jongen die melk/water drinkt
2	OR-PL	rev	verb	*Er zijn twee jongens en twee tantes. De tantes kietelen een jongen en de tantes knijpen een jongen* TARGET: de jongen die de tantes kietelen/knijpen (of passief)
3	SR-SG	rev	noun	*Er zijn twee jongens en een prinses en een ridder. Eén jongen duwt de ridder en één jongen duwt de prinses* TARGET: de jongen die de ridder/de prinses duwt
4	SR-PL	rev	verb	*Er zijn twee jongens en twee moeders. Één jongen bijt de moeders en één jongen kust de moeders* TARGET: de jongen die de moeders bijt/kust
5	OR-SG	rev	verb	Er zijn twee jongens en een juf. De juf schopt een jongen en de juf tekent een jongen. TARGET: de jongen die de juf schopt/tekent (of passief)
6	SR-SG	irrev	noun	*Er zijn twee jongens en een poes en een hond. Één jongen kamt een hond en één jongen kamt een poes* TARGET: de jongen die een poes kamt/een hond kamt
7	OR-PL	rev	verb	*Er zijn twee jongens en twee ooms. De ooms schoppen een jongen en de ooms knuffelen een jongen* TARGET: de jongen die de ooms schoppen/knuffelen (of passief)
8	SR-PL	rev	noun	*Er zijn twee jongens en twee opa's en twee oma's. Éen jongen vangt de oma's en één jongen vangt de opa's* TARGET: de jongen die de opa's/oma’s vangt
9	OR-SG	rev	noun	*Er zijn twee jongens en een buurman en een moeder. De buurman roept een jongen en de moeder roept een jongen* TARGET: de jongen die de moeder/de buurman roept (of passief)
10	SR-SG	irrev	noun	*Er zijn twee jongens. Één jongen poetst zijn schoen en één jongen poetst zijn tanden.* TARGET: de jongen die zijn schoen/zijn tanden poetst
11	OR-PL	rev	noun	*Er zijn twee jongens en twee voetballers en twee ballerina’s. De voetballers roepen een jongen en de ballerina’s roepen een jongen.* TARGET: de jongen die de ballerina's/voetballers roepen (of passief)
12	SR-SG	rev	verb	*Er zijn twee jongens en een tante. Één jongen bijt de tante en één jongen kust de tante.* TARGET: de jongen die de tante bijt/kust
13	OR-SG	rev	noun	*Er zijn twee jongens en een opa en een oma. De opa kust een jongen en de oma kust een jongen.* TARGET: de jongen die de opa/oma kust(of passief)
14	SR-SG	irrev	noun	*Er zijn twee jongens. Één jongen leest een boek en één jongen leest een krant*

				TARGET: de jongen die een boek leest/krant leest
15	OR-PL	rev	verb	*Er zijn twee jongens en twee vriendjes. De vriendjes slaan een jongen en de vriendjes tekenen een jongen* TARGET: de jongen die de vriendjes slaan/tekenen (of passief)
16	SR-PL	rev	verb	*Er zijn twee jongens en twee vaders. Één jongen slaat de vaders en één jongen aait de vaders* TARGET: de jongen die de vaders slaat/aait
17	OR-SG	rev	verb	*Er zijn twee jongens en een moeder. De moeder kietelt een jongen en de moeder knijpt een jongen* TARGET: de jongen die de moeder kietelt/knijpt (of passief)
18	SR-SG	irrev	noun	*Er zijn twee jongens. Één jongen bakt een ei en één jongen bakt een pannenkoek* TARGET: de jongen die een ei/pannekoek bakt
19	SR-PL	rev	noun	*Er zijn twee jongens en twee dokters en twee clowns. Één jongen roept de dokters en één jongen roept de clowns.* TARGET: de jongen die de dokters/clowns roept
20	SR-SG	rev	noun	*Er zijn twee jongens en een voetballer en een ballerina. Één jongen vangt de voetballer en één jongen vangt de ballerina* TARGET: de jongen die de voetballer/ballerina vangt
21	OR-SG	rev	verb	*Er zijn twee jongens en een vader. De vader slaat een jongen en de vader knuffelt een jongen.* TARGET: de jongen die de vader knuffelt/slaat (of passief)
22	SR-SG	rev	verb	*Er zijn twee jongens en een oom.Één jongen belt de oom en één jongen knijpt de oom* TARGET: de jongen die de oom belt/knijpt
23	OR-PL	rev	noun	*Er zijn twee jongens en twee clowns en twee prinsessen. De clowns bellen een jongen en de prinsessen bellen een jongen* TARGET: de jongen die de clowns/prinsessen bellen (of passief)
24	SR-PL	rev	noun	*Er zijn twee jongens en twee buurmannen en twee vaders. Eén jongen duwt de buurmannen en één jongen duwt de vaders* TARGET: de jongen die de buurmannen/vaders duwt
25	OR-SG	rev	noun	*Er zijn twee jongens en een clown en een piraat. De clown vangt een jongen en de piraat vangt een jongen.* TARGET: de jongen die de piraat/clown vangt (of passief)
26	SR-SG	irrev	noun	*Er zijn twee jongens.Één jongen filmt een bloem en één jongen filmt een lamp* TARGET: de jongen die de bloem/de lamp filmt
27	SR-SG	rev	noun	*Er zijn twee jongens en een dokter en een vader. Één jongen roept de dokter en één jongen roept de vader* TARGET: de jongen die de dokter/vader roept
28	OR-PL	rev	noun	*Er zijn twee jongens en twee piraten en twee ridders. De ridders vangen een jongen en de piraten vangen een jongen* TARGET: de jongen die de ridders/piraten vangen (of passief)
29	SR-PL	rev	verb	*Er zijn twee jongens en twee juffen. Één jongen belt de juffen, één jongen knijpt de juffen* TARGET: de jongen die de juffen belt/knijpt
30	SR-SG	rev	verb	*Er zijn twee jongens en een meester. Één jongen aait de meester en één jongen slaat de meester.* TARGET: de jongen die de meester aait/slaat

