"Once upon a time..."

Narrative ability in Dutch children with Specific Language Impairment and younger typically developing children.

Nienke Lam – de Waal
6124690
Master Thesis General Linguistics

Supervisors:
University of Amsterdam
dr. J.E. Rispens

Koninklijke Kentalis
dr. A.R. Scheper

July 2012
Acknowledgements

Once upon a time a speech language therapist was interest in Specific Language Impairment and narrative ability and conducted a study...

That is how the narrative of this thesis began, it was a long, interesting journey, but, finally, here the thesis is. Before we go to the content of this thesis I want to thank the people who have helped me.

First of all I want to thank my supervisors Judith Rispens (University of Amsterdam) and Annette Schepers (Koninklijke Kentalis). Judith, to start with you, thanks for all the critical questions about the how and why of my research design and lots of feedback on my writings. Furthermore, I owe you many thanks for guiding me through the jungle called ‘statistics’, in which I was lost sometimes. Your expertise in and enthusiasm about scientific research in linguistics was really contagious. Annette, the combination of working with children with specific language impairment and scientific research on the same group of children that you encourage is one of the reasons that I started this study and which keeps me enthusiastic. Your great knowledge of specific language impairment and linguistics are an example for me. After finishing this thesis I am really looking forward to come for my internship to the Spraak & Taal Ambulatorium of Koninklijke Kentalis to learn about the job of clinical linguist and to work with the kind of children that I studied.

Off course the 2 primary schools in Hilversum and Kortenhoef, where the data for the typically developing children was collected, could not be forgotten here. This study was not possible at all without the teachers, parents and most of all the children. So, thanks to you all, especially to the children for their help and lovely stories. It was a pleasure to work with you!

I owe many thanks to Iris Duinmeijer and Sanne van der Kleij. Iris, for making her data for the SLI children available to me. And Sanne, who does the analyses for the interrater reliability.

Furthermore I am grateful to my colleagues of Koninklijke Kentalis Audiologisch Centrum Amsterdam for supporting me.

I want to thank my colleague students for all the nice discussions, support and laughter. I owe many thanks to Emmy for providing all my writings during this year from her feedback and off course for the nice times of ‘lekker studeren’.

More closely to home I want to thank my parents, my sister Linda and my close friends who support me again and again and who believed in me. I specially want to thank my love Daniël for his attention, love, support and off course all the liters of tea. It was special to work at home together on your dissertation and my thesis.

Last but not least I want to thank the Lord, who blesses the children by Himself, for this passion for children with severe communication problems that He gave me.
Abstract

Narrative tasks are often used to examine the language development of children with Specific Language Impairment (SLI). These tasks are complex, due to both linguistically and cognitively demands. Earlier studies show that age matched typically developing children (TD) outperform children with SLI on both content (semantic-pragmatic skills) and form (morphosyntactic skills) of narratives and that children with SLI show a dissociation between form and content skills. However, there is less information about narrative skills of children with SLI in comparison to younger, language matched TD children. Therefore, this study compares narrative ability of children with SLI (age 6 & 7) to that of younger TD children (age 4 & 5) matched on passive vocabulary. Both a story generation and a story retelling task were conducted and analyzed for content and form (morphosyntactical complexity and correctness). The results show that children with SLI in the domain of narrative content only differ from younger TD children for variety of content words during story retelling. There are no differences found for morphological complexity, but significantly more morphosyntactic errors were made in the SLI group. Thereby, the children in the SLI group show most often a dissociation in which form is relatively less developed than content, where the TD children show a balanced development. Task effects were found in favor of story retelling for content variables. The results indicate that children with SLI show a delay for both content and form, but that realizing adequate morphosyntactic sentences seem to be most vulnerable during narration.

Keywords: specific language impairment, language matched, narrative ability, story generation, story retelling, semantic-pragmatic skills, morphosyntactic skills.
1. Theoretical background

1.1. Introduction

Stories are often used as a measure for detecting language impairments and evaluation of speech and language treatment (Botting, 2002). Especially, narratives, which are defined as “extended discourse forms in which at least two different events are described such that the relationship between them (temporal, causal, contrastive or other) becomes clear” (Ninio & Snow, 1996:175), are often used. The production of narratives requires an integration of both linguistic and cognitive skills. A narrator has to produce grammatical and informative sentences and the information should be coherent and organized (Tannock et al., 1993; Andreu et al., 2011). Thereby, different from conversations, a narrator does not have the opportunity to give the floor to the other speaker and get some time to think about his next utterance (Colozzo et al., 2011). Cognitively, executive functions are involved in narration like online monitoring, planning, attention and organization (Tannock et al., 1993) and in case of a sequence of pictures it is essential that the narrator comprehends the events as depicted (Trabasso & Rodkin, 1994). These narrative abilities are found to be good precursors of literacy and academic success (Tannock et al., 1991; Kadevarek & Sulzby, 2000; De Blauw & Baker, 2009; Epstein & Phillips, 2009).

Narrative tasks are often used in studies that examine the language of children with Specific Language Impairment (SLI). SLI is a language impairment that is characterized by persistent low performance on language measures in comparison to peers, although there are no intellectual disorders, severe hearing loss, sensory deficits, frank neurological disorders or other developmental problems (Leonard, 1998; Rice et al., 2005). Because a clear etiology is absent, children with SLI are diagnosed by using exclusion criteria. This results in a quite heterogeneous group of children with variation in the domain and degree of language problems. Rice et al. (2005) describe the debate on the nature of language difficulties in SLI. Is their language acquisition only delayed, or do they show a deviant pattern in comparison to typically developing children (TD children)? Hereby a delay implicates that the language characteristics of the children with SLI are similar to those of younger typically developing children on all linguistic domains. If the language characteristics of these children with SLI are different from younger typically developing children it may be the case that their language capacity and language development is deviant.

