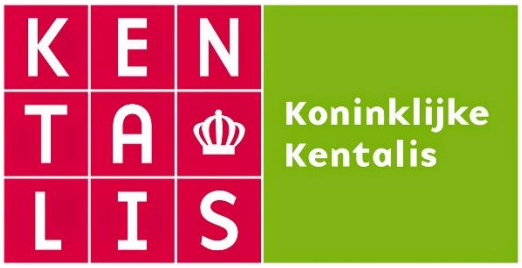
Radboud University Nijmegen and Royal Dutch Kentalis

The efficacy of Story Grammar Training in children with SLI

An efficacy study of narrative intervention as well as the establishment of the predictive value of executive and working memory ability on narrative scores

Radboud University Nijmegen



Name:	Lonneke Janssen
Student number:	4347609
Course:	Master's Thesis
Version:	Final version
Date:	24-06-2017
University:	Radboud University Nijmegen
Master:	Language and Speech Pathology
Institution:	Speech and Language Centre, Royal Dutch Kentalis, Utrecht
First reader:	Dr. A. Scheper
Second reader:	Dr. E. Janse

Foreword

The paper lying before you is the thesis 'The efficacy of Story Grammar Training on children with SLI', which is both an extension into the previous efficacy research of Story Grammar Training (SGT) as well as a study which gives direction for further research into the more general field of children with SLI and their narrative abilities. This thesis has been written to complete my Master's in Language and Speech Pathology at Radboud University in Nijmegen. The wish to further investigate SGT came from my primary supervisor, Annette Scheper, and from the institution she is representing, Royal Dutch Kentalis. I have been engaged in writing this thesis starting from February 2017 until June 2017.

Starting my internship at Kentalis, I faced many practical difficulties which prevented me from collecting data right from the start. Luckily, there was an extensive amount of literature I could start to read up on while waiting for the necessary requirements in working for Kentalis. The result of this quest you will find in the Introduction section of this thesis. Fortunately, my supervisor was always keen to answer my questions, even after numerous e-mails and nervous phone calls.

When conducting research, there is always a point where dreams become reality. I dreamed that I could answer all of my research questions perfectly clearly, without any doubt. Quite soon, however, I was pushed back into reality by the sometimes confusing results of my analyses. This, I quickly learned, is intrinsic to the practice of scientific research. In deciphering the results of my analyses, my supervisor was always happy to help too.

I hope the research I have conducted here will be of help to future Master students, to the clinical practice at Kentalis, and of course to the children experiencing the difficulties of Specific Language Impairment.

I would like to thank my supervisor, my fellow Language and Speech Pathology students for cheering me up, and the rest of my friends for their support and quite useful debates. Their Master's in Communications views have certainly helped me to view things from a different perspective. I would also like to thank my father for keeping me going, asking every time I saw him whether the thesis was 'done already'. Thank you, dad!

I hope you enjoy your reading.

Lonneke Janssen

Nijmegen, June 14, 2017

Abstract

Topic and purpose. The goal of this thesis is to investigate the efficacy of Story Grammar Training (SGT) in improving narrative ability at the macro- and microstructural level, in children with Specific Language Impairment (SLI). Additionally, several cognitive abilities of the same children are evaluated to see whether improvement took place. Further, this study is an initial attempt to investigate a predictive relationship between cognitive and narrative ability.

Method. This study was conducted using data from clinical practice at Royal Dutch Kentalis in Utrecht and Eindhoven. Data was collected starting from the years 2013 to 2016. The subjects of the study were thirty-six children with SLI. Four groups were formed based on the frequency of SGT they had been given. Pre- to post-intervention comparisons of narrative and cognitive ability were conducted, as well as regression analyses to establish possible cognitive predictors of narrative ability.

Conclusions. The results show that children with SLI, regardless of how many times they have received SGT, improve on several narrative variables in the story retelling task. These variables include plot score, MLU5 and the percentage of non-fluencies. Results further showed that some groups improved in the ability of cognitive flexibility and switching, as well as visual attention. It was established that for the group who had completed all sessions of SGT, the task for cognitive flexibility and switching could predict plot score and the mean length of utterance. These results suggest that there is a relationship between pre-intervention cognitive ability and post-intervention narrative scores, meaning that cognitive ability could, at least partly, influence how well the training is being learned.

Recommendations. Further research is needed to ensure the efficacy of SGT, since there was little control over the groups who had not received SGT eight times, thus making it difficult to draw firm conclusions. It is recommended that more research should be done to establish the predictive relationship between pre-intervention cognitive ability and post-intervention narrative ability. Additionally, there should be more participants included who have received SGT eight times, as well as a possible control group which has received a different intervention method, for example to improve morpho-syntactic skills.

Table of contents

1. Introduction

In this section, an extensive background will be given regarding previous research on SLI. First, a short focus of the present study will be discussed, with the goal to immediately have the correct frame of reference while reading literature relevant to SLI and, specifically, to the present study. Next, some general statistics regarding the prevalence of SLI and the effect on later life will be discussed, followed by some theories that try to explain the observed difficulties with language that children with SLI face. In the section following that, the question will be laid forward whether SLI is really specific to language, or might there be difficulties in other aspects of the human mind? Next, the narrative abilities of children with SLI will be discussed, because the present study is specific to these abilities. After this, an overview of different cognitive and narrative intervention studies will be discussed, followed by earlier research into the efficacy of the intervention programme at hand in the present study. In the last part of the Introduction the research questions and hypotheses will be laid forward.

1.1. Focus of this study

This study focuses on a particular group of children with a disorder of language: Specific Language Impairment (SLI). Several definitions for this disorder have been proposed, however the most commonly accepted one is that SLI is an impairment in the development in language, whilst nonverbal ability is within normal limits and no further hearing, vision or neurological damage is present. Some children with SLI have particular difficulty in telling stories (or narratives), whether it's in a social setting or in a more formal setting such as a classroom. A new intervention method aiming to train and thus enhance this narrative ability has been developed: Story Grammar Training (SGT; originally developed by MindWing Concepts in 1991 as Story Grammar Marker (SGM)), in which children are trained on telling the most important features of a story, such as the protagonist, main event, consequences and feelings, also called story grammar elements. Children can use these newly learned features and employ them in their everyday life. SGT has been developed by Daamen, De Groot and Scheper (2013), with SGM as the basis. However, SGM was more logopaedic, while SGT focuses more on the linguistic foundations. In this study, the efficacy of Story Grammar Training will be tested, by means of a pre- and post-test of children who have received training over an eight-week period. In addition to this, children's executive functioning has been assessed and it will be tested in what way these abilities contribute to the child's narrative ability.

But first, a literature review of all relevant aspects regarding narrative and cognitive functions intervention and characteristics of SLI will be discussed.

1.2. Specific Language Impairment: prevalence and effect on (later) life

According to the frequently cited epidemiologic study conducted by Tomblin, Records, Buckwalter, Zhang, Smith and O'Brien (1997), approximately 7.4% of the U.S. kindergarten population have specific language impairment. Specific Language Impairment (SLI) is a disorder of language whilst nonverbal ability as measured by intelligence quotient (IQ) is within normal limits (generally speaking, this means above 85). Additionally, hearing ability should be normal and vision should also not be impaired, as well as the absence of brain damage in any form (Conti-Ramsden, 2008). These exclusionary criteria are described in order to make clear that, in principle, these children should not have any problems in developing language typically, however they do have an atypical or severely delayed language development.

Having SLI can have a negative impact on further life. Conti-Ramsden (2008)

investigated literacy, academic achievement, friendships and emotional health in a longitudinal study following children that were diagnosed with SLI from seven to sixteen years old. Only 8% of the adolescents that were followed did not present any difficulty in any of the four areas at age sixteen. Most adolescents had difficulties in two areas, mostly in literacy and academic achievement. These results suggests that it is important to identify and treat SLI starting at a very young age, in order to minimise the negative impact on later life.

1.3. Theories regarding SLI

Several theories that try to explain the observed impairment in SLI have been proposed. The first is a linguistically oriented theory, stating that grammatical difficulties are the primary deficit. It has been found repeatedly that children with SLI have disproportional difficulty with tense-marking phonemes (Rice, 2000). In a study by Miller, Leonard and Finneran (2008), it was observed that adolescents that were diagnosed with SLI as a child still had more difficulty identifying grammatical violations. Additionally, children with SLI make more errors on verb inflection than typically developing children (Blom, De Jong, Orgassa, Baker & Weerman, 2013). However, in regarding SLI only as a grammatical deficit, one may not fully account for the variation in linguistic difficulties, for instance, word finding difficulties have also been reported. Therefore a theory emphasizing a general process underlying language ability could be more easily accepted and could account for the variability in the language disabilities that are observed. The second theory is such a theory.

This theory regards a more general cognitive ability, namely that limited processing capacity causes language difficulties. Ellis Weismer (2000) explains this theory as follows:

> Cognitive resources which can be allocated to different tasks are limited and when demands exceed available resources, the processing and storage of linguistic information is compromised. (Ellis Weismer 2000, p. 163).

There have been multiple studies supporting this theory, indicating more difficulty with processing in children with SLI than typically developing children (Dodwell & Bavin, 2008; Colozzo, Gillam, Wood, Schnell & Johnston, 2011; Montgomery, 1995). However, this theory is very difficult to test in a controlled, experimental setting, as nearly all kinds of observed impairments could in some way be explained by a limited processing capacity theory.

The third and final theory that seeks to explain the observed disorder in SLI is the Procedural Deficit Hypothesis (PDH; Ullman & Pierpont, 2005). It criticises the impaired grammar and the processing theory in that both do not account for the heterogeneity and variability within the SLI population and that neither of these theories can explain all observed impairments. They pose that individuals with SLI have an impairment in the procedural memory system. Procedural memory is implicated ‘in the learning of new, and the control of long-established, motor and cognitive ‘skills’, ‘habits’, and other procedures […].’ (Ullman et al., 2005). Grammar in the typically developing population is understood to be stored in the procedural memory system, as well as the process of lexical retrieval, mental imagery, working memory and rapid temporal processing (Ullman et al., 2005, p. 402). On the other hand, lexical knowledge are understood to be stored in the declarative memory system. The PDH is essentially a broader theory and it allows for more of the observed variation and heterogeneity to be explained, as all linguistic and cognitive processes associated with SLI are dependent of procedural memory.

The most recent evidence for the PDH is laid forward by Lum, Conti-Ramsden, Page and Ullman (2012). They showed that in typically developing children, lexical abilities were related to declarative memory, as was expected, and grammatical abilities were related to procedural abilities, which was also expected. However, in children with SLI, grammatical

abilities were related to declarative memory and not to procedural memory. This gives further evidence for the hypothesis that the procedural memory is impaired, as it is unable to store grammar. The declarative memory system functions as a compensatory system in children with SLI, storing lexical items as well as grammar. Thus, the declarative system takes over functions of the impaired procedural system. These findings support the PDH and are further evidence for procedural memory deficits in individuals with SLI, because grammatical difficulties and impairments are one of the key linguistic features of SLI.

1.4. SLI: specific to language?

As can be deducted from the last two theories regarding SLI, there is considerable debate about whether SLI is, in fact, specific to language, or if there are other deficits such as a general processing deficit or a procedural memory deficit. Within this line of thinking, researchers explored the fields of executive functioning and working memory, hypothesising that these aspects of cognitive ability could possibly be impaired in children with SLI. Theoretical debate about the definition of and distinction between executive functioning (EF) and working memory (WM) is still present, so one should be cautious about the definitions used for EF and WM when reviewing the literature. Sometimes, a study exploring EF in children with SLI also use measures for WM, thus assuming that WM is a part of EF. According to Baddeley (2010), WM refers to 'the system or systems that are assumed to be necessary in order to keep things in mind while performing complex tasks such as reasoning, comprehension and learning' (p. 136). As such, it can be argued that WM is necessary to complete EF tasks, instead of WM being a distinct aspect of EF. Rather, EF is dependent on WM. However, different studies and experiments use different definitions of WM and EF. This is unfortunate, but also inevitable. In the following paragraphs, EF will first be discussed, followed by WM.

According to a meta-analysis conducted by Pauls and Archibald (2016), inhibition and cognitive flexibility are impaired in children with SLI when scores are compared to age-matched typically developing children or to norm standards. According to moderator analyses, children with SLI seem to be impaired in inhibition and cognitive flexibility regardless of the task's linguistic demands. These linguistic demands can often be a confounding factor in examining cognitive abilities in children with SLI, so this finding is an important one.

Lukács and Ladányi (2015) investigated seven-year-old children's working memory, updating, inhibition and fluency skills. They found significant differences in verbal complex WM and verbal fluency tasks, but not in updating and inhibition. Noteworthy is that these impairments in complex WM and fluency were absent in the non-verbal domain, which they also investigated. Thus, these researchers did find a difference, although this difference was in the verbal domain only.

Victorino and Schwartz (2015) examined auditory attention control and inhibition in children with SLI of around ten to eleven years old. The results showed that children with SLI were not less accurate on the auditory attention and inhibition task. However, reaction times were significantly slower in children with SLI compared to typically developing children. The authors pose that this could be due to the linguistic nature of the task and not, more intuitively, to a processing deficit. Children with SLI, they say, could be slower in recognizing the auditory word, retrieving the lexical representation and comparing it with the visual stimulus.