J. List of codings for the elicitation of relative clauses

SR			
Analysable target SR-ambiguous			
nr	*Category*		*example*
1	SR/OR ambiguous with object	+head	Ik ben liever de jongen die de vader roept
2		-head	Ik ben liever **ø** die de vader roept
3	SR (PL for SG)	+head	Ik ben liever de jongen die de vaders roept
4		-head	Ik ben liever **ø** die de vaders roept
5	SR - no direct object (DO)	+head	Ik ben liever de jongen die kust
6		-head	Ik ben liever **ø** die kust (if verb alternates)
7	SR – verb second (V2)	+head	Ik ben liever de jongen die roept de vader
8		-head	Ik ben liever **ø** die roept de vader
9	Role reversal passive (OR for SR)	+head	Ik ben liever de jongen die geroepen wordt (door de moeder)
10		-head	Ik ben liever **ø** die geroepen wordt
17	'wie' instead of 'die'	+head	Ik ben de jongen wie de vader roept
19		-head	Ik ben **ø** wie de vader roept
18	'wie' instead of 'die'– V2	+head	Ik ben de jongen wie roept de vader
20	'welke' instead of 'die'	+head	Ik ben de jongen welke de vader roept
21		-head	Ik ben **ø** welke de vader roept
22	Head reversal		Ik ben liever de ridder die (de jongen) vangt
24	'wie' instead of 'die' - no DO	+head	die ik ben liever de kind wie belt naar de oom
25		-head	die ik ben liever **ø** wie belt naar de oom
26	'welke' instead of 'die' – V2	+head	Ik ben liever de jongen welke roept de vader
27		-head	Ik ben liever **ø** welke roept de vader
28	SR (PL for SG) – V2	+head	Ik ben liever de jongen die vangt de ballerina's
29		-head	Ik ben liever **ø** die vangt de ballerina's
Analysable non-target SR-ambiguous			
15	Role reversal (PL for SG)	+head	Ik ben de jongen die de moeders roepen
16		-head	Ik ben **ø** die de moeders roepen
Non-analysable SR-ambiguous			
11	NA – no relative clause		Ik ben lief dus ik ga de mama kussen Ik ben liever de vader vangen ofzo Ik duw liever de vader Ik ben liever de vader-jongen Ik ben liever voor de vader Ik ben liever de eerste Ik ben liever die jongen De kietelende jongen (reduced relative?)
12	NA – no relative marker		Ik ben liever de jongen de buurman roept
13	NA – 'waarvan'		Ik ben liever het meisje waarvan degene belt
14	NA – 'Ik heb liever'		Ik ben liever dat de jongen de juf belt
23	NA-other		
99	Excluded		Teveel weggegeven of niet goed doorgevraagd

SR-PL			
Analysable target SR-PL			
nr	*Category*		*example*
1	SR_PL	+head	Ik ben liever de jongen die de voetballers roept
2		-head	Ik ben liever ø die de voetballers roept
3	SR/OR ambiguous (SG for PL)	+head	Ik ben liever de jongen die de voetballer roept (SG for PL)
4		-head	Ik ben liever ø die de voetballer roept (SG for PL)
5	SR - no direct object (DO)	+head	Ik ben liever de jongen die kust (if verb alternates)) (evt. + voorzetsel en object – met de vader)
6		-head	Ik ben liever ø die kust (if verb alternates)
7	SR – verb second (V2)	+head	Ik ben liever de jongen die roept de voetballers
8		-head	Ik ben liever ø die roept de voetballers
9	Role reversal passive (OR for SR)	+head	Ik ben liever de jongen die geroepen wordt (door de voetballers/voetballer)
10		-head	Ik ben liever ø die geroepen wordt
17	SR (SG for PL) - V2	+head	Ik ben liever de jongen die roept de voetballer
18		-head	Ik ben liever ø die roept de voetballer
19	SR/role reversal – error SVA		Ik ben liever de jongen die de buurman duwen
20	'wie/welke/wat' instead of 'die'	+head	ik ben liever de jongen wie de voetballers duwt
21		-head	ik ben ø wie de voetballers duwt
22	'wie/welke/wat' (SG for PL)	+head	ik ben liever de jongen wie de voetballer duwt
23		-head	ik ben ø wie de voetballer duwt
24	Head reversal		Ik ben liever de opa's die (de jongen) vangen
25	Nr. reversal		Ik ben liever de jongens die de voetballers roept
27	'wie/welke/wat' – V2	+head	ik ben liever de jongen wie/welke/wat kust de vaders
		-head	ik ben liever ø wie/welke/wat kust de vaders
29	'wie/welke/wat' (SG for PL) – V2	+head	ik ben liever de jongen wie/welke/wat kust de vader
30		-head	ik ben liever ø wie/welke/wat kust de vader
Analysable non-target SR-PL			
15	Role reversal	+head	Ik ben liever de jongen die de voetballers roepen
16		-head	Ik ben liever ø die de voetballers roepen
Non-analysable SR-PL			
11	NA – no relative clause		Ik ben lief dus ik ga de voetballers kussen Ik ben liever de voetballers vangen Ik duw liever de voetballers Ik ben liever de voetbal-jongen Ik ben liever voor de voetballers Ik ben liever de eerste (met de voetballers) Ik ben liever die jongen De kietelende jongen (reduced relative?)
12	NA – no relative marker		Ik ben liever de jongen de voetballers roept
13	NA – 'Waarvan'/'waarbij'		Ik ben liever de jongen waarvan degene belt
14	NA – 'Ik heb liever'		Ik ben liever dat de jongen de voetballers roepen
26	NA-other		
99	Excluded		Teveel weggegeven of niet goed doorgevraagd

OR-ambiguous			
Analysable target OR-ambiguous			
nr	*Category*		*example*
1	OR/SR Ambiguous	+head	Ik ben liever de jongen die de vader roept
2		-head	Ik ben liever **ø** die de vader roept
3	OR – by case marking	+head	Ik ben liever de jongen die ze belt (de juf belt een jongen en de juf schopt een jongen)
4		-head	Ik ben liever **ø** die ze belt (de juf belt een jongen en de juf schopt een jongen)
9	OR_PL	+head	Ik ben liever de jongen die de vaders roepen (PL for SG)
10		-head	Ik ben liever **ø** die de vaders roepen (PL for SG)
28	'Welke/wat/wie' instead of 'die'	+head	ik ben liever de jongen welke de vader roept
Analysable target OR-ambiguous – passive			
5	Passive with Agent	+head	Ik ben liever de jongen die geroepen wordt door de vader
6		-head	Ik ben liever **ø** die geroepen wordt door de vader
7	Passive without Agent	+head	Ik ben liever de jongen die geroepen wordt (if verb alternates)
8		-head	Ik ben liever **ø** die geroepen wordt (if verb alternates)
29	Passive incorrect		Ik ben liever de jongen die roept door zijn moeder, ik ben liever gekust door mijn oma
Analysable non-target OR-ambiguous			
11	SR for OR – Head reversal	+head	Ik ben liever de vader die de jongen roept
12		-head	Ik ben liever **ø** die de jongen roept
13	SR for OR – role reversal	+head	Ik ben liever de jongen die roept (if verb alternates) (naar, met, tegen, aan de vader)
14		-head	Ik ben liever **ø** die roept (if verb alternates)
15	SR for OR – error SVA	+head	Ik ben liever de jongen die roept de vader
16		-head	Ik ben liever **ø** die roept de vader
25	SR for OR – head reversal – V2	+head	Ik ben liever vader die roept een jongen
26		-head	Ik ben liever **ø** die roept een jongen
Non-analysable OR-ambiguous			
17	Reduced relative		Ik ben liever de gekietelde jongen
18	NA – no relative clause		Ik ben liever kietelen Ik wil gekieteld worden Ik ben liever de jongen met de vader Ik ben liever de vader Ik ben liever knuffelt een vader Ik ben liever voor de vader Ik ben liever de eerste Ik ben liever die jongen Ik ben liever opa
19	NA – no relative marker		Ik ben liever de jongen de vader roept
20	NA – 'waarvan/waarbij' (+ resumptive		Ik ben liever de jongen waar(van) de vader/degene (hem/de jongen) belt