In this study the narrative ability of children with SLI are compared to younger typically developing (TD) children in order to get more insight in the nature of SLI and in the narrative abilities of young TD children. Two narrative tasks are used to gain this insight, namely a story generation task (Frog Story, Mayer, 1969) and a story retelling task (Bus Story, Renfrew, 1997).
This first chapter will give a theoretical background about narrative development and ability in typically developing children and in children with SLI. Subsequently differences between story generation and story retelling will be discussed shortly.

1.2. Development of narratives in typically developing children

From the age of 2, children already start telling stories about personal situations. Even before this age narrative precursors, like the child’s capacity to elaborate in discussions of joint focus and fantasy play, can be found in the communication between parents and toddlers (Uccelli et al., 2005; De Blauw & Baker, 2009). These skills lay the foundation for the narratives told by the children themselves, because discussion of joint focus will introduce non-present elements like differences in time and non-present persons while fantasy play introduces plot structures, characters and settings to the children. Apparently, children who are able to elaborate in discussions and fantasy play are at a later age more capable to use these modeled elements by themselves in comparison to children who were less capable at a younger age (Uccelli et al., 2005). These last group will tell therefore less complete and well-structured narratives.

As described in the introduction a narrative is more than talking about an event or person in the non-present. In narratives a relationship between at least two events is expressed. The first narratives are personal narratives which are short and children need intensive support of parents or other adults who experienced the same event. This support consists of questions about personalities, activities, environments and the temporal and causal ordering of the events (Roelofs, 1998). From the age of four children start to tell narratives about the non-present without help of adults (Beals & Snow, 1994; Baker et al., 2000).

For a well-formed narrative both linguistic cohesion (micro level of individual and adjacent clauses) and thematic coherence (macro level of plot organization) are important (Berman & Slobin, 1994). For thematic coherence semantic-pragmatic skills are important, e.g. realization of the overall structure of the narrative, quantification of the amount of information (not too much or too less details) and taking the knowledge of the listener into account. Another element of semantics which is essential during narration is the capability to use the correct content words likes nouns and verbs to describe the events properly. Linguistic cohesion on the micro level of a narrative concentrates more on morphosyntactic features, in which it is important that the utterances are correct and complete. Additionally the ability to use complex morphosyntax is important, because for example the use of conjunctions like ‘because’, ‘therefore’ or ‘then’ enables a child to express temporality and causality during narration.
1.2.1. Content of narratives in typically developing children

In narrations the temporally ordered events and their hierarchical relations are important in the information exchange. Trabasso and colleagues (Trabasso et al., 1989; Trabasso & Rodkin 1994) developed the Causal Network Model which can describe such a hierarchical plot structure and enables scholars to quantify the hierarchical relation in narrations told by children and adults. They have based their model on the picture book ‘Frog, where are you?’ (Mayer, 1969), which is one of the narrative tasks used in the current study. This book tells the story of a boy and his dog who search for their frog which escaped when they were asleep. During their search the boy and the dog look in different places for the frog, meet several characters and have several experiences. Finally, they find the frog together with a frog family and the boy and the dog take the frog or another frog back home. This story can be structured in terms of a hierarchical goal plan with several layers, because there is a super ordinate goal (to retrieve the frog) that has to be achieved by the protagonist and this motivates subordinate goals, which are the different search attempts of the protagonist. Figure 1.1. shows a schematic view of this Causal Network Model. The arrows show the connecting relations between the elements which can be either causal relationship and enabling conditions.

![Causal Network Model](image)

**Figure 1.1. Causal network model of Trabasso, van den Broek & Suh (1989) (In: Trabasso & Rodkin, 1994:89)**

As shown in the model, a coherent narration contains several elements: Setting (S), Event (E), Internal Response (IR), Goal (G), Attempt (A) and Outcome (O). Each narrative starts with an explanation of the setting (S), which enables the first episode (between the brackets) of a narration which can contain all other elements. Thus, for the Frog Story the setting is the description of the three characters and the escape of the frog. The enabled episodes are the various search attempts of the boy and the dog. The basic structure of an episode is a goal-attempt-outcome sequence, whereby, the outcome can either be successful, unsuccessful or neutral (+, -, 0). An example of such a goal-attempt-outcome sequence is that the boy wants to find out whether his frog is in the hole of a tree (goal) and climbs in this tree to look in the hole (attempt), but did not find the frog there (unsuccessful outcome). The elements event and internal response of the protagonist can give extra information or input to the episode. Thereby all elements can cause other elements of the same type.
(thus an internal response can be caused by an internal response), but can also cause or enable elements of a different type (Trabasso et al., 1989; Trabasso & Rodkin, 1994).