Reichenbach, Bastian, Rohrbach, Gross and Sarrar (2016) investigated executive function skills and working memory skills of children with SLI, aged around five years. Tasks measured cognitive flexibility, inhibition, processing speed and phonological short-term memory. Results showed that children with SLI differed from typically developing children only in phonological short-term memory.

This review of available literature is certainly not exhaustive, but the goal was to report recent literature and different executive functions. It is noteworthy that Reichenbach et al. (2016) did not find significant differences in processing speed, this might be further evidence that the found difference in reaction time in Victorino et al. (2015) is not due to a slower processing speed, but rather to slower recognition and retrieving of lexical and phonological representations. Together with the evidence found by Lukács et al. (2015), many found differences in EF might be due to linguistic demands of these tasks. This does not seem to be the case in cognitive flexibility and inhibition, as was made clear from the meta-analysis by Pauls et al. (2016). Impairments in cognitive flexibility and inhibition regardless of linguistic demands is further evidence to support the claim that the language impairment in SLI might not be as specific to language as we have thought in the past.

Within WM, a few methodological distinctions can be made. The first important distinction is verbal versus non-verbal WM. A second distinction is the possibility to assess phonological WM, which can be seen as a subtype of verbal WM. These three types of working memory will be discussed in this section.

Studies have shown that children with SLI have an impaired phonological WM (Girbau, 2016; Montgomery, 1995). Montgomery (1995) found that there seems to be a length effect present in children with SLI, where their performance on non-word repetition (the most common way to assess phonological working memory) decreases as the amount of syllables in the non-word increases. This is evidence for an impaired WM, as repeating becomes increasingly difficult when the non-words increase in length, and thus the amount of syllables to be recalled. Girbau (2016) found similar results, showing that Spanish-speaking children with SLI produced significantly less non-words correctly than typically developing children, especially in non-words with three, four or five syllables.

In measuring WM, the same confounding factors can play a role as with measuring EF. Many of the tasks said to measure WM are verbal or linguistic in nature, and this could put children with SLI at a disadvantage from the very start. For this reason, Botting, Psarou, Caplin and Nevin (2013) designed an experiment where they asked children with SLI and typically developing children to complete four tasks, each task increasing in the amount of verbal encoding necessary to complete the task. Their prediction was that there would be no difference on performance in the purely non-verbal task, and that differences would become greater as tasks became more verbal. The first prediction turned out to be true, as no differences were found in a block span task (a non-verbal way to assess WM). However, no differences were found between the other three increasingly verbal tasks in children with SLI, showing no decreasing performance as they became more verbal. The children with SLI did perform worse than their TD peers. Because they performed similarly on the non-verbal task and they did not find the predicted difference in the verbal tasks, it cannot be said with certainty that children with SLI have impaired WM. However, the authors put forward that children with SLI may have difficulty with verbal encoding, more so than their TD peers. Executive function abilities might also be an explanation for their performance, children with SLI may not be as readily able to manipulate the input into another modality (visual to verbal) and this deficit may lower performance (Botting et al., 2013). This explanation is further supported by impaired cognitive flexibility found in the meta-analysis by Pauls et al. (2016).

Lum, Conti-Ramsden, Page and Ulman (2012) also found impaired WM in children with SLI, showing lower performance in central executive tasks and an impaired phonological loop. The phonological loop is an important feature of WM, and even more so in describing the deficits in children with SLI, because the phonological loop is used to complete a non-word repetition task. Importantly, Lum et al. (2012) found that the impairments of children with SLI as compared to typically developing children stayed significant when they entered language ability as a covariate. Their WM skills seem to be impaired regardless of these

abilities, as was also shown for inhibition and cognitive flexibility in the meta-analysis conducted by Pauls et al. (2016).

A conclusion that can be drawn from this paragraph on working memory and executive functioning, is that we cannot say with absolute certainty that nonverbal WM or aspects of EF are impaired. Some studies do find impairments in the nonverbal domains of, for instance, cognitive flexibility, inhibition, central executive functioning and phonological loop (Pauls et al., 2016; Lum et al., 2012). However, other studies find impairments only in the verbal domain, like Montgomery (1995), Girbau (2016), Reichenbach et al. (2016) and Victorino et al. (2015). Botting et al. (2013) failed to demonstrate increasingly lower performance with increasing verbal load, and even showed no differences in the nonverbal task.

1.5. Narrative abilities in children with SLI

Next, a literature review of narrative abilities in children with SLI will be discussed, as narrative abilities is the subject of the present study. Narrative tasks can be seen as one of the most complex linguistic tasks to perform, as many different processes all have to come together at once. According to Büttner (2016), several features are seen to be the prerequisites for both production and comprehension of stories. These features include sequentiality (causal-temporal ordering of events), taking into account the mental states of different characters, the generation of various types of inferences, and the ability of switching perspectives (Büttner, 2016, p. 64). On top of this, one needs adequate linguistic ability to be able to form grammatically correct sentences. As can argued from the information that was discussed in paragraph 1.4, cognitive flexibility and, in some cases, WM may be impaired in children with SLI, making it difficult for them to fulfil these prerequisites of producing and comprehending a narrative. For instance, Friend and Bates (2014) conducted a study to establish the relationship between narrative and executive ability. Results showed that executive skills predicted later narrative ability and narrative ability predicted executive skills (Friend et al., 2014), showing that the two may have a relationship and may be complementary to each other.

Berman and Slobin (1994) conducted a crosslinguistic developmental study in which they studied the development of narrative structure across five languages. In this study, they too highlighted the complexity of performing a narrative task. In their case, the narrative task was the telling of a story by means of the Frog Story, a wordless picture book. They say that picture books impose ‘a particular kind of cognitive demand by requiring them to translate spatially static visual sequences into temporally dynamic verbal output’ (Berman et al., 1994, p. 41). Especially the temporal ordering is of special interest here, as the narrator is constantly faced with the problems of deciding, namely how to make references to each of the characters clear to the listener, and ‘how to handle the task of ordering the reporting of parallel activities while advancing thematic progression’ (Berman et al, 1994, p. 42). These references and causal-temporal orderings have to be translated into linguistic features on the one hand, and the narrator has to switch between the picture they see and the language needed to describe the picture, in reference to the previous and next picture on the other hand, as well as constantly evaluating whether the listener comprehends the story, and switching between characters.

Narrative ability has a profound impact on a person’s life, as it plays a ‘central role in autobiographical reasoning, episodic memory, the formation of our individual history and establishing of one’s own identity’ (Büttner, 2016, p. 63). Not to mention the influence of adequate storytelling on developing interpersonal relationships, which are already very important starting from a young age.

Narrative abilities are frequently examined in a set number of ways. Often story length, information or content score, cohesion, grammatical accuracy and complexity is examined. In addition to this, a number of studies investigated non-verbal and executive abilities that are

thought of to be related to narrative ability, such as memory skills, working memory, (auditory) attention and inhibition. In this section, narrative abilities are reported first, followed by the non-verbal and executive abilities that relate to narrative abilities. More often than not, both are investigated at the same time within one study.

In children with SLI, story length seems to be shorter (Pearce, James & McCormack, 2010; Colozzo, Gillam, Wood, Schnell & Johnston, 2011; Duinmeijer, De Jong & Scheper, 2012). However, Bliss and St. Pierre (1997) found equal story lengths in the narratives produced by children with SLI compared to typically developing children. This may have been due to the nature of the task. In this task, children were elicited to produce personal narratives, such as past experiences. Because these stories are more relatable to both groups, and children are more intrinsically motivated to communicate the experience and produce an amusing story, story length could also be expected to be longer than is usually measured in children with SLI. This type of narrative task is not frequently used, partly because it is difficult to control the conditions and prevent children from going 'off-topic'. However, it is a very natural way to produce a narrative and it closely resembles the way children tell stories in real life.

Information or content scores are frequently measured by the amount of plot elements that are produced by children with SLI as compared to typically developing children. Often, children with SLI produce fewer plot elements (Pearce et al., 2010; Colozzo et al., 2011; Duinmeijer et al., 2012). There does not seem to be a difference in the amount of plot elements produced by children with SLI depending on the narrative task, as shown by Duinmeijer et al. (2012). They found no differences in the amount of plot elements produced between a story retelling task and a story generation task.

Cohesion and coherence was only examined by one study that is reported here (Pearce et al., 2010). They found that children with SLI produced more erroneous ties within the story, meaning that children with SLI made more mistakes in referencing to earlier-mentioned subjects and objects or made more mistakes in consistently labelling a lexical item (Pearce et al., 2010, p. 642). It can be imagined that the correct use of ties contributes immensely to the comprehensibility of the narrative.

Grammatical accuracy is often considered as a key impairment of children with SLI. In the context of narratives, grammatical accuracy is often reduced within the SLI groups (Colozzo et al., 2011; Duinmeijer et al., 2012). Again, Duinmeijer et al. (2012) found no differences in grammatical accuracy between the two narrative tasks.

Complexity can be measured in a number of ways. For each study, the measurement of complexity will also be explained. Bliss et al. (1997) found that typically developing children produced more complex narratives than children with SLI. The narratives produced by the typically developing children contained more internal states, goals and consequences than the narratives of children with SLI. However, children with SLI did produce complex stories, just not as often. This indicates a reduced capacity to produce complex stories frequently, and perhaps the children with SLI turned to less complex narrative schemes to communicate their experiences. Pearce et al. (2011) found that an age-matched control group produced more goal-directed and elaborated narratives and the SLI group more non-goal-directed narratives. In Colozzo et al. (2011), a measure for syntactic complexity was used, and it was found that when children with SLI produce more elaborate (in terms of content) and syntactically complex stories, they made more errors per communication unit (p. 1621). Thus, there seems to be a trade-off in children with SLI, where if they try to produce more complex stories, the grammatical accuracy is reduced. The authors pose that this indicates a limitation in processing capacity (Colozzo et al., 2011, p. 1622). Duinmeijer et al. (2012) found that children with SLI produced less complex narratives as measured by the reduced amount of embedded clauses. Here, they did find a difference in the type of narrative task, where the

story retelling task generated more complex narratives. Perhaps when they have heard the story before, they are better able to produce a linguistically complex story.

Lastly, different kinds of non-verbal and executive functions have been related to narrative ability. Dodwell and Bavin (2008) found that a central executive task (where children had to retrieve information from their memory and execute a manipulation on that information) correlated significantly with a story recall task, and a story generation task correlated significantly with the central executive task too, as well as a word span task. These results indicate that children who performed well on memory tasks, also performed well on the narrative tasks. Duinmeijer et al. (2012) found that the two narrative tasks (recall and generation) seem to pose demands on different cognitive abilities. Retelling plot elements was relatively more related to verbal WM, where generation of plot elements was relatively more related to sustained auditory attention. An explanation for this is the length of the story book for each task. The story generation task contained 24 pictures and has two protagonists, while the story retelling task has only 12 pictures and one protagonist (Duinmeijer et al., 2012, p. 553). Thus, sustained auditory attention seems more important in the story generation task.

As can be concluded from this paragraph on narrative abilities, children with SLI seem to be severely impaired on these kinds of tasks. Different nonverbal and executive functions have been related to these tasks, indicating underlying processes that may play a part in this deficit. In the next section, different intervention methods will be discussed. First, two intervention methods aiming to improve working memory ability and executive functioning will be reviewed based on existing evidence for these training programs. Next, interventions aiming to improve narrative ability will be discussed.

1.6. Cognitive and narrative intervention methods

1.6.1. Cognitive intervention methods

The goal of any study that seeks to describe certain impairments in children with language difficulties, is to provide therapy that is more suitable for these children, so that certain impairments can be managed and perhaps abilities can be improved. In this light, there are two interventions which are directed at improving cognitive abilities. The first one is called Cogmed Working Memory Training (CWMT; www.cogmed.com). This is ‘a computerized training program designed to improve working memory’ through a game-like interface (Chacko et al., 2014, p. 247). It targets both storage plus processing and manipulation components of verbal and nonverbal WM (Chacko et al., 2014). It has been specifically designed for children with Attention-Deficit/Hyperactivity Disorder (ADHD), of which WM impairment is one of the core characteristics (Chacko et al., 2014). In studies examining this particular training, it has often been reported that the training works well on trained WM outcomes, but mixed findings have been reported for other cognitive outcomes (Chacko et al., 2014).

The second intervention targeting cognitive abilities that will be reported here, is Braingame Brian (Prins et al., 2013), also a computerized training program with a game interface. These games train inhibition, cognitive flexibility and working memory, all three of which are possible impairments in children with SLI. In this section, evidence for the effectiveness in various
(non-)clinical groups of these two interventions will be laid forward.

CWMT has been tested in various clinical and non-clinical groups. Different clinical groups will be reported here, they include children with fragile X-syndrome, ADHD, symptomatic epilepsy and a non-clinical group. The syndromes described in the clinical groups are all associated with attentional and WM deficits, so in theory they may benefit from a WM training. In reporting these findings, the focus has been on whether there was any general improvement of WM abilities, whether there are any long term effects (is there still

improvement after a few months?) and whether these effects are near-transfer (on trained WM tasks) or far-transfer effects (on untrained WM tasks, or on tasks associated with WM such as EF).