	pronoun/NP)		
21	NA – 'dat/die' (+ resumptive pronoun/NP)		Ik ben liever de jongen dat de vader (hem/de jongen) roept, ik ben de jongen die juf teken mij
22	NA-'waardoor'		Ik wil het liefst de jongen zijn waardoor hij geroepen wordt door de vader
23	NA- 'Ik heb liever'		Ik ben liever (de jongen) dat de vader hem/de jongen roept
24	NA-other		Ik ben liever de jongen waarvan degene roept Ik ben liever degene die mij roept Ik ben liever de jongen die mij roept Ik ben liever het meisje
27	Ambiguous OR/SR – error SVA		Ik ben liever de jongen die de vader roepen
99	Excluded		Teveel weggegeven of niet goed doorgevraagd

OR-PL			
Analysable target OR-PL			
nr	*Category*		*Example*
1	OR_PL	+head	Ik ben liever de jongen die de voetballers roepen
2		-head	Ik ben liever **ø** die de voetballers roepen
3	OR – V2	+head	Ik ben liever de jongen die roepen de voetballers
4		-head	Ik ben liever **ø** die roepen de voetballers
5	OR/SR ambiguous with object (SG for PL or PL for SG)	+head	Ik ben liever de jongen die de voetballer roept (SG for PL)
6		-head	Ik ben liever **ø** die de voetballer roept (SG for PL)
33	OR – nr. reversal		Ik ben liever de jongens die de oom knuffelt
37	'waarmee'		Ik ben de jongen waarmee de ooms knuffelen
39	'wie/wat/welke' instead of 'die'	+head	Ik ben de jongen wie de clowns bellen
40		-head	Ik ben **ø** wie de clowns bellen
45	OR_PL – 'wie/wat/welke '		Ik ben liever de jongen wie/welke/wat de voetballers roepen
46	OR – 'wie/wat/welke' - V2		Ik ben liever de jongen wie/welke/wat roepen de voetballers
Analysable target OR-PL– passive			
7	Passive	+head	Ik ben liever de jongen die geroepen wordt door de vaders
8		-head	Ik ben liever **ø** die geroepen wordt door de vaders
9	Passive (SG for PL)	+head	Ik ben liever de jongen die geroepen wordt door de vader (SG for PL)
10		-head	Ik ben liever **ø** die geroepen wordt door de vader (SG for PL)
11	Passive without Agent	+head	Ik ben liever de jongen die geroepen wordt
12		-head	Ik ben liever **ø** die geroepen wordt
34	Passive incorrect	+head	Ik ben liever de jongen die door de voetballers roept
35		-head	Ik ben liever **ø** die door de voetballers roept
44	Passive – 'wie/wat/welke'		ik ben het meisje wat/welke/wie door de prinsessen wordt gebeld

Analysable non-target OR-PL			
13	SR for OR – Head reversal	+head	Ik ben liever de juffen die de jongen tekenen
14		-head	Ik ben liever ø die de jongen tekenen
15	SR for OR – Head reversal (SG for PL)	+head	Ik ben liever de juf die de jongen tekent (SG for PL)
16		-head	Ik ben liever ø die de jongen tekent (SG for PL)
17	SR for OR – Role reversal	+head	Ik ben liever de jongen die de voetballers roept (waar jongen eigenlijk object is)
18		-head	Ik ben liever ø die de voetballers roept (waar jongen eigenlijk object is)
19	SR for OR –Role reversal– V2	+head	Ik ben liever de jongen die roept de voetballers
20		-head	Ik ben liever ø die roept de voetballers
21	SR for OR – intransitive	+head	Ik ben liever de jongen die gaat voetballen
22		-head	Ik ben liever ø die gaat voetballen
36	SR for OR - role reversal - no DO		Ik ben liever de jongen die belt (naar de prinses)
41	Head reversal (SG for PL) - V2		Ik ben liever die vriendin die tekent een meisje
43	Head reversal – V2		Ik ben liever die vriendinnen die tekenen een meisje
Non-analysable OR-PL			
23	Reduced relative		Ik ben liever de gekietelde jongen
24	NA – no relative clause		Ik ben liever kietelen Ik wil gekieteld worden Ik ben liever de jongen met de voetballers Ik ben liever de voetballers De buurman roept de jongen Ik ben liever voor de voetballers Ik ben liever de eerste Ik ben liever die jongen
25	NA – no relative marker		Ik ben liever de jongen de voetballers roepen
26	NA – 'waarvan/waarbij' (+ resumptive pronoun/NP)		De jongen waarvan de voetballers hem roepen Ik wil zijn de jongen waarvan de voetballers naar hem roepen
27	NA – 'dat/die' (+ resumptive pronoun)		Ik ben liever de jongen dat de voetballers hem roepen
28	NA-'waardoor'		Ik wil het liefst de jongen zijn die, waardoor hij gevangen wordt door de ridders
30	NA- 'Ik heb liever'		Ik ben liever dat de jongen de voetballers roept Is liever dat de jongens een clown bellen Ik ben liever dat de ridders een jongen vangen
31	NA-other		Ik ben liever de jongen waarvan degene roepen Ik ben liever degene die mij roepen Ik ben liever de jongens die vangt de oma Ik ben liever dat de juf zich tekent Ik ben liever de jongen met de tante die knijpt Ik ben liever het meisje die de tante haar kietelt Ik ben liever de meisje die de prinsessen een meisje bellen
32	Ambiguous OR/SR – error SVA		Ik ben liever de jongen die de prinses bellen
99	Excluded		Teveel weggegeven of niet goed doorgevraagd

Summary in English

Persistent grammatical difficulties in SLI. Deficits in knowledge or in knowledge implementation?