Several studies used this model of Trabasso and colleagues to examine the narrative development of children (e.g. Berman & Slobin, 1994; Roelofs, 1998; Blankenstijn & Scheper, 2003; Duinmeijer, 2010). Young TD children, around the age of 3 or 4 years old, at the start of self-employed narrations, are not capable to organize their narrations in a hierarchical manner. The narrations of young TD children are characterized by isolated events together with some short strains of 2 or 3 events (Hudson & Shapiro, 1991; Berman & Slobin, 1994; Roelofs, 1998; Pearce et al., 2010). Beside this, these children have difficulties to focus on the given story and will associate to their own experience or describe persons, objects or stages which are not related to the theme of the story (Trabasso & Rodkin, 1994; Pearce et al., 2010). Trabasso & Rodkin (1994), Roelofs (1998) and Blankenstijn & Scheper (2003) found that children of 4 years old introduce the characters of a narration and describe some of the initiating events, but they have difficulties with realizing complete goal-attempt-outcome sequences and do not mention at all the super ordinate goal at the end of the story. However, they can answer questions about this goal properly (Trabasso & Rodkin, 1994). Children of 5 years old realize about half of both the initiating events and attempts and express temporal relations. They show some attention to the super ordinate goal (in around one third of the narrations), which shows a development toward a hierarchical ordered narrative. But difficulties with focusing on the main events of the narrative are still observed and children give too much side information (Trabasso & Rodkin, 1994; Roelofs, 1998; Blankenstijn & Scheper, 2003 Pearce et al., 2010). At the age of 8 or 9 children are more capable to organize their narrations in a hierarchical, coherent and causal manner (Hudson & Sapiro, 1991; Berman & Slobin, 1994; Trabasso & Rodkin, 1994; Roelofs, 1998).

Beside the realization of plot elements, the use of content words is an important semantic-pragmatic skill as well. Semantically in the sense that events, characters and objects must be labeled correctly and pragmatically in that way that variation in the use of content words causes a more lively narration. The narratives of young TD children are featured by less variation in content words in comparison to older children (Miller, 1991 in Botting, 2002). This may be inherent to the fact that the vocabulary of children still grows. Furthermore, around the age of 3 or 4 children organize their vocabulary more and more in semantic networks, whereby more differentiation in the same semantic category occurs, which may cause more variation in the use of content words (Elbers & Van Loon – Vervoorn, 2000).
1.2.2. Morphosyntax in narratives of typically developing children

During narration both content and form are important. For form morphosyntactic complexity and correctness are important. Children have to learn to organize their narratives in individual and adjacent clauses properly. The ability to produce linguistic structures such as embedded clauses to express causal relations, producing anaphoric pronouns correctly to refer clearly, is for this reason indispensible. It is clear that when the morphosyntactic ability of children develops this will cause longer and linguistically more complex narrations (Norbury & Bishop, 2003). In contrary to the development of plot structure, there are no clear steps linked with age which describe the development on morphosyntactic level in detail. Additionally, great variety in the skills of TD children is known (Norbury & Bishop, 2003). Roughly, at the start of making narratives at age 4, children will use mostly short, less complex main clauses without subordinated clauses (Berman & Slobin, 1994; Botting, 2002; Blankenstijn & Scheper, 2003). For the use of connectives both a development is seen in the type of relations which are described by the connectives as a increased variety of connectives in order to express relations between sentences more specifically (Hudson & Sapiro, 1991; Peterson & McCabe, 1991; Berman & Slobin, 1994). Children of 3 years old use more connectives that express spatial coherence than temporal sequencing and most of the markers for connectivity (like conjunctions) are used sentence initially. At the age of 5 children show more temporal sequencing but it is still restricted to a local level of organization. Children of 9 years old use more causal connectors to establish cohesion and organize their narrative on a more global level. Subordinated conjunctions and prepositional phrases emerge during the school period as well.

Berman & Slobin (1994) examined grammatical tense during story generation. For telling a narrative a dominant tense must be chosen, i.e. present or past tense. Young children (age 3 or 4) do not make such a choice but rather mix the different tenses in one narration. Subsequently, the young (English) children still make some morphological errors in past-tense forms and in releasing the third person present tense (Berman & Slobin, 1994; Botting, 2002). At the age of 5 children are somewhat more stable in choosing a tense, but Berman & Slobin (1994) found more heterogeneity than homogeneity in this group.

Blankenstijn & Schep (2003) found on a narrative task for children of 4 years old that around 40% of their utterances is ungrammatical. Less errors were found around the age of 8 as the percentage of ungrammatical utterances was halved.

1.3. Narrative ability of children with Specific Language Impairment

Although there is much research concerning narrative ability in children with SLI it is difficult to compare the results of different studies, due to a great variability in age of children and narrative tasks. Furthermore, some studies compared children with SLI with age matched TD children. Others
compared the SLI group with a group of language matched TD children, who are younger than the SLI group, but have the same language level. Therefore the results of the different studies must be treated with care. In general Kadevarek & Sulzby (2000) demonstrated that narratives in young children with SLI (2 to 4 years old) emerge later than in TD children of the same age. In this paragraph the semantic-pragmatic skills of children with SLI will be discussed first, followed by a discussion of the morphosyntactic skills.

1.3.1. The content of narratives of children with Specific Language Impairment

When children with SLI, in the age of 6 to 11 years old, are compared with age matched TD children most studies demonstrated that the children with SLI realize significantly less plot elements during story generation (Merrit & Liles, 1987; Olley, 1989; Bishop & Donlan, 2005; Duinmeijer, 2010). They seem to have specifically problems with realizing complete episodes (Merrit & Liles, 1987; Olley, 1989). Not all scholars disentangle the different types of plot elements, but Olley (1989) did and found for the SLI group more omitted initiating events, attempts, internal responses and consequences compared to age matched TD children. On the other hand, Merrit & Liles (1987) did not find significant differences at this level and Norbury & Bishop (2003) did not found differences in realizing initiation, attempt and resolution. In young children with SLI, mean age of 5;6 years, Pearce et al. (2010), found that they do not realize goal-directed narratives. Instead, they used a more “loosely formed description of characters, things, actions or events that are not goal-directed” (Pearce et al., 2010:641). For story retelling less plot elements for SLI, in the age of 6 to 11, in comparison to TD peers are found as well (Merrit & Liles, 1987; Duinmeijer, 2010), but Dodwell & Bavin (2008) did not find significant differences between children with SLI and TD children of 6 years old. This suggests that children with SLI have more problems with realizing the important components of a story, which may be a part of a cognitively problem with planning and organization as well as a semantic-pragmatic problem.