In short, most CWMT studies do report improvement in WM ability (Chacko et al., 2011; Au, Berkowitz-Sutherland, Schneider, Schweitzer, Hessl & Hagerman, 2014; Fuentes & Kerr, 2017; Hitchcock & Westwell, 2017). However, these improvements in WM ability can only be useful if these newly acquired skills can be implemented in untrained tasks and when they are maintained over a longer period of time. Long term improvement is only reported in one study, examining children with symptomatic epilepsy (Fuentes et al., 2017). In addition to the necessity of long term effects, it is important that these effects are transferred to unrelated tasks. This is in none of the reported studies the case. All reported improvements are found in the tasks that were measured during the training fase, and were not found when untrained tasks were administered (Fuentes et al., 2017; Hitchcock et al., 2017, Au et al., 2014; Chacko et al., 2011). In addition, two of the four studies report no improvement in parent- or teacher-rated evaluations of children's behaviour following the intervention (Chacko et al., 2011, Hitchcock et al., 2017), although one study of children with fragile X syndrome does report improved parental perception of WM (Au et al., 2014). In addition to this, Hitchcock et al. (2017) found no improvement in control of attention in the classroom. In conclusion, one can argue whether CWMT is a reliable intervention method, as it rarely has long term effect and cannot be transferred to different tasks or behavioural aspects.

The effects of Braingame Brian (Prins et al., 2013) has been examined in two clinical groups that are reported here, namely children with ADHD and children with autism spectrum disorder (ASS). One study, conducted by Prins, Dovis, Ponsioen, Ten Brink and Van der Oord (2011) did not specifically use Braingame Brian to train WM, as in the year of publication the game was not developed. However, they did use a game interface that trained WM, as well as motivation and performance (as measured by fewer errors during the game). In this way, it could be viewed as a precursor of Braingame Brian, as the first author is the same as the author who developed the game. When evaluating these studies, the same criteria have been applied as with the studies examining CWMT. In short, the results reported in these studies are mixed. Dovis, Van der Oord, Wiers & Prins (2015) did find improved visuospatial short term memory and WM, as well as long term effects. However, these long term effects applied to all conditions, including conditions where WM was not being trained actively. Inhibition only improved in two conditions, the full-active condition and the partially-active condition. No far-transfer effects were reported for EF. De Vries, Prins, Schmand and Geurts (2015) did not find improvement of WM and flexibility training in children with ASS, nor did they find any far-transfer effects. Other variables such as parent-rated executive functioning, social behaviour, ADHD behaviour and quality of life did not improve significantly either. Prins et al. (2011) found an improvement in WM and also found far-transfer effects. However, this study did not examine EF as measured by inhibition or cognitive flexibility. In conclusion, in the studies reported here, Braingame Brian does not seem to induce any far-transfer effects, which are desirable in any intervention method, and only one study reported long term effects (Dovis et al., 2015).

In conclusion, these interventions more often than not do not yield any long term effects, nor do the acquired short term effects transfer to untrained tasks.

1.6.2. Narrative intervention methods

When reviewing the literature, careful attention has been paid to whether there was an improvement in narrative ability after the intervention, which kind of narrative was analysed (story retelling or story generation) and whether any maintenance (long term effects) had been measured.

All reviewed narrative intervention studies have shown improvement in narrative ability

(Hayward & Schneider, 2000; Spencer & Slocum, 2010; Petersen, Gillam, Spencer & Gillam, 2010; Petersen, Gillam & Gillam, 2008). However, the individual studies differed in which type of narrative was elicited and whether maintenance was tested. Retelling narrative was elicited in three studies (Hayward et al., 2000; Spencer et al., 2010; Petersen et al., 2008). Within these three studies, Hayward et al. (2000) showed an increased amount of story information (content score) and episodic complexity. In the study by Spencer et al. (2010) children were taught story grammar elements (character, setting, initiating event, character response, plan, actions, consequence, complication, and resolution) to help structure their narratives better. Spencer et al. (2010) found an increase in the number of story grammar elements produced by the children following the intervention. The study by Petersen et al. (2008) was primarily a study to validate and test the reliability of a new scoring instrument for narrative tasks. To do this, they set up an intervention study and scored the narratives by means of this new scoring instrument. The children in this study also received training in story grammar, as well as training in causality, temporal concepts and dialogue. Results showed that scores on the narrative retelling task increased significantly.

The remaining two studies elicited a narrative generation task, in the case of Spencer et al. (2010) this was in addition to the retelling task. The results of Spencer et al. (2010) showed that in three out of five participants, the self-generated stories were considerably more complete than pre-intervention (Spencer et al., 2010, p. 193). Petersen et al. (2010) also elicited a story generation task. They found that the story's macrostructure (as measured by story grammar elements and episodic format) improved significantly. Complexity of the narrative's microstructure improved, as well as the amount of story grammar elements and the children improved in their ability to narrate in complete episodes (Petersen et al., 2010, p. 974). Petersen et al. (2010) measured microstructure too, generally this refers to the linguistic properties of the narrative. It was shown that all participants improved in their ability to include causal relationships (Petersen et al., 2010, p. 975). Additionally, the number of adverbs and elaborated noun phrases increased, as well as the use of mental and linguistic verbs, mean number of communication units, total number of words and number of different words. Not all measures of microstructure improved.

Unfortunately, only one study reported maintenance effects (Spencer et al., 2010). They reported that children maintained their narrative scores above baseline level after a 2-week break (Spencer et al., 2010, p. 193).

In conclusion, the effects of narrative intervention seems positive, as all studies reported improvement in the language impaired children. However, only one study included long term effects. These were positive as well, but this study did not give any statistical analysis whether these data points differed significantly from baseline data points. It seems that these narrative interventions yield bigger improvements in ability than trainings aiming to improve cognitive ability, as one study did not even report significant differences in scores directly after intervention (De Vries et al., 2015).

1.7. Story Grammar Training: previous efficacy studies

Two efficacy studies on Story Grammar Training have been conducted previously, the first in 2014 by Duijf and the second in 2015 by Boersma. The results of these two studies will be discussed in this paragraph. In Story Grammar Training (SGT), narrative microstructure and narrative macrostructure are targeted. The level of microstructure defines the 'local design of a story at sentence level and refers to causal and temporal subordinating conjunctions, coordinating conjunctions, adverbs, cognitive and linguistic verbs, and mean length of utterance' (Boersma, 2015, p. 2). Narrative macrostructure refers to 'the global content of a story and consists of the number of story utterances and the various plotting elements, which form the basis of a coherent narrative' (Boersma, 2015, p. 3). In measuring narrative ability pre- and post-intervention, a story retelling task and a story generation task is administered.

Duijf investigated the efficacy of SGT on a small group of children with SLI in a broad age range (*N*=11, age range 7;0 – 11;02, M_{age}=8;06, *SD*=1;04). On the level of macrostructure, six of the eleven children improved their number of plot elements, however this finding was not significant. On the level of microstructure, there were various individual children who improved on the variables that were examined. Mean length of utterance (MLU), mean length of the five longest utterances (MLU5), number of grammatical correct sentences, number of complex utterances, adverbs, subordinating conjunctions (SC) and coordinating conjunctions (CC) all improved in most of the children, but none of these improvements were significant. The absence of significant results, could be due to a number of reasons. In seven children, the post-intervention measurement was six months after the last training session, where in four children it was directly after the last training session. These two differing time points were treated as two different groups, however no differences were found between these groups either. The broad age range and the fact that gender differences were not representative for the SLI population (five boys and six girls), could all have influenced the absence of significant results.

Boersma (2015) therefore studied the efficacy of SGT in a narrower age range (9;02 – 11;02), but still in a small sample (*N*=7). Boersma (2015) studied the number of plot elements and story length on the level of macrostructure. On the level of microstructure, the number of verb-related errors (VRE), non-verb related errors (N-VRE) and grammatical complexity (GC). GC was subdivided in MLU, MLU5, number of CC and number of SC. Also, receptive narrative skills were assessed through comprehension questions (CQ). The comprehension of the narratives improved, but the difference between pre- and post-intervention was insignificant. The number of PE produced in the retelling task improved for four out of seven children, but this difference was insignificant. The number of PE produced in story generation increased significantly.

The number of CC and SC did not improve significantly from pre- to post-intervention. The usage of the additive CC ‘en’ did improve significantly in the retelling task. The total scores for N-VRE and VRE showed no significant improvements from pre- to post-intervention. However, when examining the different variables more closely, there was a significant decline in the adverb ‘er’ in the retelling task, a significant increase of deletion of nouns in the story generation task, a significant increase in argument structure deletion and congruency errors in VRE. These three last results are not positive, since it is not expected for the narrative performance to decrease after intervention. The increase of congruency errors could be explained by a limited processing theory, stating that when the children have to focus on completing a complex narrative task (the story generation task in this case), show certain trade-offs. In this case, this trade-off was directed towards congruency. Also, the increase of noun deletion may be due to the fact that the story generation task demands more of a ‘independent inquiry of content related words’ (Boersma, 2015, p. 12), which could be impaired in children with SLI.

In conclusion, Boersma (2015) showed slightly more positive results in the improvement of narrative ability than Duijf (2014). However, the small sample size is still a problem, as well as the short intervention period and the limited frequency with which the training was given (Boersma, 2015, p. 11). In the present study, a bigger sample size and more homogeneity on the length of training period and training frequency are desired.

1.8. Research questions and hypotheses

The goal of the present study is to further investigate the efficacy of SGT. Because the training has been given to children with SLI for over three years now, and courses are being given to speech and language therapists to ensure the correct SGT procedures are applied, we hope to find more (significant) improvements in the narrative ability of children with SLI. The first research question is, therefore: to what extent do the narrative skills of children with SLI

from the ages of seven to twelve improve after being given SGT looking at the macro- and microstructural level? The hypothesis accompanying this research question is that the narrative skills of children with SLI improve on both the macro- and the microstructural level, as we have seen in the section on narrative intervention studies that narrative abilities do generally improve after intervention.

The second research question is: to what extent will auditory and visual attention, working memory and executive functioning ability in children with SLI from seven to twelve years old improve after being given SGT? In the section discussing executive ability in children with SLI, we have seen that quite a few executive functioning abilities and working memory capacities can be impaired. Because SGT trains children to explicitly use the story grammar elements in their story, it could be hypothesised that executive functioning and/or working memory ability might improve after training as the usage of story grammar elements might enhance cognitive flexibility and working memory, and maybe also attentional skills. This is because children have visual and sensorial input from the story grammar icons, so these become easier to use within a story. Therefore, children can focus more on the linguistic parts of telling a narrative. However, as can be read in the literature describing working memory and executive functioning interventions, almost no long term or far transfer effects can be seen. From the reviewed literature, there is also no reason to expect attentional skills to improve, but this variable will be taken into account because in this way a more complete picture of cognitive ability will be visible. Thus, it is hypothesised that auditory and visual attention, working memory ability and executive functioning ability will not improve after SGT, as measured by the standard tests to assess these abilities.

The third research question involves to what extent pre-intervention auditory and visual attention, working memory ability and executive ability can predict the narrative scores post-intervention. It can be hypothesised that when individual children have poor attentional, executive functioning or working memory abilities, they will not be able to fully learn to use the story grammar elements efficiently, therefore not benefiting fully from the therapy. It is hypothesised that the attentional, working memory and executive functioning abilities of children with SLI will predict the narrative scores post-intervention, as is expected that impaired executive and working memory ability will block the ability to fully learn and benefit from Story Grammar Training.

2. Methodology

In this section, the methodology of this study is discussed. Firstly, the participants of the study will be addressed based on their relevant characteristics. Secondly, the different assessments to establish the scores on narrative skills, attention, working memory and executive functioning pre- and post-intervention will be discussed, as well as the specific variables that were gathered from these assessments. After this, the procedure of the study will be laid forward. Specifically, this means the therapy at hand will be explained further. In the next section, the experimental design of this study will be discussed. The different dependent and independent variables will be made explicit, for each of the three research questions. Lastly, the statistical procedures to answer the research questions will be explained.

2.1. Participants

In total, thirty-six children participated in this study. All of the children in this study have been diagnosed with SLI by trained therapists at the Speech and Language Center of the Royal Dutch Kentalis in the cities of Utrecht and Eindhoven. Their scores on the Peabody Picture Vocabulary Test III (Dunn & Dunn, 1997, Dutch adaptation by Schlichting, 2005), a non-word repetition test (Rispens & Baker, 2012) and mean ages are given in table 1. All children had normal hearing and vision, suffered no neurological damage that could explain their language difficulties and had Dutch as their native language. As can be seen in table 1, all children had average to above average IQ scores. IQ scores were measured with the WISC-III-NL (Kort, Schittekatte, Bosmans et al., 2005) or SON-R 2½-7 (Tellegen, Winkel, Wijnberg-Williams & Laros, 1998) or SON-R 6-40 (Tellegen and Laros, 2010). Participants often had impaired visual and auditory attention skills, working memory skills and executive function skills. Means, standard deviations and 95% confidence intervals for these skills are given in Table 1 as well. All parents of the children who participated in this study gave their informed consent. The participating children had not received SGT before.