If children have significant difficulties learning or applying the linguistic rules of their language while there is no clear aetiology for these difficulties (e.g., a hearing impairment, limited cognitive abilities or a neurological deficit), they are often diagnosed with the disorder Specific Language Impairment (SLI). Children with SLI are characterized as having frequent errors in their language production; difficulties in the grammatical domain are often regarded as a core symptom of the disorder. Which grammatical elements are impaired is, however, partly dependent on the typology of the language they are acquiring. In Dutch, grammatical errors in SLI are, for instance, apparent in the form of the definite determiner (grammatical gender marking), in verb inflections (e.g., subject verb agreement) and in the construction of complex clauses (e.g., relative clauses). These errors are also found in early acquisitional stages of typical development, but they persevere longer in children with SLI. It is however unclear until what age these grammatical errors persist. A first aim of this study was therefore to examine whether grammatical difficulties in children with SLI are persistent into adolescence.

Grammatical errors in the language production of children with SLI have often been interpreted as a deficit in grammatical knowledge, but several studies have noted that children with SLI show variability in grammatical performance. Grammatical aspects are sometimes produced correctly but in other contexts contain an error. More so than in typical development, grammatical performance in SLI seems to be dependent on the complexity of the linguistic context or task in which grammatical aspects are produced. This suggests that children with SLI not only have problems with the acquisition of grammatical knowledge, but also have problems with the implementation of knowledge once it has been acquired.

In 1994, Bishop proposed the hypothesis that variability in grammatical performance in SLI originates in a ‘limited capacity system that is handling several operations at the same time’. This hypothesis, known as the Vulnerable Markers Hypothesis (VMH), suggested that the same processing problems that are often assumed to underlie deficits in grammatical knowledge in SLI may also have an effect on grammatical performance once grammatical knowledge is in place. As a

result, grammatical markers may not be produced in linguistic contexts that involve a higher processing load. A second aim of this thesis was to test the claims of the VMH by examining firstly under what conditions individuals with SLI are impeded in the implementation of grammatical knowledge, and secondly whether variability in performance is related to (limited) processing abilities.

The two aims of this thesis were linked. We assumed that the investigation of the predictions of the VMH could be tested best in an older population of adolescent subjects in which grammatical knowledge has relatively stabilized. Although adolescents might still acquire grammatical knowledge, their access to implicit learning mechanisms is assumed to be limited. Differences between knowledge (grammatical knowledge) and performance (knowledge implementation) might therefore show up more clearly in adolescence than in childhood, when grammatical acquisition is still very much in development. To this end, two groups of individuals with SLI were examined on their grammatical performance: a younger group of children and an older group of adolescents. The two age groups were matched on severity of the language problems to validate examination of development in a cross-sectional design (more details concerning the rationale of this study can be found in *Chapter 1*).

In order to formulate hypotheses in relation to the research questions of this thesis, *Chapter 2* provided background information on three relevant topics. First of all, the literature on Dutch SLI was reviewed and the choice for the three grammatical variables tested in this thesis was motivated. Secondly, the literature on the persistence of grammatical difficulties in SLI was summarized to define the empirical background in which the results of this study had to be placed. Lastly, the concept 'processing complexity' was defined and the literature on linguistic factors affecting grammatical performance was reviewed. On the basis of these reviews, the following hypotheses were formed. First of all, we expected to find differences between individuals with and without SLI in all three grammatical variables (grammatical gender, subject verb agreement and relative clauses) and we expected these differences to be persistent into adolescence. Secondly, we expected to find variability in performance on the basis of the linguistic context factors that we varied in the grammatical tasks. This variability was expected to be larger in SLI than in TD. Thirdly, we expected to find that variability in performance would be related to (limited) processing abilities.

In *Chapter 3*, the methodology used to recruit and select subjects was presented in detail, as well as the type of tasks, the general procedures of administration, transcription, scoring and coding and the statistical analyses. As was mentioned above, the persistence of grammatical difficulties over time was tested cross-sectionally by comparing a younger and an older group of individuals with SLI. Furthermore, the two SLI-groups were matched to two TD groups on age and gender to examine differences between SLI and TD in grammatical performance. An adult group was added to check the validity of some of our tasks. Chapter 3 also provided information on the general language abilities, the non-verbal cognitive abilities, the socio-economic status (SES) and the developmental history of the individuals in our child and adolescent participant groups. The SLI and TD groups appeared to differ in general language ability, as expected, but also differed in terms of non-verbal IQ and SES. Differences were largest in the younger groups, but were still present in the older groups despite recruitment from similar types of school and environment. Non-verbal IQ was therefore corrected for in all the statistical analyses that were performed in this thesis (because non-verbal IQ and SES were highly related, we did not correct for SES).

Before we could discuss the grammatical abilities of our participants to see whether our three hypotheses were borne out, one assumption had to be first tested. The claim that variability in grammatical performance is related to limitations in processing abilities in SLI entails the assumption that processing capacities are limited in SLI. This assumption was checked in *Chapter 4*, which described the information processing abilities of the children and adolescents with and without SLI. Five different processing measures were selected based on previously attested robust differences between SLI and TD. These measures involved executive control functions (motor inhibition) and working memory in different modalities (visual and verbal) and with differences in the amount of verbal load. The results on the different processing measures showed that the assumption that processing abilities are limited in SLI needs some refinement. Whether processing limitations were found seemed to be linked to the amount of verbal load in a processing task. Furthermore, differences between SLI and TD in processing abilities seemed to decrease with age and persistent differences were only found in the tasks with a high verbal load (nonword repetition and sentence repetition). These results implied that we could only expect to find (negative) correlations with variability in grammatical performance for these two verbal processing measures.

In *Chapter 5*, the results of the tests of grammatical gender were reported. They had many different facets. The variable was investigated in common and neuter nouns, in judgement and in production, and in the assignment of determiners and the inflection of adjectives. The data of determiner assignment and adjectival inflection were also combined to investigate whether a gender feature was present and whether the rule for adjectival inflection had been acquired. Finally, the effect of task and linguistic context on performance in SLI was examined, and linked to the verbal processing measures. Grammatical gender proved to be persistently difficult for individuals with SLI and knowledge of this grammatical feature seemed to fossilize before adolescence. Although performance on determiner assignment and adjectival inflection was in general very low in both SLI groups, influences of the linguistic context in which performance was tested could still be measured. Performance was affected more negatively in the SLI groups, and this effect was related to (limited) verbal processing abilities. The problems with grammatical gender in Dutch subjects with SLI clearly indicated a problem at the representational level (knowledge fossilizes) but their errors in the production of grammatical gender also reflected problems in the implementation of knowledge.