Few studies compare the narrative ability of children with SLI and younger TD children. Some studies show that the children with SLI (age 5 and 6) realize more plot elements than younger TD children and tell more often a goal-directed narrative (Dodwell & Bavin, 2008; Pearce et al, 2010). On the contrary, Olley (1989) found that children with SLI (older than 7 years) and language matched TD children do not differ in omission of initiating events, attempts, internal responses and consequences. For story retelling no differences with language matched TD children were found by Dodwell & Bavin (2008). Factors which can explain why children with SLI outperform younger TD children in organizing their narratives are maturation and life experience (Pearce et al., 2010). On the other hand Botting (2002) suggests that difficulties with working memory may explain the explicit difficulties of children with SLI with narratives, because keeping a story structure in mind during the
production of linguistic structures requires a high degree of working memory capacity. It has been demonstrated that children with SLI have limited memory skills (Ellis Weismer, 1996; Norbury & Bishop, 2002; Dodwell & Bavin, 2008) even in comparison to language matched TD children (Gathercole & Baddely, 1990).

Two studies, investigated the variety of the used content words during story generation in children with SLI. For children between 2 and 4 years old no differences in variety were found between children with SLI and TD peers (Khadevarek & Sulzby, 2000), but for children of 6 to 8 years old there was less variety of content words in the narratives of the language impaired children in comparison to age matched TD children (Fey et al., 2004). The different outcomes of both studies may be due to differences in age of the participants.

1.3.2 Morphosyntax in narratives of children with Specific Language Impairment
Several studies have demonstrated that children with SLI produce less grammatically complex utterances compared to TD peers (Gillam & Johnston, 1992; Botting, 2002; Norbury & Bishop, 2003). Their utterances are shorter as well in comparison with age matched TD children during story generation (Kadevarek & Sulzby, 2000; Bishop & Donlan, 2005). However, Van Gils (2010) did not find significant differences for morphosyntactic complexity during story generation when she compared children with SLI to age matched TD children. Contrastively, in story retelling significant differences were found for complexity whereby TD children used especially more embedding than children with SLI (Van Gils, 2010). Gillam & Johnston (1992) and Olley (1989) found that younger children matched for language ability (age 7) outperform children with SLI (age 10) on correct realization of grammatical complexity (coordination, subordination, relative clauses, complement sentences, conjunctions and cohesive devices).

For young children with SLI, age 2 to 4, it is found that they use less past tense forms during story generation than age matched TD children (Khadevarek & Sulzby, 2000). In general children with SLI make more morphosyntactic errors during story generation than TD peers like errors in the use of past tense, determiners, argument structure, word order, agreement and verb morphology (Gillam & Johnston, 1992; Botting, 2002, Norbury & Bishop, 2003; Duinmeijer, 2010; Van Gils, 2010; Verhoeven et al., 2011). For story retelling significant differences in realization of obligatory arguments, verb morphology and use of determiners were found between SLI and age matched TD children (Van Gils, 2010). Children with SLI tend to make more errors in all these categories.

The studies discussed in this paragraph compare children with SLI with age matched TD children. In this study the morphosyntax of children with SLI in comparison to language matched TD children will be examined.
1.3.3 Content versus form in narratives of children with Specific Language Impairment

Colozzo et al. (2011) state that although children with SLI show problems in both content and form that this does not automatically imply that both language domains are equally affected. This was observed earlier by Rice et al. (2005) who mention that children with SLI score similar on content measures in comparison to younger TD children, but worse on morphosyntactic measures. Colozzo et al. (2011) investigated the relationship between form and content problems in children with SLI in more detail. They elicited 2 narratives, 1 based on a single picture and 1 based on a sequence of pictures and conducted an in-depth story grammar and linguistic analysis for these narratives. A scoring framework was used in order to get a measure for content, in which for each utterance was analyzed for how much story elements it contains. For the form analysis they computed the mean number of errors per utterance. With those 2 measures a Relative Strength of Form was computed. Colozzo et al. (2011) found in the SLI group more unbalanced Relative Strength of Form than in the age matched TD group. They state that “the children with SLI are distributed more or less evenly between the low form accuracy and the low content elaboration subgroups” (Colozzo et al., 2011:1621). However, their results do show differences between these dissociations for children with SLI of about 10 to 15%, whereby Low Form – High Content occurs most. Colozzo et al (2011) suggest that the dissociations reflect limitations in processing capacity, whereby children with SLI do not succeed in equal proficiency of both content and form during a complex task as telling a narrative.

To sum up, the literature about narrative ability in SLI indicates that both content and form are less developed in these children compared to age matched TD children which may be due to both linguistic and cognitive problems. For the comparison with language matched TD children no consensus about both form and content is found.