Table 1: Participants' characteristics

Variable	Participants (*n* = 36)
Age	
Age in years (*SD*)	9.06 (1.472)
Range	7 – 12
95% CI	[9.28 – 10.00]
Gender	
Male	18
Female	18
Intelligence quotient	
IQ (*SD*)	95.9 (11.14)
Range	83 – 128
95% CI	[93.09 – 99.61]
Receptive Skills	
PPVT-III-NL (*SD*)	105.66 (9.84)
95% CI	[102.6 – 108.6]
Phonological Short Term Memory	
Non-word repetition	16.79 (5.00)
95% CI	[15.15 – 18.42]
Auditory Attention	
Raw score (*SD*)	6.72 (2.48)
95% CI	[5.79 – 7.59]
Visual Attention	
Raw score	4.43 (1.62)
95% CI	[3.87 – 5.0]
Working Memory (Block Recall)	
Raw score (*SD*)	25.64 (3.54)
95% CI	[24.14 – 26.95]
Executive Functions	
Raw Accuracy Score 'Trollen tellen' (*SD*)	4.47 (1.75)
95% CI	[3.87 – 5.07]
Raw Time Score 'Trollen Tellen' (*SD*)	5.70 (2.25)
95% CI	[4.96 – 6.52]
Raw Score 'Zesdelentest' (*SD*)	10.32 (4.59)
95% CI	[8.44 – 12.04]

95% CI: 95% confidence interval; *SD*: standard deviation; IQ: intelligence quotient; PPVT-III-NL: Peabody Picture Vocabulary Test-Third Edition, Dutch version.

2.2. Assessments

In this section, a detailed description of the different assessments of auditory and visual attention, working memory and executive functioning that have been used on the participants will be given. The tests to establish narrative ability pre- and post-intervention will also be discussed.

Firstly, the test for auditory attention was the 'Tel mee!' test, this is a test for auditory sustained attention and is a part of the Test of Everyday Attention for Children (TEA-Ch, Dutch adaptation by Schittekatte, Groenvynk, Fontaine & Dekker, 2007). In this test, children are required to count the number of times they hear a specific sound on a test item. Children are not allowed to count with their fingers, only with their mind.

The test for visual attention was the 'Ruimteschepen' test, this is a test for visual selective attention and is a part of the TEA-Ch (Schittekatte et al., 2007) as well. In this test, children

are required to find pairs of the same spaceships on a sheet of paper designed for this test, while also paying attention to the number of times they hear a computer sound during the task. The requirements for this test are thus twofold, making it more complex.

The test to assess working memory ability was a Block Recall task, this is a test to assess the visuo-spatial sketchpad, which holds information in visual and spatial form. Here, there has been a very conscious choice for a test without any explicit verbal output, because, as can be read in the literature review, this may be a confounding factor for children with SLI. The Block Recall test is part of the Working Memory Test Battery for Children (WMTB-C, Pickering & Gathercole, 2001, Dutch adaptation by Velner, 2006). In this test, children are first required to recall the location of only one block. Following this, sequences of two and more blocks are presented. Children have to remember the exact order in which the blocks are presented and where they were located.

The tests to assess executive functioning were twofold. Firstly, 'Trollen tellen' was administered, this is a test for cognitive flexibility and the ability to switch between tasks. This test is also a part of TEA-Ch (Schittekatte et al., 2007). In this test, children are required to count the number of trolls that are hidden inside their caves. The children need to switch between counting forward and counting backward. Secondly, the 'Zesdelentest' was administered, which is the Dutch version of a six elements test for children. This test is a part of the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C, Emslie, Wilson, Burden, Nimmo-Smith & Wilson, 2003, Dutch adaptation by Emslie, Tjeenk-Kalff & Krabbendam, 2006). In this test, children are required to complete three tasks, each divided into two separate tasks. The first task is a simple calculating task, the second is a picture naming task and the third is an item search. The assignment is to do a small part of each of the six tasks in total, in a matter of five minutes. They are not required to complete every task, rather than to just do some of it. Not every order of doing the tasks is allowed, among some other specific instructions. The goal of this test is not to see how well the child can calculate, spell or find items, but how well the child can organise doing a small part of every task, and how well the child is able to follow the instructions within the allotted time of five minutes.

To assess narrative ability, two tests have been administered pre- and post-intervention. First, the Bus Story Test (Renfrew, 1997). In this test, children are required to recall and retell the Bus Story that has been told to them by the speech and language therapist. A number of variables are gathered after the children retell the Bus Story, namely, plot score (PS), mean length of utterance (MLU), mean length of the five longest utterances (MLU5), morpho-syntactic scores (MSA), and a non-fluency variable (NV), which depicts the amount of non-fluencies within the child's speech.

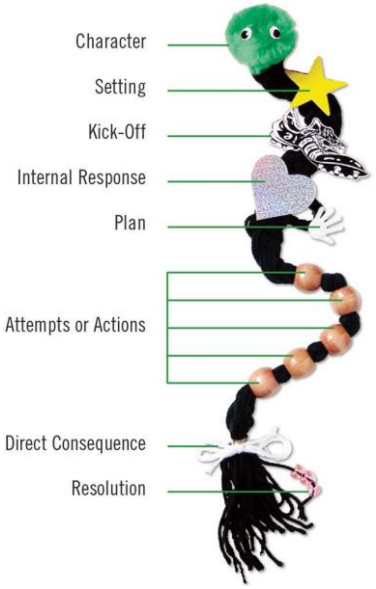
The second test to assess narrative ability is the Frog Story Test (Mayer, 1969). The Frog Story is a wordless picture book. Children have to tell the story about the frog using these pictures. This is a story generation task, as opposed to the Bus Story Test, which is a story recall task. The same variables are gathered when the Frog Story has been administered as the Bus Story Test.

2.3. Procedure

2.3.1. Therapy

The participants received individual Story Grammar Training (SGT) by trained and professional speech and language therapists. SGT is originally from the United States and was developed by Maryellen Rooney Moreau of MindWing Concepts (1991) under the name 'Story Grammar Marker' (SGM). The logopaedic approach of SGM has been translated and adapted for training the narrative skills of Dutch children with SLI by Daamen et al. (2013), focusing more on a linguistic basis for the therapy. The training frequency is once a week, for

about thirty to forty-five minutes, during eight weeks. SGT is given by means of a story braid. This is a tactile-kinaesthetic instrument and it consists of visual and tactile 3D-symbols, each symbolising a different element of a story (for reference, see picture on the next page). There are symbols for the main character, setting, kick-off, feelings, plan, actions or events, direct consequence and resolution. These story elements are seen as the basis for telling a story (the macrostructural level). Because the child sees the different elements in the correct order, and the child is able to feel and shift these elements on the story braid, the macrostructural level is being made more explicit to the child. The therapist chooses a few picture books to train with before SGT starts, which have all the story elements and thus are good training books. Also beforehand, it has been established which conjunctions, nouns and verbs will be provided, and the therapist tries to evoke the child to use these words. This is training on the microstructural level. Each two to four training sessions, a different picture book is picked.



Picture: © MindWings concepts, inc, 1991.

2.4. Experimental design

This study consists of four groups in total. Participants were assigned to each group based on the amount of Story Grammar Training sessions each of them have had. As such, there is a first control group of children who have not received SGT (*N*=7). There is a second control group of children who have received four SGT sessions (*N*=9), a third group which has received five or six SGT sessions (*N*=7) and lastly the experimental group which has received the prescribed eight sessions of SGT (*N*=13). The reasons for not being given the mandatory eight sessions can be varying, the most important one being that the therapist feels that the child may benefit more from a different therapy. The design consists of a pre-intervention examination, where the narrative assessments and neuropsychological assessments have been administered, followed by eight weeks of intervention (of which is zero, four, five or six, or eight times SGT), then a post-intervention examination where the same assessments have been administered, six months after the last intervention. Thus, the post-interventions examinations were not immediately after the last therapy session. In the six months following the intervention, children visit a speech and language therapist frequently to keep practicing with the intervention and techniques they have acquired at Kentalis. After six months children are then asked to return to Kentalis for the post-intervention examination.

2.5. Statistical procedures

For the first research question, which addresses the question of whether the narrative skills of the children from seven to twelve years old improve on the micro- and macrostructural level after being given SGT, a repeated measures ANOVA will be performed. The within subjects variables are the narrative scores as measured by the Bus and Frog Story Test pre- and post-intervention. The between subjects variables are the different groups to which the participants have been assigned. Thus, a repeated measures ANOVA of the narrative scores pre- and post-intervention has been performed with the factor 'Group' as the between subjects variable.

For the second research question, the same experimental design will be used as for the first research question. This research question involves whether or not there is an improvement in the executive abilities of the children. Thus, for this research question, a repeated measures ANOVA with the attentional, working memory and executive function scores as within subjects variable has been performed with the factor 'Group' as the between subjects variable.

For the third research question, it will be examined whether the attentional, executive and working memory skills are predictors for the narrative scores, specifically plot score (PS), morpho-syntactic accuracy (MSA) and mean length of utterance (MLU) for both narrative tasks. Because of the recall component of the non-word repetition task used to describe the participants in the Participants paragraph, this variable will also be used as a predictor. As such, PS, MSA and MLU for both the Frog Story and the Bus Story are the dependent variables and the cognitive scores, including non-word repetition, are the independent variables. The choice has been made to focus on PS, MSA and MLU because the goal of SGT is to make the children better storytellers. This is best demonstrated in an increased PS post-intervention. Secondly, from clinical practice at Kentalis it has been noticed that the quickest progress during therapy is made in MSA. Third, previous research conducted at Kentalis by De Wael (2017), demonstrated good correlations between these narrative and cognitive variables.

These regression analyses will be conducted only for the experimental group, which has received the mandatory eight sessions of SGT. This is because the control groups have received different treatments instead of or next to SGT. As such, the relationship between SGT and narrative scores are not reliable, because different treatments may have influenced the narrative scores post-intervention. Which cognitive variables are chosen to predict which narrative score is further discussed in the Results paragraph. All variables have been entered into the regression analysis stepwise, so that it is clearly visible how much explained variance each variable adds to the regression model.

Before any pre- to post-intervention comparisons are performed, the assumptions for a repeated measures ANOVA will be checked. If a variable does not seem to be normally distributed, a non-parametric Related Samples Friedman's ANOVA by Ranks will be performed to compare the means of these specific variables.

3. Results

3.1. Narrative scores on the macro- and microstructural level

In this paragraph, the narrative scores on the macro- and microstructural level are tested for improvement after the children have or have not received SGT. At first, the assumptions for a Repeated Measures ANOVA, meaning normality and homogeneity, are tested. If the variables violate the normality assumption, a non-parametric statistical test has been performed, as was discussed in the Statistical Procedures section. Post-hoc analyses have been performed when Repeated Measures ANOVA could be used, by applying a Tukey's HSD test. All comparisons have been made at a significance level of .05. Missing values have been excluded and extreme outliers have been adjusted to the mean score of the specific variable. For all narrative variables, the results for the Bus Story are reported first, followed by the Frog Story.

3.1.1. Microstructure: complexity on the sentence level: mean length of utterance (MLU)

Firstly, the assumptions for a Repeated Measures ANOVA have been tested on the data relevant for this variable. These include the pre- and post-intervention MLU scores for both the Bus Story (Renfrew, 1997) and the Frog Story (Mayer, 1969). According to the Kolmogorov-Smirnov test for normality, all measurement points in time and for both stories were normally distributed (all $p > .05$).

According to Levene's Test for Homogeneity, the pre-intervention Bus Story and the post-intervention Frog Story MLU scores were homogeneous ($p > .05$). However, the post-intervention Bus Story and pre-intervention Frog Story MLU scores were not homogeneous ($p < .05$). Thus, these scores are violating the homogeneity assumption.

Table 2: Means and standard deviations on MLU for the Bus Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	7.1 (.98)	7.23 (.78)
Four times SGT	7.68 (1.0)	7.76 (.38)
Five or six times SGT	7.15 (.47)	7.35 (.34)
Eight times SGT	7.07 (.74)	7.57 (.94)

No main effect of time was found (pre-intervention M=7.21, SD=.80; post-intervention M=7.47, SD=.71) on the mean length of utterances in the Bus Story ($F(1,23)=1.958$, p=.175, η^2=.023). There was no main effect of group either ($F(3,23)=.781$, p=.517, η^2=.092). Subsequently, no interaction was found between time and group ($F(3,23)=.404$, p=.752, η^2=.050).

Table 3: Means and standard deviations on MLU for the Frog Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	6.89 (1.01)	7.21 (.58)
Four times SGT	7.44 (.74)	7.54 (1.25)
Five or six times SGT	7.40 (.34)	7.18 (1.13)
Eight times SGT	7.19 (.65)	7.38 (.47)

No main effect of time was found (pre-intervention M=7.2, SD=.65; post-intervention M=7.32, SD=.81) on the mean length of utterances in the Frog Story ($F(1,23)=.541$, p=.47, η^2=.023). No main effect of group was found either ($F(3,23)=.431$, p.732, η^2=.053). Subsequently, no interaction was found between time and group ($F(3,23)=.717$, p=.552, η^2 =.085).

3.1.2. Microstructure: complexity on the sentence level: mean length of the five longest utterances (MLU5)

Firstly, the assumptions for a Repeated Measures ANOVA have been tested on the data relevant for this variable. These include the pre- and post-intervention MLU5 scores for both the Bus Story (Renfrew, 1997) and the Frog Story (Mayer, 1969). According to the Kolmogorov-Smirnov test for normality, all measurement points in time and for both stories were normally distributed (all $p > .05$).

According to Levene's Test for Homogeneity, the pre-intervention Bus Story, and the pre and post-intervention Frog Story MLU5 scores were homogeneous ($p > .05$). However, the post-intervention Bus Story MLU5 scores were not homogeneous ($p < .05$). Thus, these scores are violating the homogeneity assumption.