Chapter 6 presented the results of the tests for subject verb agreement. On the different subject verb agreement tasks the rules for subject verb agreement had generally been mastered by adolescents with SLI, but persistent differences were still found. Adolescents with SLI improved in terms of accuracy and the amount of finite lexical verbs produced (instead of auxiliary constructions, which were often used by younger subjects with SLI), but they still made some inflection errors in the singular domain. Errors occurred more often in phonologically and syntactically complex contexts and these influences of linguistic context appeared to be related to (limited) processing abilities. This confirmed the idea that grammatical aspects that have been acquired in SLI are still vulnerable in performance when the processing load increases.

In *Chapter 7*, the results on the comprehension, judgement and production of relative clauses were described. The children and adolescents with SLI had persistent problems with the production of relative clauses in general, although their abilities improved over time. Object relatives were harder than subject relatives, and this difference was larger in SLI. The enhanced effect of linguistic context on the production of relatives in SLI was linked to processing abilities. In contrast to what is reported in the cross-linguistic literature on the comprehension of relatives in SLI, children and adolescents with SLI were equally able to interpret object relatives as TD peers if object relatives had typical characteristics (an inanimate object and an animate subject). They did, however, show stronger negative effects when object relatives were non-typical or when the cue for interpretation was in sentence-final position.

Chapter 8 evaluated whether the three hypotheses that had been tested in this thesis were confirmed. The first hypothesis that persistent grammatical difficulties in SLI would be found in all grammatical variables studied in this thesis was confirmed for grammatical gender marking and the production of (object) relative clauses. These grammatical aspects revealed not only persistent differences between SLI and

TD, but also seemed to indicate genuine difficulties in the grammatical domain. For subject verb agreement, differences between SLI and TD were also persistently found but these differences could not really be classified as difficulties. The scores were generally quite high in both SLI groups and the rules for subject verb agreement seemed to have been acquired before adolescence. Grammatical aspects thus remain vulnerable in SLI, but the specific grammatical profile changes over time.

The second hypothesis that grammatical performance would vary on the basis of linguistic context factors and that this variability would be larger in SLI than in TD, was confirmed for many of the context factors that were implemented in the grammatical tasks. Performance appeared to be generally lower in contexts that were assumed to be more complex, and effects of these context factors were often larger in SLI. We can now list a range of factors that more strongly influence grammatical performance in Dutch SLI.

The third hypothesis that the larger influence of these linguistic context factors in SLI were related to (limited) processing abilities was examined by performing correlational analyses between the context effects and the verbal processing measures that showed persistent difficulties in SLI (sentence repetition and nonword repetition, as was discussed in Chapter 4). Almost all context effects were related to the verbal processing abilities that showed limitations in SLI. Bishop's idea that "errors would occur when the speech production system is stressed by the need to produce output that makes heavy demands on its processing capacity" (Bishop, 1994, p. 528) was therefore supported. We do not claim that SLI is *explained* by a deficit in processing capacities, nor do we claim that the relationship with processing that we found in this study has a clear causality. The results of this study do, however, indicate that errors in language production in SLI do not necessarily reflect deficits in grammatical knowledge but might sometimes stem from difficulties to implement grammatical knowledge in production when the processing load is increased. Future research should further examine under what conditions knowledge implementation is impeded or facilitated in SLI.

Samenvatting

Persistente grammaticale problemen bij kinderen met specifieke taalontwikkelingsstoornissen (SLI). Een gebrek aan grammaticale kennis, of een probleem in de toepassing ervan?

Wanneer een kind ernstige problemen heeft in de verwerving van de moedertaal zonder dat daar een duidelijke reden voor is (zoals een probleem in het gehoor, een beperkt IQ of een neurologische afwijking) wordt vaak de diagnose Specifieke Taalontwikkelingsstoornis (S-TOS) toegekend. In het Engels wordt deze stoornis aangeduid met de term Specific Language Impairment (SLI). Kinderen met SLI laten veel fouten zien in hun taalproductie, met name in de grammatica. Grammaticale problemen worden daarom vaak gezien als een belangrijk symptoom van de stoornis. Welke grammaticale fouten kinderen met SLI maken verschilt echter van taal tot taal. In het Nederlands zijn er bijvoorbeeld vaak problemen met het lidwoord (de markering van grammaticaal geslacht), werkwoordvervoegingen (bijvoorbeeld in de afstemming tussen onderwerp en werkwoord) en in de constructie van complexe zinnen (relatieve bijzinnen bijvoorbeeld). De volgende zin illustreert de grammaticale problemen van Nederlandse kinderen met SLI:

We hebt de konijn in de hok gestopt om hij moet slapen

De fouten die kinderen met SLI maken kunnen ook in vroege stadia van de normale taalontwikkeling worden teruggevonden, maar fouten houden dikwijls langer aan bij kinderen met SLI. Het is echter onduidelijk tot welke leeftijd grammaticale fouten aanhouden. Veel onderzoek naar SLI heeft zich gericht op de basisschoolleeftijd en weinig studies kijken naar grammaticale problemen in de adolescentie. Het *eerste doel van* deze studie was daarom om te onderzoeken of de grammaticale fouten die Nederlandse kinderen met SLI maken ook in de adolescentie nog (vaker dan gewoonlijk) gevonden worden.

Grammaticale fouten in de taalproductie van kinderen met SLI worden vaak geïnterpreteerd als een indicatie dat zij grammaticale kennis missen. Veel studies hebben echter ook opgemerkt dat kinderen met SLI variatie laten zien in hun grammaticale presteren. Dezelfde grammaticale aspecten worden soms correct geproduceerd en worden in andere contexten in foutieve vorm aangetroffen. Meer dan bij een normale taalontwikkeling lijkt het grammaticale presteren van kinderen met SLI afhankelijk te zijn van de complexiteit van de talige context of de taak waarin een grammaticaal element wordt geproduceerd. Die variatie kan niet worden verklaard door ontbrekende kennis, en doet vermoeden dat kinderen met SLI niet alleen problemen hebben met de verwerving van grammaticale kennis maar ook met de toepassing ervan.