1.4. Story generation versus story retelling

In the literature a great variety is reported on (type of) tasks used to measure narrative ability, which may tap different skills and therefore provide other information (e.g. Merrit & Liles, 1987; Pearce, 2003; Epstein & Phillips, 2009). Epstein & Phillips (2009) list some narrative elicitation tasks: single-scene pictures, wordless picture books, films with and without dialogue, story stems and conversation elicitation procedures. Another kind of task is a narrative that must be retold by a child with or without pictures (Gillam & Carlile, 1997; Norbury & Bishop, 2002; Dodwell & Bavin, 2008; Andreu et al., 2011). As stated by Hudson & Shapiro (1991: 127) “narrative skill cannot be considered as a general skill applicable to all narrative genres and elicitation contexts”, but it is not always clear which skills are needed for which kind of narrative task. Therefore it is important to make a conscious
choice for the kind of task that is used in science and in the clinical practice. In this study a story generation and a story retelling task were used, both with a wordless picture book. These type of tasks will be discussed shortly in this paragraph.

During story generation with a wordless picture book the child does not have, in contrast to story retelling, an example of the story told by an adult which it can use as a scaffold. Therefore it has to formulate the plot line and the linguistic elements as sentences and words by itself. This seems to put a higher demand on the linguistic competence of a child. Therefore story generation may give a more realistic impression of the narrative abilities of a child (Norbury & Bishop, 2003). However, story generation is cognitive demanding as well. Executive functions like working memory, and attention are involved. For example properly use of tense and locations appeals on working memory. Attention is important in keeping the outline of the narrative in mind during narration.

Story retelling, on the other hand may put especially a demand on cognitive skills, like described by Purvis & Tannock (1997:134):

“it requires attention to the incoming information, extraction of meaning and relevance, encoding of the input into memory and reconstruction from memory using effort and judgement so that the information is organized coherent and sensitive to the needs of the listener”

Beside this, it is important that a child understands the story that is told correctly in order to be able to retell the events correctly (Merrit & Liles, 1987; Hayiou-Thomas et al., 2006; Andreu et al., 2011). This is not the case for story generation, although it is essential that children interpret the visual information on the pictures correctly (Trabasso & Rodkin, 1994).

Studies in which story generation and story retelling are compared show that story retelling elicited a higher amount of information about the plot structure (Botting, 2002) and more linguistic complex utterances (Duinmeijer, 2010). Thereby both children with SLI and age matched TD children tend to make more morphosyntactic errors during story generation than during story retelling (Duinmeijer, 2010; Van Gils, 2010). However, a higher percentage of non-fluency was found as well as a lower MLU5 (Mean Length of the 5 longest Utterances) in the SLI group for story retelling in comparison to story generation (Duinmeijer, 2010; Van Gils, 2010). Correlational analyses between the narrative scores and the cognitive verbal working memory and sustained auditory attention revealed differences between the two tasks in the way the outcomes were related to the cognitive measures. For story generation the realized plot elements correlate with auditory attention and for story retelling a correlation was found between realized plot elements and verbal working memory (Duinmeijer et al., 2012). The comparison of story generation and story retelling suggest that the task effects that were found are due to a greater linguistic demand for story generation, as argued by Duinmeijer (2010). The realization of the plot structure by themselves require a great deal of effort for their linguistic capacities, beside the cognitive load of both tasks.
This study uses the Frog Story (Mayer, 1969) for story generation and the Bus Story (Renfrew, 1997) for story retelling to compare the narrative ability of children with SLI with those of language matched TD children. Both stories are reported in many scientific and clinical settings and are used in the Netherlands by speech language therapists and clinical linguists to examine the complex language skills of children. It is reported that stories whereby visual support is provided by a causal sequence of pictures, may elicit more complex narratives than when no visual scaffold is presented (Hudson & Shapiro, 1991; Epstein & Philips, 2009). The Frog Story and Bus Story are both supported with a causal sequence of pictures and therefore a comparison could be made without interfering information due to differences in visual information. The Frog Story seems to be linguistically more demanding because this story is longer (24 picture instead of 16) and it contains more protagonists and events that have to be described and related to each other than the Bus Story (Duinmeijer, 2010). Subsequently the Frog Story appeals on attention, whereby the Bus Story appeals on verbal working memory (Duinmeijer et al., 2012).

1.5. Research questions and hypotheses

In this study story generation and story retelling are compared between children with SLI and younger TD children. The 2 groups were matched on passive vocabulary. This study was conducted in order to get an answer to the following 3 research questions:

1. Does the group of children with Specific Language Impairment differ significantly from language matched, younger typically developing children on form and content of narratives in story generation and/or story retelling?

2. Is there a significant task effect between story retelling and story generation in children with SLI and language matched, younger typically developing children?

3. Is there a form-content dissociation in the narratives of children with Specific Language Impairment and language matched, younger typically developing children as described by Colozzo et al. (2011) in story generation and/or story retelling?

In the literature contradicting results are reported for the comparison of narrative ability between children with SLI and language matched TD children. Therefore it is not straightforward whether differences between SLI and the younger TD children matched on passive vocabulary can be expected in this study. In this study, matching was carried out with a lexical-semantic variable (passive vocabulary) and therefore it seems plausible that the content of the narratives between the 2 groups is comparable. But, form aspects of narratives like morphosyntactic correctness and complexity may be different, especially because the literature mentions more difficulties with
morphosyntactic than with semantic skills in SLI. Based on these arguments it is hypothesized in this study that children with SLI and TD children will not show significantly differences in the content of narratives, though significant differences are expected for the morphosyntax of narratives. Hereby it is hypothesized that the younger TD children make less morphosyntactic errors and use more complex utterances than the children with SLI.