Table 4: Means and standard deviations on MLU5 for the Bus Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	10.8 (1.46)	11.78 (1.51)
Four times SGT	11.76 (2.62)	12.32 (.89)
Five or six times SGT	10.4 (1.70)	11.73 (1.42)
Eight times SGT	10.87 (1.48)	12.18 (2.71)

A main effect of time was found (pre-intervention M=10.91, SD=1.72; post-intervention M=12.0, SD=1.83) on the mean length of the five utterances in the Bus Story (F(1,23)=8.11, p=.009, η^2=.261). This means that after being given SGT, the mean length of the five longest utterances increased from 10.9 words to 12 words on average. No main effect of group was found (F(3,23)=.372, p=.774, η^2=.046). Subsequently, there was no interaction between time and group (F(3,23)=.221, p=.881, η^2=.028).

Table 5: Means and standard deviations on MLU5 for the Frog Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	12.31 (3.11)	13.63 (2.18)
Four times SGT	12.56 (2.22)	13.6 (3.18)
Five or six times SGT	12.73 (1.07)	12.4 (2.53)
Eight times SGT	12.84 (1.26)	12.62 (1.30)

No main effect on time was found (pre-intervention M=12.63, SD=1.93; post-intervention M=13.02, SD=2.17) on the mean length of the five longest utterances in the Frog Story (F(1,23)=1.078, p=.310, η^2=.045). No main effect of group was found either (F(3,23)=.094, p=.963, η^2=.012). Subsequently, there was no interaction between time and group (F(3,23)=1.007, p=.408, η^2=.116).

3.1.3. Microstructure: morpho-syntactic accuracy (MSA)

Firstly, the assumptions for a Repeated Measures ANOVA have been tested on the data relevant for this variable. These include the pre- and post-intervention MSA scores for both the Bus Story (Renfrew, 1997) and the Frog Story (Mayer, 1969). According to the Kolmogorov-Smirnov test for normality, the pre-intervention MSA scores in the Bus Story were not normally distributed (p=.024). However, all other scores were normally distributed (all p>.05).

According to Levene's Test for Homogeneity, the pre-intervention MSA scores on the Bus Story were not homogeneous and thus violate this assumption (p=.008). However, all other scores did not violate this assumption (all $p \geq .05$).

Table 6: Means and standard deviations on MSA for the Bus Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	8.14 (2.34)	7.14 (3.24)
Four times SGT	7.0 (3.08)	6.40 (2.07)
Five or six times SGT	3.17 (.753)	4.17 (2.93)
Eight times SGT	7.78 (3.87)	6.78 (3.83)

Because the pre-intervention MSA scores in the Bus Story were not normally distributed, the Friedman's ANOVA statistic is reported here. It revealed no significant difference between the MSA scores pre-intervention (*M*=6.70, *SD*=3.36) and post-intervention (*M*=6.22, *SD*=3.25), $\chi^2(1)=.043$, *p*=.835.

Table 7: Means and standard deviations on MSA for the Frog Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	20.29 (6.73)	19.86 (9.23)
Four times SGT	14.80 (6.54)	13.0 (6.89)
Five or six times SGT	8.50 (4.14)	6.17 (3.06)
Eight times SGT	22.67 (14.60)	17.22 (10.27)

No main effect of time was found (pre-intervention *M*=17.44, *SD*=10.3; post-intervention *M*=14.67, *SD*=9.38) on the number of ungrammatical utterances within the Frog Story (*F*(1.23)=2.017, *p*=.169, η^2=.081). A main effect of group was found (*F*(3,23)=3.849, *p*=.023, η^2=.334). Post-hoc analysis revealed a significant mean difference (*MD*=12.74, *SE*=4.42) between the group that has not received SGT and the group that has received SGT five or six times (*p*=.039). Overall, the group that received SGT five or six times produced 12.74 less ungrammatical utterances in the Frog Story. Furthermore, post-hoc analysis revealed a significant mean difference (*MD*=12.61, *SE*=4.19) between the group who received SGT eight times and the group who received SGT five or six times (*p*=.029). Overall, the group that received SGT eight times produced 12.61 more ungrammatical utterances than the group who received SGT five or six times. No interaction effect between time and group was found (*F*(3,23)=18.099, *p*=.718, η^2=.056).

3.1.4. Microstructure: non-fluencies on word- and sentence level (NV)

Firstly, the assumptions for a Repeated Measures ANOVA have been tested on the data relevant for this variable. These include the pre- and post-intervention MSA scores for both the Bus Story (Renfrew, 1997) and the Frog Story (Mayer, 1969). According to the Kolmogorov-Smirnov test for normality, all scores were normally distributed (all *p*>.05).

According to Levene's Test for Homogeneity, all variables were homogeneous (all *p* >.05). Post-hoc analyses have been performed on the Repeated Measures ANOVA using Tukey's HSD to reveal mean differences between groups. Missing values are excluded and extreme outliers have been adjusted to the mean score.

The results of the plot scores for the Bus Story are given first, followed by the results of the plot scores for the Frog Story. In these results, percental scores are given in addition to the raw scores, because of the nature of the variable being non-fluencies, it is important to know the relativity of the number of non-fluencies in relation to the total number of words or syllables in the utterance. This makes the interpretation of the raw scores more telling.

Table 8: Means and standard deviations of the raw scores (RS) and percental scores (pct) on non-fluencies in the Bus Story

	Pre-intervention RS *M* (*SD*), pct *M* (*SD*)	Post-intervention RS *M* (*SD*), pct *M* (*SD*)
No SGT	29.86 (10.0), 14.54% (4.91)	30.14 (13.35), 13.93% (5.74)
Four times SGT	30.20 (9.99), 13.74% (5.51)	37.40 (15.49), 14.82% (5.03)
Five or six times SGT	27.00 (22.35), 13.42% (9.12)	21.67 (11.61), 9.12% (4.39)
Eight times SGT	31.33 (10.57), 18.21% (6.77)	32.33 (24.81), 12.69% (5.54)

No main effect of time was found (pre-intervention *M*=29.78, *SD*=13.10; post-intervention *M*=30.33, *SD*=17.94) on the raw non-fluency scores in the Bus Story ($F(1,23)=.049$, $p=.827$, $\eta^2=.002$). However, the percentage of non-fluencies in the utterances of the children did differ over time (pre-intervention *M*=15.37, *SD*=6.69; post-intervention *M*=12.61, *SD*=5.37) ($F(1,23)=5.486$, $p=.028$, $\eta^2=.193$). However, no main effect of group was found on the raw scores ($F(3,23)=.550$, $p=.653$, $\eta^2=.067$) or the percental scores ($F(3,23)=.710$, $p=.556$, $\eta^2=.085$). Subsequently, no interaction was found between time and group was found for either the raw scores ($F(3,23)=.436$, $p=.729$, $\eta^2=.054$) or the percental scores ($F(3,23)=2.449$, $p=.089$, $\eta^2=.242$).

Table 9: Means and standard deviations of the raw scores (RS) and percental scores (pct) on non-fluencies in the Frog Story

	Pre-intervention RS *M* (*SD*), pct *M* (*SD*)	Post-intervention RS *M* (*SD*), pct *M* (*SD*)
No SGT	46.57 (18.52), 10.69% (2.47)	51.86 (26.89), 11.63% (4.49)
Four times SGT	49.20 (27.07), 13.08% (7.37)	55.20 (20.14), 14.28% (3.79)
Five or six times SGT	29.50 (15.86), 7.85% (3.64)	28.50 (21.77), 7.02% (2.97)
Eight times SGT	72.00 (44.48), 13.89% (6.59)	60.22 (43.23), 12.26% (5.85)

No main effect of time was found (pre-intervention *M*=51.74, *SD*=33.38; post-intervention *M*=50.07, *SD*=32.40) on the raw non-fluency scores in the Frog Story ($F(1,23)=.006$, $p=.938$, $\eta^2=.000$). No main effect of time was found (pre-intervention *M*=11.57, *SD*=5.60; post-intervention *M*=11.30, *SD*=5.04) on the percental scores of non-fluencies either ($F(1,23)=.010$, $p=.920$, $\eta^2=.000$). No main effect of group was found on the raw scores ($F(3,23)=.550$, $p=.653$, $\eta^2=.067$) or on the percental scores ($F(3,23)=2.299$, $p=.104$, $\eta^2= .231$). Subsequently, there was no interaction effect between time and group on the raw scores ($F(3,23)=.892$, $p=.460$, $\eta^2=.104$) or the percental scores ($F(3,23)=.803$, $p=.505$, $\eta^2=.095$).

3.1.5. Macrostructure: Plot Score (PS)

Firstly, the assumptions for a Repeated Measures ANOVA have been tested on the data relevant for this variable. These include the pre- and post-intervention MSA scores for both the Bus Story (Renfrew, 1997) and the Frog Story (Mayer, 1969). According to the Kolmogorov-Smirnov test for normality, the pre-intervention plot scores in the Frog Story were not normally distributed ($p=.017$). However, all other scores were normally distributed (all $p>.05$).

According to Levene's Test for Homogeneity, all variables were homogeneous (all $p>.05$).

Table 10: Means and standard deviations on PS for the Bus Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	14.57 (4.20)	14.29 (4.92)
Four times SGT	12.40 (5.98)	15.60 (1.67)
Five or six times SGT	10.33 (5.57)	13.67 (6.44)
Eight times SGT	11.00 (3.97)	14.00 (5.35)

A main effect of time was found between pre-intervention plot scores (*M*=12.00, *SD*=4.79) and post-intervention plot scores (*M*=14.29, *SD*=4.85) across all groups ($F(1,24)=6.456$, $p=.018$, $\eta^2=.212$). Except for the group who did not received SGT, all groups increased the amount of plot elements produced in the Bus Story post-intervention. A main effect of group was not found ($F(3,24)=.480$, $p=.699$, $\eta^2=.057$). Furthermore, no interaction effect was found ($F(3,24)=.953$, $p=.431$, $\eta^2=.106$).

Table 11: Means and standard deviations on PS for the Frog Story

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	6.86 (2.61)	7.43 (2.57)
Four times SGT	5.80 (3.56)	4.40 (1.67)
Five or six times SGT	4.83 (2.23)	5.83 (3.37)
Eight times SGT	5.40 (2.12)	6.90 (3.11)

Because the pre-intervention plot scores for the Frog Story were not normally distributed, the Friedman's ANOVA statistic is reported here. It was found that there was no significant difference in the amount of plot elements produced pre- (*M*=5.71, *SD*=2.52) to post-intervention (*M*=6.36, *SD*=2.90), $\chi^2(1)=.391$, $p=.532$.

3.2. Attention, working memory and executive functioning

In this paragraph, the scores on both attention tasks, the working memory task (Block Recall) and the executive functioning tasks have been compared for improvement post-intervention. At first, the assumptions for a Repeated Measures ANOVA, meaning normality and homogeneity, are tested. If the variables violate the normality assumption, a non-parametric statistical test has been performed, as was discussed in the Statistical Procedures section. Post-hoc analyses have been performed when Repeated Measures ANOVA could be used, by applying a Tukey's HSD test. All comparisons have been made at a significance level of .05. Missing values have been excluded and extreme outliers have been adjusted to the mean score of the specific variable.

3.2.1. Auditory attention

According to the Kolmogorov-Smirnov test for normality, both the pre-intervention scores and the post-intervention scores were not normally distributed (both $p<.05$). According to Levene's test for homogeneity, both variables were homogeneous (both $p>.05$).

Table 12: Means and standard deviations of the 'Tel mee!' test for auditory attention

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	7.57 (1.81)	6.86 (2.12)
Four times SGT	8.17 (1.33)	7.25 (.50)
Five or six times SGT	5.60 (2.88)	7.60 (2.88)
Eight times SGT	5.91 (2.81)	7.55 (2.07)

No difference was found in the mean scores pre- (*M*=6.72, *SD*=2.48) and post-intervention (*M*=7.33, *SD*=1.78), $\chi^2(1)=.200$, *p*=.655.

3.2.2. Visual attention

For the visual attention test, scores were given for accuracy and the amount of time it took to complete the task. Both variables will be analysed in this section. According to the Kolmogorov-Smirnov test for normality, all variables were not normally distributed (all *p*<.05). According to Levene's test for homogeneity, all variables were homogeneous (all *p*>.05).

Table 13: Means and standard deviations of the 'Ruimteschepen' test for visual attention

	Pre-intervention Accuracy *M* (*SD*)	Post-intervention Accuracy *M* (*SD*)	Pre-intervention Time *M* (*SD*)	Post-intervention Time *M* (*SD*)
No SGT	19.43 (1.13)	18.43 (2.23)	6.51 (2.38)	5.59 (2.33)
Four times SGT	19.43 (.976)	17.67 (1.75)	5.50 (1.51)	6.16 (2.42)
Five or six times SGT	18.83 (1.47)	19.40 (.894)	5.85 (2.02)	5.01 (1.20)
Eight times SGT	19.00 (.894)	18.83 (1.12)	6.18 (2.58)	5.09 (1.79)

No difference was found in the mean accuracy scores pre- (*M*=19.16, *SD*=1.07) and post-intervention (*M*=18.60, *SD*=1.57), $\chi^2(1)=1.471$, *p*=.225. A main difference was found in the mean time scores pre- (*M*=6.04, *SD*=2.15) and post-intervention (*M*=5.41, *SD*=1.93), $\chi^2(1)=4.840$, *p*=.028. Follow-up analyses revealed that the group who has received SGT four times needed less time on average to complete the task (pre-intervention *M*=5.50, *SD*=1.51; post-intervention *M*=6.16, *SD*=2.42; $\chi^2(1)=4.00$, *p*=.046). The same was found for the group who has received SGT eight times (pre-intervention *M*=6.18, *SD*=2.58; post-intervention *M*=5.09, *SD*=1.79; $\chi^2(1)=6.400$, *p*=.011).