In 1994 publiceerde Dorothy Bishop een artikel waarin zij de variatie in het grammaticale presteren van kinderen met SLI aan de kaak stelde. Zij opperde het idee dat deze variatie misschien kan worden verklaard vanuit beperkingen in de informatieverwerkingscapaciteiten van kinderen met SLI, die vooral naar voren komen wanneer er op het talige vlak veel van hen gevraagd wordt. Deze hypothese, beter bekend als de Vulnerable Markers Hypothesis (VMH), sloot aan bij het reeds bestaande idee dat SLI voortkomt uit een algemener probleem in het verwerken van (talige) informatie – een theorie die de laatste decennia in zwang is geraakt. In tegenstelling tot eerdere informatieverwerkingstheorieën veronderstelde Bishop dat problemen in de informatieverwerking niet alleen impact hebben op het verwerven van grammaticale kennis, maar ook op de toepassing ervan in het taalproductieproces. Zij voorspelde dat grammaticale elementen vooral kwetsbaar zijn in contexten die een hoger beroep doen op de informatieverwerking. Het toetsen van de voorspellingen van de VMH was het *tweede doel* van deze studie. Dat deden we door te testen in welke condities kinderen met SLI de meeste grammaticale fouten maken en door te onderzoeken of de mate van variatie in presteren was gerelateerd aan de informatieverwerkingscapaciteit van de groepen.

De twee doestellingen van deze studie waren sterk met elkaar verbonden. In onze optiek konden de voorspellingen van de VMH het best worden getoetst in een wat oudere populatie van adolescenten met SLI omdat hun grammaticale kennis relatief stabiel is. Hoewel adolescenten nog steeds grammaticale kennis kunnen verwerven wordt vaak verondersteld dat de toegang tot impliciete leermechanismen is afgenomen. Verschillen tussen grammaticale kennis en de toepassing ervan zouden daarom duidelijker zichtbaar kunnen zijn in de adolescentie dan in de kindertijd, waarin grammaticale kennis nog volop in ontwikkeling is. Om die reden werden er twee groepen met SLI onderzocht: een groep kinderen in de basisschoolleeftijd en een groep adolescenten in de middelbare schoolleeftijd. We zorgden ervoor dat de twee SLI-groepen vergelijkbaar waren in de ernst van de taalstoornis, zodat we op een 'cross-sectionele' manier iets konden zeggen over de grammaticale ontwikkeling (in *Hoofdstuk 1* wordt de onderzoeksrationale in meer detail besproken).

Om hypothesen te vormen met betrekking tot de onderzoeksvragen van deze studie was het belangrijk om op drie terreinen de achtergrondliteratuur te bespreken.

Deze achtergrondinformatie kan worden gevonden in *Hoofdstuk 2*. Het hoofdstuk geeft een overzicht van de literatuur over Nederlandse kinderen met SLI, om de keuze voor de grammaticale variabelen die in dit onderzoek worden getoetst toe te lichten. Er worden in dit onderzoek drie variabelen getoetst: grammaticaal geslacht, congruentie tussen onderwerp en werkwoord en relatieve bijzinnen. Vervolgens wordt de (cross-linguïstische) literatuur over de persistentie van grammaticale problemen in SLI samengevat om de uitkomsten van deze studie goed te kunnen plaatsen in het onderzoeksveld. Tenslotte wordt geschetst wanneer we spreken van 'een complexere talige context' en wordt besproken van welke linguïstische contextfactoren eerder een effect op het grammaticale presteren is aangetoond. Op basis van deze achtergrondinformatie werden drie hypothesen gevormd. Allereerst verwachtten we een verschil tussen proefpersonen met en zonder SLI in de drie grammaticale aspecten die we toetsten en we verwachtten dat deze verschillen in de adolescentie nog steeds meetbaar zouden zijn. Ten tweede verwachtten dat de grammaticale prestaties van individuen zouden variëren, afhankelijk van de linguïstische context waarin grammaticale elementen werden getoetst. Hierbij verwachtten we dat de variatie groter zou zijn in de SLI-groepen dan in de groepen met een normale taalontwikkeling. Tenslotte verwachtten we een relatie tussen de mate van variatie in grammaticaal presteren en (beperkingen in) de informatieverwerking.

In *Hoofdstuk 3* wordt in detail besproken hoe de proefpersonen werden geworven en geselecteerd. Het hoofdstuk bespreekt ook welk type taken werd gebruikt, hoe de taken werden afgenomen, hoe de data werden getranscribeerd, gescoord en gecodeerd, en hoe de data statistisch werden geanalyseerd. Zoals hierboven genoemd werd de persistentie van grammaticale problemen getest op een 'cross-sectionele' manier: door een jongere en een oudere groep proefpersonen met SLI te vergelijken die overeen kwamen in de ernst van hun taalstoornis. De twee SLI-groepen werden beiden ook gekoppeld aan een groep proefpersonen met een normale taalontwikkeling (typical development, hierna TD) in de zelfde leeftijd (met hetzelfde aantal jongens en meisjes). Op die manier werd getoetst in hoeverre de SLI-groepen afweken van de norm in grammaticaal presteren. Een groep volwassenen werd toegevoegd om de validiteit van sommige van onze testen te toetsen.

Hoofdstuk 3 bevat ook gedetailleerde informatie over de algemene taalvaardigheid, het non-verbale IQ, de socio-economische status (SES) en de ontwikkelingshistorie van de verschillende groepen. Behalve in taalvaardigheid verschilden de SLI- en TD-groepen ook in SES en in hun non-verbale cognitieve vaardigheden. De verschillen waren het grootst in de jongere groepen, maar waren ook aanwezig in de oudere groepen, ondanks de poging deze variabelen constant te houden door SLI- en TD-leerlingen op dezelfde scholen te werven. In alle statistische analyses van grammaticale verschillen en verschillen in de informatieverwerking werd daarom gecorrigeerd voor verschillen in het non-verbale IQ (door deze maat mee te nemen als covariaat). Omdat non-verbaal IQ en SES sterk gecorreleerd waren werd niet ook nog voor SES gecorrigeerd.