As argued in paragraph 1.4. the Frog Story seem to be both linguistically and cognitively demanding, whereby this task seem to be linguistically more demanding than the Bus Story because of the length of the narrative and the more complex content. For TD children in the age of 4 and 5 years, as in this study, it is shown that the linguistic part of narration is still developing. On the other hand memory and attention are not fully developed as well. For children with SLI research so far shows that they have both linguistic and cognitive difficulties. Because both groups in this study may not have fully developed the executive functions yet, it is hypothesized that they will differ especially in managing linguistic demand. Therefore it is hypothesized here that the SLI children and the TD children will show a task effect in favor of story retelling.

The results of Colozzo et al. (2011) do suggest that, due to a limited processing capacity, children with SLI show form-content dissociations during narration. It is not made abundantly clear, however, that there are as much SLI children with a High Form – Low Content and a Low Form – High Content dissociation, because a higher percentage of Low Form – High Content was found in this group. Therefore, in this study it is hypothesized that children with SLI will show a form-content dissociation in which Low Form – High Content is expected to occur most, in line with literature that describe that children with SLI are especially impaired in the morphosyntactic domain. For TD children it is expected that they score, as shown by Colozzo et al. (2011) balanced for form and content, because they do not have a limited processing capacity.
2. Method

In order to answer the research questions two groups of participants were selected: the experimental group of children with Specific Language Impairment and the control group of typically developing children matched on passive vocabulary. To measure the narrative ability of the children both a story generation and a story retelling task were conducted. This chapter will discuss the selection of participants, the narrative tasks and their analysis, the statistical analysis and the interrater reliability.

2.1. Participants

2.1.1. SLI Subjects

The subjects with Specific Language Impairment were recruited at the Spraak & Taal Ambulatorium (Speech and Language Centre) of Koninklijke Kentalis (Royal Dutch Kentalis) in Eindhoven and Utrecht, the Netherlands. This centre is specialized in the examination and treatment of children with SLI who do not show improvement in the standard speech and language therapy offered by a speech language therapist and those children for whom the underlying deficit is not clear yet (Duinmeijer, 2010). During the intervention trajectory of 3 months at this centre, these children are intensively tested and treated by a multidisciplinary team of a clinical linguist, neuropsychologist, pedagogic councilor, audiologist, speech language therapist and a communicative councilor. The goal of this intervention trajectory is to get more insight in the specific problems of the children with SLI and to conduct a treatment strategy. The data of the children used in this study are collected at the start of the trajectory. The children were first selected by Duinmeijer (2010).

The children who are treated at the Spraak & Taal Ambulatorium are diagnosed with SLI. They score considerably below age norms on multiple language domains in spite of normal nonverbal intelligence and the absence of severe hearing impairments (Duinmeijer, 2010). A few children at the Spraak & Taal Ambulatorium have specific phonological problems or are diagnosed with impairments like ADHD and autism. These children were not included in this study, because it is difficult to disentangle the language features that are part of SLI and those that are due to a psychiatric classification like ADHD or autism. At the same time, those children with multiple diagnoses often show a somewhat different language profile. The children with specific phonological problems only were excluded as well, because the scope of this study is narrative ability. Due to this scope the children with SLI that were examined with both the diagnostic tools, the Bus Story and the Frog Story, were included. Additionally, as many as possible children must be tested for non-word repetition, because this measure was taken as a control measure to check for SLI. Only children with Dutch as mother tongue were selected. The selected group of children with SLI contains 12 children,
9 children of 6 years old and 3 children of 7 years old (table 2.1.). The group consists of 1 girl and 11 boys, a distribution of gender which is expected in SLI groups.

Table 2.1.: Characteristics of the SLI and TD group and results of the Mann Whitney U Test

<table>
<thead>
<tr>
<th></th>
<th>SLI (N=12)</th>
<th>TD (N=16)</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (months)</strong></td>
<td>Range</td>
<td>(s.d.)</td>
<td>Range</td>
<td>(s.d.)</td>
</tr>
<tr>
<td></td>
<td>73-93</td>
<td>(6.07)</td>
<td>51-69</td>
<td>(6.15)</td>
</tr>
<tr>
<td><strong>PPVT (raw score)</strong></td>
<td>75-85</td>
<td>(2.91)</td>
<td>59-86</td>
<td>(9.42)</td>
</tr>
<tr>
<td><strong>PPVT (Q-score)</strong></td>
<td>75-102</td>
<td>(9.37)</td>
<td>91-121</td>
<td>(8.47)</td>
</tr>
<tr>
<td><strong>Non-word repetition</strong></td>
<td>4-13</td>
<td>(3.23)</td>
<td>5-30</td>
<td>(6.58)</td>
</tr>
</tbody>
</table>

For the raw score of the PPVT a Q-score can be computed whereby a child is compared to peers. Hereby a Q-score of 100 is average, s.d. = 15.

2.1.2 Control Subjects

16 TD control subjects were matched with the SLI group based on passive vocabulary, measured with the Peabody Picture Vocabulary Test-III-NL (PPVT-III-NL; Schlichting, 2005). These children were required from 2 public primary schools in Hilversum and Kortenhoef, a small city and a village in the middle of the Netherlands. The raw scores on the PPVT-III-NL of the SLI children are between -1 and +1 standard deviation of the scores of the control subjects as shown in table 2.1. A Mann Whitney U Test showed that the children with SLI and the TD group indeed do not differ for passive vocabulary. When the raw scores are converted into standardized Q-scores, whereby children are compared to TD peers, we see that the SLI children score significantly lower than the TD children. Although the mean score of both SLI and TD children is in the average range, the SLI children score low in this range and the TD children high.