3.2.3. Working memory

According to the Kolmogorov-Smirnov test for normality, both variables were normally distributed (both *p*>.05). According to Levene's test for homogeneity, the post-intervention Block Recall scores were not homogeneous (*p*<.05).

Table 14: Means and standard deviations of the Block Recall test for working memory

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	24.43 (3.69)	28.00 (1.73)
Four times SGT	27.00 (3.61)	27.20 (5.22)
Five or six times SGT	27.00 (2.65)	25.17 (5.54)
Eight times SGT	24.89 (3.59)	24.89 (2.09)

No main effect of time was found (pre-intervention *M*=25.67, *SD*=3.62; post-intervention *M*=26.19, *SD*=2.94) in the Block Recall scores pre- and post-intervention ($F(1,17)=.104$, *p*=.751, $\eta^2=.006$). No main effect of group was found ($F(3,17)=1.362$, *p*=.288, $\eta^2=.194$). Subsequently, there was no interaction effect between time and group ($F(3,17)=2.035$, *p*=.147, $\eta^2=.264$).

3.2.4. Executive functioning: 'Trollen tellen' for cognitive flexibility and switching

For this variable, there was an accuracy score and a time score as well. Both variables will be analysed in this section. According to the Kolmogorov-Smirnov test for normality, all variables were not normally distributed (all $p<.05$). According to Levene's test for homogeneity, the post-intervention time variable was not homogeneous ($p<.05$), but all other variables did not violate this assumption (all $p>.05$). Thus, for this variable, the non-parametric Friedman's ANOVA statistic is reported.

Table 15: Means and standard deviations of the 'Trollen tellen' test for cognitive flexibility and switching

	Pre-intervention Accuracy *M* (*SD*)	Post-intervention Accuracy *M* (*SD*)	Pre-intervention Time *M* (*SD*)	Post-intervention Time *M* (*SD*)
No SGT	3.14 (1.86)	5.71 (.951)	5.76 (2.64)	4.30 (1.31)
Four times SGT	5.00 (2.00)	4.80 (2.78)	6.23 (3.13)	6.06 (1.81)
Five or six times SGT	5.60 (1.14)	5.50 (.837)	4.94 (1.12)	7.38 (5.80
Eight times SGT	4.45 (1.37)	5.42 (1.98)	5.69 (2.01)	5.71 (2.85)

A main effect of time was found for the accuracy variable (pre-intervention $M=4.47$, $SD=1.76$; post-intervention $M=5.40$, $SD=1.71$), meaning that the overall mean accuracy score increased over time, $\chi^2(1)=8.909$, $p=.003$. Follow-up tests revealed that the group who had not received SGT increased the accuracy score significantly, $\chi^2(1)=6.00$, $p=.014$.

No main effect of time was found in the time variable (pre-intervention $M=5.70$, $SD=2.25$; post-intervention $M=5.77$, $SD=3.36$), meaning that the time score did not differ significantly post-intervention, $\chi^2(1)=1.80$, $p=.180$.

3.2.5 Executive functioning: 'Zesdelentest' for organising and planning

For this task, multiple variables were measured, however, due to a too small sample size and too many missing values only the total score is analysed. According to the Kolmogorov-Smirnov test for normality, both variables were not normally distributed (both $p<.05$). According to Levene's test for homogeneity, both variables were not homogeneous (both $p<.05$). Thus, the non-parametric Friedman's ANOVA is reported.

Table 16: Means and standard deviations of the 'Zesdelentest' for organising and planning (total score)

	Pre-intervention *M* (*SD*)	Post-intervention *M* (*SD*)
No SGT	13.50 (.577)	11.25 (3.86)
Four times SGT	11.17 (3.55)	14.00 (1.00)
Five or six times SGT	9.33 (5.85)	12.33 (1.75)
Eight times SGT	9.00 (5.05)	11.56 (3.96)

No difference was found in the total score on the Zesdelentest pre-intervention ($M=10.32$, $SD=4.59$) and post-intervention ($M=12.05$, $SD=3.12$), $\chi^2(1)=.053$, $p=.819$.

3.3. Regression analyses for the Bus Story

In this section, the regression analyses for the Bus Story are discussed. Significance values are set at the .05 level. Ninety-five percent confidence intervals are reported for all coefficients. For further explanation of the statistical procedures, these can be read in the Methodology section. First the regression analysis for PS is discussed, next the MSA, and lastly the MLU

scores.

For the Bus Story, the variables Block Recall (working memory), non-word repetition (phonological short term memory), the accuracy and time score for Trollen Tellen (cognitive flexibility and switching) and the total score for the Zesdelentest (planning and organising) were taken as independent variables. The plot score in the Bus Story post-intervention was taken as dependent variable in this model. All variables were entered stepwise in the order presented above.

Table 17: Linear model of predictors of Bus Story PS post-intervention, with 95% confidence intervals reported.

	B	*SE B*	*β*	*p*
Constant	6.307 [-3.149 – 15.763]	3.864		$p = .154$
Accuracy Score 'Trollen Tellen'	1.961 [.005 – 3.916]	.799	.708	$p = .005$

The multiple regression showed that the accuracy score for 'Trollen Tellen' explained 50% percent of the variance found in the Bus Story PS post-intervention ($F(1,6)=6.017$, $p=.05$). All other variables taken in to the analysis were excluded from the analyses, because the *t*-statistics were not at the .05 significance level. The accuracy score for 'Trollen Tellen' showed to be a significant predictor of the Bus Story PS post-intervention ($B=1.961$, $p=.005$).

For the Bus Story MSA scores post-intervention, none of the above mentioned variables were entered into the model. This means that none of the variables had a significant *t* statistic and thus these variables could not predict the MSA scores post-intervention.

For the Bus Story MLU scores post-intervention, the variables that were chosen as possible predictors were the accuracy and time scores for 'Trollen Tellen', the total score for the 'Zesdelentest', the Block Recall scores and the auditory attention scores. These were the independent variables. The dependent variable was the post-intervention MLU score. All variables have been entered stepwise.

Table 18: Linear model of predictors of Bus Story MLU scores post-intervention, with 95% confidence intervals reported.

	B	*SE B*	*β*	*p*
Constant	5.811 [4.019 – 7.603]	.732		$p < .001$
Accuracy Score 'Trollen Tellen'	.417 [.046 - .787]	.151	.747	$p = .033$

The multiple regression showed that the accuracy score for 'Trollen Tellen' explained 55.8% of the variance found in the MLU scores post-intervention ($F(1,6)=2.754$, $p=.033$). All other variables were excluded from the model, since their *t*-statistics were not significant. The accuracy score for 'Trollen Tellen' showed to be a significant predictor of the Bus Story MLU scores post-intervention ($B=.417$, $p=.033$).

3.4. Regression analyses for the Frog Story

In this section, the regression analyses for the Frog Story are discussed. Significance values are set at the .05 level. Ninety-five percent confidence intervals are reported for all coefficients. For further explanation of the statistical procedures, these can be read in the Methodology section. First the regression analysis for PS is discussed, then the MSA scores,

and lastly the MLU scores.

For the Frog Story, the variables visual attention (measured by the 'Ruimteschepen' task), Block Recall (working memory), accuracy and time score for 'Trollen Tellen' (cognitive flexibility) and the total score for the Zesdelentest (planning and organising) are taken as independent variables. The Frog Story PS post-intervention has been taken as dependent variable. All variables were entered stepwise into the regression model.

None of the variables were entered into the model, meaning that none of the *t* statistics were significant and thus could not predict the PS post-intervention.

For the Frog Story MSA scores post-intervention, the same variables were used. The same results were found; none of the variables were entered into the regression model.

For the Frog Story MLU scores post-intervention, the same variables were used as the Bus Story MLU scores. None of the variables were entered into the model, meaning that none of these variables is a significant predictor of the post-intervention MLU scores in the Frog Story.

4. Discussion

4.1. Narrative scores on the macro- and microstructural level

In this particular part of this study, four groups were compared with each other based on their mean scores on various narrative variables, on both a macro- (story level) and micro- (sentence level) structural level. First, significant results will be discussed here for further interpretation. They will also be related to the results found in previous SGT efficacy studies, as well as to previous narrative intervention studies, which were both reviewed in the Introduction. Second, relevant insignificant results will be reported and discussed for further interpretation, again comparing these results to earlier research. Third, some factors which may have impacted the study in a negative way will be discussed. Lastly, some recommendations for further investigations will be proposed.

A significance difference in time was found on the MLU5 scores in the Bus Story. However, there was no difference between groups, meaning that each group increased in the mean length of the five longest utterances after eight weeks of intervention, be it eight weeks of SGT or otherwise. On the MSA scores in the Frog Story, additional significant results were found. Compared to children who had received SGT five or six times, the group who had not received SGT produced around thirteen more ungrammatical utterances. However, this finding was not time-specific, meaning that there was no interaction effect between these groups and the pre- and post-intervention scores. Ungrammatical utterances were not more or less in number post-intervention for either group. Furthermore, it was found that the MSA scores in the Frog Story differed between the group who had received SGT five or six times and the group who had received SGT eight times. The group who had received SGT eight times also produced around thirteen more ungrammatical utterances than the group who had received SGT five or six times. Again, this finding was not time-specific, meaning there was no significant difference pre- and post-intervention.

A difference was found between the percentage of non-fluencies in the Frog Story over time. This means that all groups decreased in the amount of non-fluencies relative to the total amount of words or syllables in the utterance post-intervention. On average, percental scores went from 15.37% pre-intervention to 12.61% post-intervention. This means that all groups benefited from intervention, be it SGT or otherwise. Another difference was found on plot score in the Bus Story. All but the group who had not received SGT, the plot score increased pre- to post-intervention. The plot score of the group who had not received SGT actually decreased from a mean of 14.57 to a mean of 14.29, but apparently this decrease was also significant, since there was no difference in the effect of time between groups.

Previous efficacy studies of SGT or narrative intervention studies have not examined the increase or decrease of the amount (or percentage) of non-fluencies in a narrative retelling task. Thus, it is difficult to relate the findings of the present studies to previous ones. However, it might be expected for the percentage of non-fluencies to decrease within the story retelling task, since it is the least complex of the two tasks and non-fluencies are usually a sign of insecurity or word retrieval difficulties, as well as an inability to apply grammatical rules or to fluently produce sentences. After having had intervention at Kentalis (be it SGT or otherwise), children might have learned to store and retrieve a story more efficiently, resulting in less non-fluencies while retelling the story. In a story generation task, this seems more unlikely because of the complex nature. Apparently, even after intervention, this task is still too complex for the percentages of non-fluencies to decrease, thus children are still unsure or still experience word retrieval difficulties.

Looking at the previous narrative intervention studies discussed in the Introduction, the results on plot score in the present seem to be in line with these studies. Hayward et al. (2000) reported an increased amount of story information (content score) in a story retelling task, which is comparable to the Bus Story. In the present study, a significant increase in plot score

was also found. The same goes for the study by Spencer et al. (2010) and Petersen et al. (2008), which also found an increase in the number of story grammar elements in a story retelling task. Compared to previous SGT efficacy studies, this study did manage to find significant results in plot score. The very first SGT efficacy study, carried out by Duijf (2014), failed to find significantly increased plot score post-intervention for the Bus Story, and the same went for Boersma (2015). This may be due to a larger sample size. However, as was also discussed above, there were no differences between groups, so that the group who did not receive SGT did not differ significantly from the group who received SGT eight times.

While looking at the significant increase of MLU5 in the Bus Story, comparable outcomes can be found in intervention studies training complexity, since MLU5 is a measure of complexity of an utterance (a longer utterance generally means a more complex one). In this case, this result is in line with Hayward et al. (2000) too, who found an increased amount of episodic complexity in a retelling task. Compared to the previous SGT efficacy studies, the significant increase of MLU5 in the Bus Story is a first. Duijf (2014) and Boersma (2015) failed to find this difference. This again, may be due to a larger sample size in the present study. It cannot be said with certainty that it is due to the fact that SGT is being given in a more structured and controlled way, because there was no significant difference between the group who did not receive SGT and the group who received it eight times.

The absence of a significant result in the MSA of the Bus Story, might lie in the finding that there was a significant result in the plot scores of the Bus Story. Colozzo et al. (2011) conducted a study examining the relationship between content and form in the narratives of school-aged children. It was found that children with SLI had a more unbalanced content-and-form-relationship than TD children, meaning that when children with SLI produced longer, more elaborate stories, they made more errors per C-unit (thus, more morpho-syntactic errors which are on the microstructural level). It is hypothesised in this study that children with SLI have a limited processing capacity, which prevents them from telling a highly elaborated story and making few errors at the same time. It seems that children with SLI concentrate on either content accuracy, or morpho-syntactic accuracy. In the present study as well as in Colozzo et al. (2011), it seems that children concentrate more on the plot elements (producing a more elaborated story) than on the morpho-syntactic accuracy (hence, there is no main time effect for the MSA scores) when generating a story by themselves. Apparently, children find it more important to transfer enough plot elements for the listener to be able to follow the story, than to produce it with, let's say, 80% morpho-syntactic accuracy. Boersma (2015) also found a trade-off in congruency on the one hand and completing a complex narrative task (story generation) on the other hand. Her results showed that when the narrative task was more complex, more congruency mistakes were made.