Voor het toetsen van de hypothese dat variatie in grammaticaal presteren samenhangt met beperkingen in de informatieverwerkingscapaciteiten in SLI was het belangrijk om te onderzoeken of de informatieverwerkingscapaciteiten in de SLI-groepen ook daadwerkelijk beperkt zijn. Deze assumptie wordt getoetst in *Hoofdstuk 4,* waarin de informatieverwerkingscapaciteiten van de SLI- en TD-groepen werden getoetst en vergeleken. Op basis van eerder gevonden robuuste verschillen tussen SLI en TD werden vijf informatieverwerkingsmaten geselecteerd. Vier maten testten het werkgeheugen in verschillende modaliteiten (visueel en verbaal) en met verschillen in het beroep op talige vermogens (cijfers, nonsenswoorden of zinnen herhalen). Eén maat richtte zich op het kunnen onderdrukken van automatische (motorische) responsen (inhibitie, een onderdeel van de executieve functies). De uitkomsten van deze verschillende testen lieten zien dat de aanname dat ‘informatieverwerkingscapaciteiten beperkt zijn in SLI’ enige nuancering behoeft. Of er sprake was van beperkingen (verschillen met de TD-groep) hing sterk samen met het beroep op talige vermogens. Hoe groter het beroep op talige vermogens, hoe groter de verschillen tussen SLI en TD in het werkgeheugen. Er werd geen verschil gevonden in het onderdrukken van (motorische) responsen. Daarnaast leken de verschillen tussen SLI en TD in informatieverwerkingscapaciteiten af te nemen naarmate kinderen ouder worden. In de adolescentie was er alleen bij zinsrepetitie en nonsendwoordenrepetitie een significant verschil tussen de SLI- en de TD-groep. Deze uitkomsten hadden implicaties voor de derde hypothese, die stelde dat variatie in grammaticaal presteren zou zijn gerelateerd aan problemen in de informatieverwerking. Op basis van de uitkomsten in Hoofdstuk 4 verwachtten we alleen een (negatieve) correlatie te vinden met de twee verbale informatieverwerkingstaken.

Hoofdstuk 5 rapporteert de uitkomsten op de testen voor grammaticaal geslacht. Het grammaticaal geslacht van een woord (zijdig of onzijdig) bepaalt onder andere de vorm van het lidwoord (*de* of *het*) en de vorm van het bijvoeglijk naamwoord (een *grote* man/een *groot* paard). Correct gebruik van lidwoorden is een kwestie van opslag, terwijl er voor de bijvoeglijk naamwoorden ook een regel verworven moet worden (wanneer je een bijvoeglijk naamwoord in een attributieve, onbepaalde, onzijdige, enkelvoudige context gebruikt laat je de vervoeging /ə/ weg).

In dit onderzoek werden zowel lidwoorden als bijvoeglijk naamwoorden getoetst. De resultaten werden gecombineerd om te onderzoeken of geslacht als abstract kenmerk van een zelfstandig naamwoord was opgeslagen, en of de regel voor het vervoegen van bijvoeglijk naamwoorden was verworven.

Grammaticaal geslacht bleek ook in de adolescentie nog moeilijk voor de proefpersonen met SLI (zowel de opslag van lidwoorden als de regel voor bijvoeglijk naamwoorden) en de verwerving van deze grammaticale variabele leek voor de adolescentie al te fossiliseren. Daarbovenop zagen we in de productie van lidwoorden en bijvoeglijk naamwoorden een duidelijk effect van de talige context op het grammaticale presteren. Kinderen en adolescenten met SLI produceerden in het algemeen weinig onzijdige lidwoorden (*het paard*) en correct vervoegde bijvoeglijk naamwoorden (*een groot paard*) en wanneer de linguïstische context complexer werd (*het grote, bruine paard/een groot, bruin paard*) kelderden hun scores terwijl de TD-groepen dit effect niet of nauwelijks lieten zien. De SLI groepen werden dus meer gehinderd in het toepassen van grammaticale kennis en de mate van hinder correleerde (negatief) met de verbale informatieverwerkings-capaciteiten. De problemen die Nederlandse kinderen en adolescenten met SLI ondervinden in het toekennen van grammaticaal geslacht komen dus enerzijds voort uit een gebrek aan grammaticale kennis, maar fouten in de productie van grammaticaal geslacht kunnen ook een probleem in het toepassen van grammaticale kennis betekenen.

In *Hoofdstuk 6* worden de resultaten voor de tweede grammaticale variabele van dit onderzoek, de afstemming tussen werkwoord en onderwerp, gepresenteerd. In het Nederlands stemmen we de vorm van het werkwoord af op de kenmerken van het onderwerp. Zo duiden we aan of het onderwerp enkelvoud of meervoud is, en maken we in het enkelvoud ook onderscheid naar persoon (*ik loop* versus *hij loopt*). De resultaten lieten zien dat de regels voor het vervoegen van werkwoorden over het algemeen verworven zijn als kinderen met SLI de adolescentie hebben bereikt. Wel werden er blijvende verschillen gemeten tussen de SLI- en de TD-groepen in het aantal vervoegingsfouten dat gemaakt werd. Adolescenten met SLI produceerden meer correcte werkwoordsvormen dan de kinderen met SLI en vervoegden vaker het lexicale doelwerkwoord (in plaats van het inzetten van een hulpwerkwoord-constructie, wat vaak door jongere kinderen met SLI werd gedaan). In het enkelvoud maakten zij echter nog steeds een aantal vervoegingsfouten. Fouten kwamen vaker voor wanneer het werkwoord een complexere fonologische vorm had, of wanneer het werkwoord werd uitgelokt in een complexere zin. Deze invloeden van de linguïstische context waren gecorreleerd met de (beperkte) verbale informatiecapaciteiten. Dat bevestigt het idee dat ook grammaticale aspecten die verworven zijn in SLI nog steeds ‘kwetsbaar’ zijn in het taalproductieproces wanneer er een groter beroep wordt gedaan op de informatieverwerkingscapaciteiten van een kind.

Hoofdstuk 7 bespreekt tenslotte de resultaten van het begrip en de productie van relatieve bijzinnen. De problemen met relatieve bijzinnen die vaak worden gerapporteerd in de literatuur over SLI bleken persistent, hoewel het (correct) gebruik van deze complexe zinsstructuren wel toenam in de oudere groepen. In de literatuur wordt een onderscheid gemaakt tussen relatieve bijzinnen waarbij het zinselement waarop de bijzin betrekking heeft het onderwerp is in de bijzin (*dit is het meisje dat het ijsje eet)* en relatieve bijzinnen waarbij het element waarop de bijzin betrekking heeft in de bijzin lijdend voorwerp is (*dit is het ijsje dat het meisje eet)*. Het eerste type zinnen wordt vaak aangeduid als een 'subject relatief' terwijl het laatste type zinnen wordt aangeduid als aan 'object relatief'. Vanuit de literatuur is de hypothese dat kinderen met SLI vooral moeite hebben met het begrip en de productie van objectrelatieven. Objectrelatieven bleken in dit onderzoek inderdaad moeilijker te produceren dan subjectrelatieven en dit verschil was ook groter in de SLI-groepen. Het verschil tussen subject- en objectrelatieven bleek gerelateerd te zijn aan de verbale informatieverwerkingscapaciteiten van de proefpersonen: de verschillen waren groter bij kinderen met zwakkere scores op de zinsimitatietest en de nonsenswoordenrepetitietest.