The control subjects were all TD children, which means that they have no behavioral disorders, learning disorders, hearing problems and/or language disorders. Only children who have Dutch as mother tongue were included. These characteristics where checked by their teachers, who filled in a short questionnaire. Parents gave permission to test their child during school time and use the results anonymous. The TD children were 4 and 5 years old. The group consisted of 9 boys and 7 girls. Because of matching for passive vocabulary, the SLI and TD group differ as expected significantly for age.

Non-word repetition, measured with the non-word repetition task developed by Rispens (2009), was used as a background measure to check for SLI, because this is supposed to be a clinical marker (De Bree et al, 2007). During this task each child has to repeat 40 non-words, varying in length between 2 and 5 syllables. For each child the correct repeated non-words were scored. For the SLI group only 8 of the 12 participants did this task, because these children were tested between 2007 and 2009 when the non-word repetition test was not used yet at the Spraak & Taal
Ambulatorium. A Mann Whitney U Test was conducted and shows that the SLI and TD group differ significantly for non-word repetition, whereby the younger TD children outperform the children with SLI.

2.2. Materials and analysis

The narrative ability of the children was measured with the Frog Story (story generation; Mayer, 1969) and the Bus Story (story retelling; Renfrew, 1997). These tasks and the analysis done in this study will be discussed in this paragraph.

2.2.1. Story generation

Story generation was examined with the Frog Story (‘Frog where are you?’ – Mayer, 1969). This wordless picture book consists of 24 pictures and tells the story of a boy and his dog who search for their frog which escaped when they were asleep (See paragraph 1.2.1). Due to the structure of the story this book elicits a story which contains problem solving and goal-oriented behavior. The book was introduced similarly as in Pearce (2003) and Duinmeijer (2010) to the children by the following instructions (translated into English):

“This is a story about a boy, a frog and a dog (while pointing at the first page). I want you to have a careful look at all the pictures, so that you know what the story is about. If you are done, I am going to ask you to tell the story to me alongside the pictures. But for now, just have a quiet look at the pictures.” (Duinmeijer, 2010:17)

After the child looks at all the pictures it was asked to tell the story. Thereby, the examiner could not see the pictures and the child was asked to tell carefully what was happening on the pictures. Only neutral prompts of the examiner were used. Sometimes other prompts like ‘what happens next?’ were used to encourage a child when needed.

2.2.2. Story retelling

For story retelling the Bus Story by Renfrew (1997) was used. This test is 1 of the 3 subtests of the Renfrew Language Scales, which contains a word finding test and a test which elicited sentences (action picture test) as well. Jansonius – Schultheiss & Borgers (2009) translated and developed the manual for English to Dutch and are currently working on the Dutch norm-scores.

The child has to listen to a narrative that is told by the examiner (see Appendix A for the Dutch text) while it looks to a wordless picture book of 16 pictures. The Bus Story tells about a bus who escapes when the driver tries to repair him. On his ride alone he meets several characters like a train, a policeman and a cow. Finally he falls into a lake, because he does not know how to brake. At that moment the driver finds his bus again and arranges a breakdown lorry to get the bus out.
After listening to the story, the child was asked to tell the story by itself while looking at the pictures again. During story retelling the examiner used neutral prompts to support the child. Sometimes a child was helped with turning the pages.

2.2.3 Analyses
Each narrative was transcribed and utterances were segmented according to the guidelines of STAP (Spontaneous Language Analysis Procedure, Van Den Dungen & Verbeek, 1999) in standard orthography. Hereby, coordinated clauses were treated as 2 main clauses, where subordinated clauses were counted as 1 utterance. Non-communicative words such as false starts, fillers, self-repetitions and self corrections were marked and deleted from further analysis. Utterances which were not related to the narrative (like ‘last page’ or ‘I was in a forest as well’) were excluded from further analysis in order to analyze only the content and morphosyntax of the narrative relevant utterances. The last group of utterances which are excluded from further analysis are those utterances which are unintelligible at all.

The narratives are analyzed, as described in the following subparagraphs, for a range of measures which can be divided into 3 groups: content, morphosyntactic complexity and morphosyntactic correctness.

2.2.3.1. Content variables
To measure the content, percentage of realized plot elements, was constructed as a variable. For the Frog Story the model of Blankenstijn & Scheper (2003; 2012) is used, which is based on the model of story grammar of Labov & Waletsky (1967) and the causal network model of Trabasso & Rodkin (1994; see chapter 1.2.1). The story grammar of Labov & Waletsky (1967) proposes a structure, consisting of the elements orientation, complication, evaluation, resolution and coda, which can be applied to all narratives. In comparison to Trabasso & Rodkin, Blankenstijn & Scheper added 2 extra plot elements, namely the initiating event ‘the jar is empty’ and an emotional response ‘the boy is glad’. A plot element is only scored as realized when the child reflects clearly to the person who acts, describes the activity and mentions the place where the activity took place (Blankenstijn & Scheper 2003; 2012). For the Bus Story the model of Jansonius · Schultheiss & Borgers (2009) was used. This model is also based on the story grammar of Labov and Waletsky (1967). The Bus Story, however, does not have a complex causal network in a way that there is a super ordinate goal. Rather, the Bus Story is more a chronological arranged story. Therefore the plot scores of both stories could not be compared with each other. For the Bus Story incomplete plot elements can be scored as well, but for this study only the complete plot elements were taken into account. In Appendix B the plot elements of both stories are given.
Other measures to gain more information about the content of the story are the *Type Token Ratios (TTRs)*, which were computed for nouns, verbs and conjunctions. These measures give information about the variety of content words, whereby a TTR nearby 1 means that there is much variety and a TTR nearby 0 that there is hardly any variety. To give an overview, a TTR in which verbs and nouns were combined was computed as well (*TTR Content Words*).