A significant increase in MLU5 was found in the Bus Story post-intervention. This difference might be due to the improved number of plot elements that were produced. If a child produces more plot elements, the mean length of the five longest utterances might also be expected to improve. Indeed, from a simple correlational analysis (Pearson's *r*) it seemed that there was a moderate to strong correlation between the MLU5 scores and plot scores post-intervention in the Bus Story ($r = .656$). This correlation was significant at the .01 level (two-tailed). This shows that there is a linear relationship between MLU5 scores and plot scores, meaning that when plot score increases, MLU5 will also increase.

The absence of significant interaction effects between the two time points and the four groups might be due to a couple of reasons. Because these data were used retrospectively, there was no control over the additional training that was given to children who had received SGT four times or five or six times. It is not clear which therapies these children have received on top of the SGT sessions. This complicates the data, and might be a reason why no significant results are found between the group who has received SGT eight times and these

two groups. However, the absence of a significant interaction effect regarding the group who has not received SGT and the group who has received SGT eight times is more surprising. The children in the first group have not received any narrative intervention, but do not differ in the narrative tasks administered pre- and post-intervention compared to the 'eight times'-group. A reason for this might lie in the way therapy is being given at the Speech and Language Centre of Kentalis. Based on personal observations, it is noticeable that all therapy sessions are highly communicative, meaning that there is a high amount of interaction between the therapist and the child. Maybe, children who do not even receive narrative intervention as explicitly as with SGT, are trained in their narrative competence due to these training sessions, in which therapists always ask the child to tell about their day at school, or important events that are taking place in their lives or at school. This explanation might have some ground to it, as there were some significant main effects of time, meaning that all groups increased in some abilities pre- to post-intervention.

To avoid of the problem that is described above in future research, there needs to be additional research in a highly-controlled environment, where there is minimal pre-therapy interaction time between the therapist and the child and the focus is only on the (narrative or not) intervention that is being given. This might be cause for some problems, as children need to be able to feel 'at home' during therapy and should not be intimidated by the therapist and the intervention. The fact that there is so much interaction between the therapist and the child at Kentalis makes sure of that. Still, a more controlled environment to assess the improvement of narrative ability following SGT is highly desirable.

4.2. Attention, working memory and executive functioning

For this part of the study, the attentional, working memory and executive functioning scores have been compared pre- to post-intervention for all four groups as well. In this section, the results will be discussed. As can be read in the section where the research questions and hypotheses are given, no significant changes were expected pre- to post-intervention. This was based on literature that was read and discussed in the Introduction section on cognitive intervention methods, where long-term effects were rarely measured. Because the post-intervention measurements at Kentalis take place six months after the last therapy session, no long-term effects were expected here either. The results of auditory and visual attention, working memory and executive functioning will be discussed in this order.

4.2.1. Auditory and visual attention

For auditory attention, no differences were found pre- to post-intervention. This is in line with the hypothesis and is also in line with previous research where no long-term effects were measured for different cognitive functions (Chacko et al., 2011; Au et al., 2014; Hitchcock et al., 2017).

For visual attention, however, a significant difference was measured in the time variable, meaning that the group who received SGT four and eight times became significantly faster in identifying pairs of certain objects (as was required for this task) pre- to post-intervention. This could mean the children became more attentive and perhaps more sensitive to the different kinds of objects and were able to identify the pairs faster. The groups who received SGT five or six times and who received no SGT did not improve their reaction time. Because the group who received SGT five or six times did not improve, it cannot be said with certainty that this is due to SGT.

No additional literature has been found that also reported an improved reaction time on either attentional tasks or working memory tasks. As the result of the present study is surprising, it could be a new and promising field of research. It could be starting point for further research into improved reaction times within the clinical group of children with SLI. There would need to be a bigger sample size and scores should be compared to typically

developing children, to show possible discrepancies.

Even though no difference in visual attention task was expected, it is a positive outcome that at least a few children seemed to improve in the performance of this task. Unfortunately, these functions have not been tested for far-transfer effects, so there is no way to be certain that this improved task-related ability also transfers to untrained tasks. Any new research into this field should also include far-transfer measurements.

The additional result of there being no improvement on the accuracy variable of the visual attention task, is again in line with the hypothesis that was formed for this research question.

4.2.2.Working memory

For working memory, Block Recall scores have been compared pre- to post-intervention for the four groups. No main effects of time or group were found. This is in line with the hypothesis formed for the research question and it is partly in line with previous research. Fuentes et al. (2017) did report long-term effects of the CWNT (Cogmed Working Memory Training), but this study was conducted within a fundamentally different clinical group, namely children with symptomatic epilepsy. None of the other studies that used CWNT reported long-term effects. As for the studies using Brain Game Brian to improve working memory skills, they are also partly in line. Dovis et al. (2015) did report long-term effects, but this effect was also significant for a condition where working memory was not being trained actively, so there is no way to be sure whether this result is reliable. De Vries et al. (2015) did not find any long-term effects, thus having the same result as the present study.

In further research, it is recommended to extend the amount of tasks that are administered to measure working memory, so that a broader spectrum of all aspects of working memory can be compared within the clinical group of children with SLI. Block Recall is a task to measure the visuospatial sketchpad, but there are three more aspects of working memory that can be measured, namely the phonological loop, the central executive and the episodic buffer (as is described in the working memory model by Baddeley (2010)).

4.2.3. Executive functioning

Within the rather complex field of executive functions, two tasks have been examined in this study, which reflect four executive functions. First, the 'Trollen Tellen' task was administered for cognitive flexibility and switching, and second, the 'Zesdelentest' for organising and planning. In terms of organising and planning, no differences have been found pre- to post-intervention for any of the groups. This is in line with the hypothesis. For cognitive flexibility and switching, a significant difference was found in the accuracy variable for the group who has not received SGT. This means that post-intervention, this group could better switch between counting forward and counting backward than pre-intervention. It is unclear why this group in particular improved in this ability, since there was no control over which intervention this group has been given instead of SGT. This result was not expected given the hypothesis and the literature that has been discussed in the Introduction. However, a recent study conducted by Vugs, Knoors & Cuperus (2017) found improved cognitive flexibility in children with SLI after six weeks of executive function training, and found a maintenance effect after six months. Thus, cognitive flexibility might be an executive function that could be sensitive to improvement.

4.3. Regression analyses: attentional, working memory and executive functioning skills as possible predictors of narrative variables

In this part of the study, several cognitive variables have been taken to investigate whether they were possible predictors of specific narrative variables, namely plot score, MSA and MLU. These regression analyses have only been conducted for the group who has received SGT eight times. Within the research conducted at Kentalis, several others have tried to

establish predictive relationships between cognitive and narrative skills. These include the unpublished papers by De Wael (2017) and Broedelet (2016). The results of the present study will be compared to the results found in these two previous studies. First, the results of the regression analyses for the Bus Story will be discussed, followed by the results for the Frog Story.

For PS and MSA, the following pre-intervention cognitive variables have been selected as possible predictors: Block Recall scores, non-word repetition scores, the accuracy and time score for Trollen Tellen (cognitive flexibility and switching) and the total score of the Zesdelentest (planning and organising). For the post-intervention MSA scores, none of these variables were significant predictors. For the post-intervention PS, however, the accuracy score for Trollen Tellen turned out to explain 50% of the variance found in PS, with a significant *B* coefficient of 1.961, meaning that PS increased with 1.961 for each additional point in the accuracy of the Trollen Tellen task.

For MLU, slightly different pre-intervention cognitive variables have been selected, namely the accuracy and time scores of Trollen Tellen, the total score of the Zesdelentest, the Block Recall Scores and the auditory attention scores. The results showed that, again, the pre-intervention accuracy score for Trollen Tellen was a significant predictor of the post-intervention MLU scores in the Bus Story, with, again, a very high R^2, namely 55.8%. This means that the accuracy score for Trollen Tellen explained 55.8% of the variance found in the MLU scores. It was found that the MLU increased with .417 for each additional point the accuracy of the Trollen Tellen task.

For the Frog Story, the same cognitive variables have been entered for the same narrative variables as in the Bus Story. None of these variables turned out to be significant predictors of any of the narrative variables.

Since the same variable showed to be a significant predictor of PS and MLU, the relationship between telling a story and mastering the skill of cognitive flexibility and switching is one that needs further research. The results of the present study are telling, but for reliable and more powerful results there needs to be a bigger sample size (the present sample size was thirteen children that received SGT eight times). The percentage of variance explained by the accuracy score of Trollen Tellen for both PS and MLU was very high, which could be seen as a very positive result, but because the sample size was so small and there were some missing values, these percentages might not be completely reliable. With a bigger sample size more reliable and valid results might be gathered. However, the results of the present study do tell that there might be a strong predictive relationship between the pre-intervention cognitive variables and the post-intervention narrative variables. It has been shown in this study that the pre-intervention cognitive scores might be of influence on the narrative scores post-intervention, meaning that, at least for some children with SLI, the ability to have a flexible mind and to switch accurately has a positive influence on how many plot elements you produce and how long utterances are post-intervention. Hence, this cognitive skill might influence how well a child with SLI is able to learn from SGT and the way they perform post-intervention.

The fact that no significant results have been found for the Frog Story, does not mean that there are definitely no predictive relationships between cognitive and narrative variables, but it might just be due to the small sample size and the (although small) amount of missing values.

Now, the results of the present study will be compared to those of De Wael (2017) and Broedelet (2016). De Wael (2017) has conducted a study in which the goal was to ‘gain insight in the development of EF and narrative abilities’, and to investigate the relationship between these two. The development of EF and narrative abilities has been studied over a period of one year. At the first measurement point, T1, De Wael found a significant predictive

relationship between plot score and cognitive flexibility in the Bus Story. For cognitive flexibility, she also used the accuracy score of the Trollen Tellen task. In this light, the present study has replicated the results found by De Wael, but also extended them, by showing a positive predictive relationship of the cognitive flexibility at T1 (pre-intervention) and the plot score at T2 (post-intervention). De Wael showed that, at the same measurement time, the cognitive flexibility score could predict plot score, but the present study extended this result by showing that the cognitive flexibility score pre-intervention could predict the plot score post-intervention. An important difference between De Wael's results and the results of the present study, is that De Wael's results are based on children who are younger, namely six years old. The average age of the children in the present study was nine years old. However, it is telling that De Wael's results show that there is a relationship between cognitive flexibility and plot score at six years old, and the present study shows that this relationship is still present three years later. De Wael further showed a positive predictive relationship between auditory attention and MLU, a result which this study has not found. Subsequently, De Wael has not found a positive predictive relationship between cognitive flexibility and MLU.

Broedelet (2016) used a similar design as the present study, investigating (among other things) the predictive values of T1 executive functions on T2 narrative variables. For the Bus Story, she found that T1 non-word repetition was a significant predictor of T2 plot score, and that T1 grammaticality was a significant predictor of T2 grammaticality. For the Frog Story, she found that T1 grammaticality was a significant predictor of T2 grammaticality. No further significant predictors were found for plot score. Comparing these results to the results of the present study, they do not resemble each other. Broedelet (2016) did find a significant correlation between T1 cognitive flexibility and T2 plot score in the Bus Story, but no significant predictive value of cognitive flexibility on plot score was found. Thus, the present study did not replicate any of the results by Broedelet (2016). However, Broedelet as well as the present study did find significant predictors of cognitive skills on narrative variables, so further research is justified. There does seem to be some predictive value of cognitive functions on narrative variables.

4.4. Limitations of the present study and recommendations for further research

To conclude this Discussion section, a few limitations of the present study will be discussed, and recommendations for further research will be given. Some of these limitations and recommendations have already been discussed in the Discussion paragraphs. The points which are given here reflect more on the general experimental design. A serious limitation of the present study is that the post-intervention examination takes place six months after the last therapy session. After this session, children continue treatment with a local speech and language therapist, where they are expected to train more with the acquired strategies and techniques learned at Kentalis. It would be recommended that children are tested immediately after the last therapy session, so the effects of the intervention can be examined more directly. This brings us to the next limitation of this study, namely that it has dealt with retrospective data. The author of this study did not have any control over which tests were used and when they were tested. More research needs to be done that is prospective, controlled and with an immediate post-intervention measurement. This will also result in fewer missing values.

5. Conclusion

The research questions that were proposed for this study were as follows:

1. To what extent do the narrative skills of children with SLI from the ages of seven to twelve improve after being given SGT looking at the macro- and microstructural level?
2. To what extent will auditory and visual attention, working memory and executive functioning ability in children with SLI from seven to twelve years old improve after being given SGT?
3. To what extent can pre-intervention auditory and visual attention, working memory ability and executive ability predict the narrative scores post-intervention?

To answer the first research question, looking at the macrostructural level, which was measured by examining the amount of plot elements that were produced in the Frog Story and the Bus Story, a significant increase of plot score was found in the Bus Story pre- to post-intervention for groups who had received SGT, be it eight times or four, or five or six times. Because these last two, which functioned as control groups, also showed an increase in the amount of plot score, it cannot be said with certainty that this is due to SGT.