Bij eerder gevonden problemen in het begrip van object relatieven in SLI plaatst dit onderzoek echter een kanttekening. De SLI-groepen hadden niet meer moeite met het begrijpen van de structuur van een object relatief dan hun leeftijdgenoten wanneer de object relatieven prototypische kenmerken hadden. Een prototypische Nederlandse object relatief heeft een levend onderwerp en een niet-levend lijdend voorwerp, zoals in de voorbeeldzin ‘*dit is het ijsje dat het meisje eet’.* Wanneer de object relatieven minder prototypisch waren, bijvoorbeeld *dit is de prinses die de ridders vangen*, verscheen het verschil tussen SLI en TD. De problemen met het begrijpen van object relatieven lijken dus deels afhankelijk van hoe natuurlijk de zin is.

In *Hoofdstuk 8* word geëvalueerd of de drie hypothesen die centraal stonden in dit onderzoek werden bevestigd. De eerste hypothese, dat er persistente grammaticale verschillen zouden zijn tussen proefpersonen met en zonder SLI in de grammaticale aspecten die we toetsten, werd bevestigd. Voor het markeren van grammaticaal geslacht en het produceren van (object) relatieven waren de verschillen tussen SLI en TD zo groot dat kan worden gesproken van een blijvend grammaticaal probleem. Voor het vervoegen van werkwoorden werden ook persistente verschillen gevonden, maar deze verschillen kunnen niet langer als diagnostisch kenmerk worden bestempeld. Kinderen en adolescenten met SLI haalden over het algemeen hoge scores en de regels voor het vervoegen van werkwoorden lijken te zijn verworven voor de adolescentie. Grammaticale problemen zijn dus persistent in SLI, maar het grammaticale profiel (welke grammaticale aspecten uitval tonen) verandert over de tijd.

De tweede hypothese was dat we variatie zouden vinden in het grammaticale presteren van individuen, afhankelijk van de linguïstische context waarin grammaticale elementen werden getoetst. We namen aan dat deze variatie groter zou zijn in de SLI-groepen dan in de groepen met een normale taalontwikkeling. Deze hypothese werd bevestigd voor veel van de contextfactoren die we in de grammaticale taken hadden gevarieerd. Prestaties waren over het algemeen lager in contexten waarvan we hadden aangenomen dat die complexer zouden zijn. Deze effecten waren vaak groter in SLI dan in TD. Op basis van deze uitkomsten kunnen we een aantal linguïstische factoren aanwijzen die het grammaticale presteren van Nederlandse kinderen met SLI sterk beïnvloeden.

Tenslotte verwachtten we een relatie tussen de mate van variatie in grammaticaal presteren en (beperkingen in) de informatieverwerkingscapaciteiten. Deze hypothese werd onderzocht door de correlatie te berekenen tussen de effecten van de linguïstische context en de verbale informatieverwerkingstaken die blijvende verschillen lieten zien tussen de groepen met en zonder SLI. Bijna alle effecten van context waren gecorreleerd met de uitkomsten op de verbale informatieverwerkingstaken. Bij kinderen met zwakkere informatieverwerkingscapaciteiten had de complexiteit van de linguïstische context een groter effect op hun presteren (en vice versa). Daarmee werd Bishop’s idee, dat grammaticale elementen kwetsbaar zouden zijn in talige contexten die een hoger beroep doen op de informatieverwerking (Bishop 1994, p. 528), in dit onderzoek bevestigd. We beweren niet dat SLI wordt *verklaard* door een beperking in de informatieverwerking, noch beweren we dat de relatie tussen grammaticaal presteren en informatieverwerking een duidelijke causaliteit heeft. De resultaten van deze studie laten echter wel zien dat fouten in de taalproductie van kinderen met SLI niet noodzakelijkerwijs betekenen dat ze grammaticale kennis missen maar dat deze ook kunnen ontstaan vanuit onvermogen om grammaticale kennis toe te passen in de taalproductie. Dit onderzoek laat zien dat dat vooral gebeurt wanneer er veel van de informatieverwerkingscapaciteiten wordt gevraagd. Toekomstig onderzoek moet verder in kaart brengen onder welke condities de implementatie van kennis in SLI wordt gehinderd en gefaciliteerd.

Curriculum Vitae

Iris Duinmeijer was born on August 14th, 1984 in Hoorn, The Netherlands. After obtaining her high school diploma (VWO) at the Adriaan Roland Holstschool in Bergen, she followed the interdisciplinary bachelor program Beta-Gamma and specialised in Linguistics, obtaining degrees in both bachelors in 2007. She then started the research master Linguistics and graduated cum laude in January 2010.

While doing her bachelor in Linguistics, Iris developed a particular interest in language development and language pathology through an internship in the PhD project of Esther Parigger. In the research master, she specialised in these areas and did a clinical internship in the Speech & Language Centre of the Koninklijke Kentalis in Eindhoven. Within this centre, Iris also wrote her master thesis on *'Narrative abilities of children with Specific Language Impairment'*, under the supervision of Annette Scheper and Jan de Jong. The results were presented at several conferences and appeared as an article in an international journal.

Directly after her graduation, Iris started working as a clinical linguist in the Speech & Language Centre in Eindhoven. In September 2011 she started as a PhD in the project '*Persistent problems in Specific Language Impairment. Deficits in knowledge or in knowledge implementation?*', under the supervision of Fred Weerman, Anne Baker, and Jan de Jong. She collected a large dataset in regular and special educational settings. Iris combined the PhD with her clinical job till the end of 2014, when she went to London for a three-month stay at Nicola Botting's lab.

In the course of the PhD project, Iris taught part of the module 'Language and speech development' and gave several guest lectures on language pathology. Her work was presented at several international conferences, such as the *Child Language Seminar (CLS)* in Manchester (2013), the *International Association for the Study of Child Language (IASCL*) (2014) in Amsterdam and the *Symposium on Research in Child Language Disorders (SRCLD)* in Madison (2014). She also helped organising symposia such as the *Netwerk Eerste Taalverwerving (NET) Symposium* (2012).

Iris is currently working in an educational setting for children with speech and language disorders. She assists parents, pupils and schools in their official request for (special) educational support.