While looking at the microstructural level, the MLU5 scores of the Bus Story significantly increased for all groups, including the one who did not receive any SGT. The same goes for the percentage of non-fluencies in the Bus Story, all groups decreased in the percentage of non-fluencies pre- to post-intervention.

Overall, participants managed in the story retelling task to produce more plot elements, they produced longer utterances and their speech became more fluent. However, it is unclear whether this is because of SGT. Additional research with more control over the groups is needed to verify or reject the effect of SGT. Looking at the absence of significant improvements in the Frog Story Test, it seems that this task might be still too complex for chidren with SLI, since this is a story generation task. As was discussed in the Discussion section, between self-generating a story and being grammatically correct, with enough plot elements and complex sentences, a trade-off might take place. A story retelling task is less complex and might make improvements on macro- and microstructural levels more available, because the ‘generating’ part of the Frog Story task is less difficult.

In terms of the second research question, mixed results were found, which were not expected given the hypotheses. The groups who received SGT four and eight times decreased in reaction time in the task for visual attention pre- to post-intervention. Additionally, the group who has not received SGT could switch more accurately between counting forward and counting backward in the task for executive functioning. Both these results were not expected. The time between pre- and post-intervention measurements is six months. This can be seen as a maintenance or long-term effect and this was not found in the literature that was discussed in the Introduction. However, on all other tasks the participants did not become better or worse at performing.

Regarding the third and final research question, a more clear result has been found. It has been found that the pre-intervention accuracy variable of the cognitive flexibility and switching task could predict post-intervention plot scores and MLU in the Bus Story. More research to validate and replicate these results is needed, but it seems that there is a relationship between the ability to switch accurately and these narrative variables, giving more body to the debate going on about the language-specificity of SLI. The current findings support the idea that SLI might not be specific to language, by showing that cognitive flexibility skills have an influence over narrative variables in a clinical group of children.

6. Literature

Au, J., Berkowitz-Sutherland, L., Schneider, A., Schweitzer, J. B., Hessl, D., & Hagerman, R. (2014). A feasibility trial of Cogmed working memory training in fragile X syndrome. *Journal of Pediatric Genetics, 3*, 147-156. doi:10.3233/PGE-14098

Baddeley, A. (2010). Working memory. *Current Biology, 20, 4*, 136-140. doi:10.1016/j.cub.2009.12.014

Berman, R. A., & Slobin, D. I. (1994). Narrative structure. In R. A. Berman, & D. I. Slobin, *Relating events in narrative: a crosslinguistic developmental study* (pp. 39-84). Hillsdale, New Jersey / Hove, UK: Lawrence Erlbaum Associates, Publishers. doi:10.2307/415747

Bliss, L., & St. Pierre, K. (1997). Personal Narrative Structures of Children with Specific-Language Impairment. *Journal of Children's Communicative Development, 18, 2*, 11-18. doi:10.1177/152574019701800202

Boersma, A. (2015). *The efficacy of story grammar training on narrative skills of children with specific language impairment.* Utrecht / Eindhoven: Utrecht University, Master of Clinical Health Sciences and Speech and Language Centre, Royal Dutch Kentalis.

Botting, N., Psarou, P., Caplin, T., & Nevin, L. (2013). Short-Term Memory Skills in Children with Specific Language Impairment: The Effect of Verbal and Nonverbal Task Content. *Topics in Language Disorders, 33, 4*, 313-327. doi:10.1097/01.TLD.0000437940.01237.51

Broedelet, I. (2016). *Narrative abilities and executive functioning in Dutch children with Specific Language Disorder: a longitudinal study.* Amsterdam: Universiteit van Amsterdam.

Büttner, J. (2016). Neurolinguistic view into narrative processing. In N. Igl, & S. Zeman, *Perspectives on Narrativity and Narrative Perspectivation* (pp. 63-88). Amsterdam / Philadelphia: John Benjamins Publishing Company. doi:10.1075/lal.21.0int

Chacko, A., Bedard, A., Marks, D., Feirsen, N., Uderman, J., Chimiklus, A., . . . Ramon, M. (2014). A randomized clinical trial of Cogmed Working Memory Training with ADHD: a replication in a diverse sample using a control condition. *Journal of Child Psychology and Psychiatry, 55, 3*, 247-255. doi:10.1111/jcpp.12146

Colozzo, P., Gillam, R. B., Wood, M., Schnell, R. D., & Johnston, J. R. (2011). Content and Form in the Narratives of Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 54*, 1609-1627. doi:10.1044/1092-4388(2011/10-0247)

Conti-Ramsden, G. (2008). *Heterogeneity of Specific Language Impairment (SLI): Outcomes in Adolescence.* Manchester: The University Of Manchester. doi:10.1177/0265659010368750

Daamen, K., De Groot, M., & Scheper, A. (2013). *Story Grammar Training.* Eindhoven : Spraak en Taalambulatorium Kentalis.

De Vries, M., Prins, P. J., Schmand, B. A., & Geurts, H. M. (2015). Working memory and cognitive flexibility-training for children with an autism spectrum disorder: a randomized controlled trial. *Journal of Child Psychology and Psychiatry*, 566-576. doi:10.1111/jcpp.12324

De Wael, L. (2017). *EF and narrative abilities.* Amsterdam: University of Amsterdam.

Dodwell, K., & Bavin, E. (2008). Children with specific language impairment: an investigation of their narratives and memory. *International Journal of Language and Communication Disorders, 43, 2,* 201-208. doi:10.1080/13682820701366147

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. (2015). Improving Executive Function in Children with ADHD: Training Multiple Executive Functions within the Context of a Computer Game. A Randomized Double-Blind Placebo Controlled Trial. *PLoS ONE, 10, 4.* doi:10.1371/journal.pone.0121651

Duijf, A. (2014). *Story Grammar Training: effectstudie naar de vertelvaardigheid bij kinderen met een taalontwikkelingsstoornis.* Utrecht: MSc Clinical Language, Speech and Hearing Sciences.

Duinmeijer, I., De Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language and Communication DIsorders*, 542-555. doi:10.1111/j.1460-6984.2012.00164.x

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-Third Edition (PPVT-III).* Pearson.

Ellis Weismer, S. (2000). Intervention for children with developmental language delay. In D. V. Bishop, & L. B. Leonard, *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome* (pp. 157-176). Sussex: Psychology Press Ltd.

Emslie, H., Kalff, A., & Krabbendam, L. (2006). *Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C).* Amsterdam: Harcourt Assessment.

Emslie, H., Wilson, C., Burden, V., Nimmo-Smith, I., & Wilson, B. (2003). *Behavioural Assessment of the Dysexecutive Syndrome in Children (BADS-C).* Pearson.

Friend, M., & Bates, R. P. (2014). The union of narrative and executive function: different but complementary. *Frontiers in Psychology, 5, Article 469,* 1-12. doi:10.3389/fpsyg.2014.00469

Fuentes, A., & Kerr, E. N. (2017). Maintenance effects of working memory intervention (Cogmed) in children with symptomatic epilepsy. *Epilepsy & Behavior, 67,* 51-59. doi:10.1016/j.yebeh.2016.12.016

Girbau, D. (2016). The Non-word Repetition Task as a clinical marker of Specific Language Impairment in Spanish-speaking children. *First Language, 36 (1),* 30-49. doi:10.1177/0142723715626069

Hayward, D., & Schneider, P. (2000). Effectiveness of teaching story grammar knowledge to pre-school children with language impairment. An exploratory study. *Child Language Teaching and Therapy, 16, 3,* 255-284. doi:10.1177/026565900001600303

Hitchcock, C., & Westwell, M. S. (2017). A cluster-randomised, controlled trial of the impact of Cogmed Working Memory Training on both academic performance and regulation of social, emotional and behavioural challenges. *Journal of Child Psychology and Psychiatry, 58, 2,* 140-150. doi:10.1111/jcpp.12638

Kort, W., Schittekatte, M., Bosmans, M., Compaan, E., Dekker, P., Vermeir, G., & Verhaeghe, P. (2005). *Wechsler Intelligence Scale for Children-III (Dutch Version).* Pearson.

Lukács, A., & Ladányi, E. (2016). Executive Functions and the Contribution of Short-Term Memory Span in Children with Specific Language Impairment. *Neuropsychology, 30, 3,* 296-303. doi:10.1037/neu0000232

Lum, J. A., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex, 48,* 1138-1154. doi:10.1016/j.cortex.2011.06.001

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

Miller, C. A., Leonard, L. B., & Finneran, D. (2008). Grammaticality judgements in adolescents with and without language impairment. *International Journal of Language and Communication Disorders, 43, 3,* 346-360. doi:10.1080/13682820701546813

Mongomery, J. W. (1995). Sentence Comprehension in Children With Specific Language Impairment: The Role of Phonological Working Memory. *Journal of Speech and Hearing Research, 38,* 187-199. doi:10.1044/jshr.3801.187

Pauls, L. J., & Archibald, L. M. (2016). Executive Functions in Children with Specific Language Impairment: A Meta-Analysis. *Journal of Speech, Language and Hearing Research, 59,* 1074-1086. doi:10.1044/2016_JSLHR-L-15-0174

Pearce, W. M., James, D. G., & McCormack, P. F. (2010). A comparison of oral narratives in children with specific language and non-specific language impairment. *Clinical Linguistics & Phonetics, 24, 8,* 622-645. doi:10.3109/02699201003736403

Petersen, D. B., Gillam, S. L., & Gillam, R. B. (2008). Emerging Procedures in Narrative Assessment: The Index of Narrative Complexity. *Topics in Language Disorders, 28, 2,* 115-130. doi:10.1097/01.TLD.0000318933.46925.86

Petersen, D. B., Gillam, S. L., Spencer, T., & Gillam, R. B. (2010). The Effects of Literate Narrative Intervention on Children with Neurologically Based Language Impairments: An Early Stage Study. *Journal of Speech, Language, and Hearing Reseach, 53,* 961-981. doi:10.1044/1092-4388(2009/09-0001)

Pickering, S., & Gathercole, S. (2001). *Working Memory Test Battery for Children (WMTB-C).* Pearson.

Prins, P. J., Dovis, S., Ponsioen, A., Ten Brink, E., & Van der Oord, S. (2011). Does Computerized Working Memory Training with Game Elements Enhance Motivation and Training Efficacy in Children in ADHD? *Cyberpsychology, Behavior, and Social Networking, 14, 3,* 115-122. doi:10.1089/cyber.2009.0206

Prins, P., Ten Brink, E., Dovis, S., Ponsioen, A., Geurts, H., & De Vries M., e. a. (2013). 'Braingame Brian': Toward an Executive Function Training Program With Game Elements for Children with ADHD and Cognitive Control Problems. *Games Health Journal, 2,* 44-49. doi:10.1089/g4h.2013.0004

Reichenbach, K., Bastian, L., Rohrbach, S., Gross, M., & Sarrar, L. (2016). Cognitive functions in preschool children with specific language impairment. *International Journal of Pediatric Otorhinolaryngology,* 22-26. doi:10.1016/j.ijporl.2016.04.011

Rice, M. L. (2000). Grammatical symptoms of specific language impairment. In D. V. Bishop, & L. B. Leonard, *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome* (pp. 17-34). Sussex: Psychology Press Ltd.

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. doi:10.1044/1092-4388(2011/10-0263)

Rooney Moreau, M. (1991). *Methodology: Story Grammar Marker*. Retrieved from MindWing Concepts, inc.: https://mindwingconcepts.com/pages/methodology

Schittekatte, H., Groenvynk, H., Fontaine, J., & Dekker, P. (2007). Aanvullend psychometrisch onderzoek met de Test of Everyday Attention for Children (TEA-Ch): In Nederland en Vlaanderen aangepaste normen en nieuwe validiteits- en betrouwbaarheidsgegevens. Amsterdam: Harcourt.

Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL* . Amsterdam: Harcourt Assessment B.V.

Spencer, T. D., & Slocum, T. A. (2010). The Effect of a Narrative Intervention on Story Retelling and Personal Story Generation Skills of Preschoolers With Risk Factors and Narrative Language Delays. *Journal of Early Intervention, 32, 3,* 178-199. doi:10.1177/1053815110379124

Tellegen, P., & Laros, J. (2011). *Snijder-Oomen Niet-verbale intelligentietest SON-R 6-40* . Amsterdam: Hogreve Uitgevers.

Tomblin, B. J., 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*, 1245-1260. doi:10.1044/jslhr.4006.1245

Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment is not specific to language: The Procedural Deficit Hypothesis. *Cortex, 41, 3,* 399-433. doi:10.1016/S0010-9452(08)70276-4

Velner, S. (2006). *Werkgeheugen en dyslexie: Het werkgeheugen van Nederlandse dyslectische kinderen en de relatie met leesniveau, lange duurgeheugen en intelligentie.* Nijmegen: Radboud Universiteit Nijmegen.

Victorino, K. R., & Schwartz, R. G. (2015). Control of Auditory Attention in Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 58,* 1245-1257. doi:10.1044/2015_JSLHR-L-14-0181

Vugs, B., Knoors, H., & Cuperus, J. (2017). Executive function training in children with SLI: A pilot study. *Child Language Teaching and Therapy, 33, 1,* 47-66.