### 2.2.3.2. Variables of morphosyntactic complexity

Concerning morphosyntactic complexity several measures were computed. Firstly, the *number of utterances* and the *number of words* were counted in order to define the differences in elicited story lengths between the 2 tasks and the 2 groups. Secondly, the *mean length of utterances (MLU)* and the *mean length of the 5 longest utterances (MLUS)* were computed. Both measures give indirect information about the complexity of the utterances. However, a remark is needed that these measures are somewhat doubtful, because there is not always a direct relation between length and complexity. Beside this, above a certain age it seems not to distinguish between ages anymore. Therefore *embedding, compound sentences* and *direct speech* were taken into account as well. Percentages were computed for these categories whereby the number of sentences with embedding, compounds or direct speech were divided by the total number of utterances. For embedding the utterances were counted which contain a subordinator and/or the word order of a dependent clause. In the example of the Bus story told by the examiner 13 subordinated clauses were present (see Appendix A). Compound sentences are those sentences in which 2 main sentences are combined and the subject or verb of the second sentence is deleted. For the measure ‘direct speech’ the complex sentences which contain 2 subjects and 2 inflected main verbs, whereby 1 verb is a verb like to say, to shout, to think etc, were counted. Again an overview measure, *percentage of linguistic complex utterances*, was computed in which the 3 types of complex sentences were combined.

### 2.2.3.3. Variables of morphosyntactic correctness

In order to give an answer to the question if the errors of the children with SLI differ from younger TD children a variety of measures for morphosyntactic correctness was computed. The general measure is the *percentage of ungrammatical utterances*, whereby all ungrammatical utterances were counted as described in the STAP-manual (Van den Dungen & Verbeek, 1999). Important to mention here is that errors in the use of prepositions and combinations of a verb and a preposition or adverb were included as well. The *mean number of errors per utterance* was computed to make a repetition of Colozzo et al (2011) possible (see paragraph 1.3.1 & 2.5.2). For this measure all errors were counted and divided by the total number of utterances.
Subsequently the general measures for morphosyntactic correctness were separated in specific type of errors. Firstly, the percentage of gender errors was computed whereby omitted and wrong determiners were taken into account as well as wrong demonstrative pronouns. For this measure all errors were counted and divided by the sum of obligated determiners and realized demonstrative pronouns. Secondly, the percentage of past tense errors was computed, whereby all past tense errors were divided by the past tense forms that were realized. To get some information about the amount of past tense verbs used by the child the measures percentage of past tense forms and percentage of correct past tense forms were included as well, whereby the number of past tense forms (for the first measure either correct or incorrect and for the second only the correct forms) were divided by the total amount of finite verbs. For past tense it was examined as well whether children use the correct tense in contexts in which an temporal adverb obligates a certain tense. For Dutch the adverb ‘toen’ (then/at that time) requires past tense and ‘dan’ (then) requires present tense. Situations in which there is no correspondence in time between 2 main verbs in 1 utterance were taken into account as well. Thirdly, the percentage of verb agreement errors was taken into account, whereby all errors in congruency between subject and verb were divided by the total number of utterances which contain a subject and a verb. Fourthly, percentage of word order errors was computed for both stories. In Dutch the verb in main clauses must be realized in second position, eventually with a infinitive or participle in final position. The situations in which the verb was not in second position were counted and divided through the total number of utterances for the measure percentage of word order errors. Finally, the percentage of omitted obligatory arguments was used as a measure for morphosyntactic correctness. Most verbs need a number of arguments like for example ‘to give’ which requires a subject, direct object and an indirect object. It was counted how often 1 or more of these obligatory arguments were deleted. In order to get a percentage this amount was divided by the total amount of utterances which contain a main verb.

2.3. Procedure

The SLI subjects were tested by the linguists of the Spraak & Taal Ambulatorium. For most children, the 2 narratives were elicited at different moments: the Frog story was used during the intake meeting before the screening session and the Bus story was used to elicit a narrative during the screening session. However, some children did both tasks in the same session. The order of the narratives varied from child to child. The screening sessions took place in an individual situation in a quiet room, whereby the children knew that their parents watched them on a video scheme in the adjacent room.

The TD children were tested in an individual situation at their school. The 2 narratives were elicited at 2 different moments on different days or 1 in the morning and 1 in the afternoon of the
same day. All children performed firstly the Bus Story and secondly the Frog Story. After testing was done all children received a sticker as a reward for their participation.

The sessions at the Spraak & Taal Ambulatorium were videotaped and the sessions of the TD children were audio taped in order to be able to transcribe the narratives afterwards.

2.4. Interrater Reliability

The narratives were analyzed by the experimenter who knew which child belongs to the experimental or control group and who tested the TD children. To see whether the analysis is objective and conform the developed score protocol interrater reliability for the measures ‘plot score’ and ‘ungrammatical utterances’ for both story generation and story retelling was computed. For the interrater reliability check 6 of the 28 children were analyzed by a second student of general linguistics. She did an internship at the Spraak & Taal Ambulatorium and had some experience with analyzing spontaneous language and narratives. The 6 children were randomly chosen by just writing down 6 numbers under 28, whereby the second examiner did not know which child belongs to which group. Both for the plot score and ungrammatical utterances interrater reliability was checked by computing Pearson correlations (table 2.2.). Hereby a correlation of ≥ 0.800 was interpreted as acceptable.

<table>
<thead>
<tr>
<th></th>
<th>Story generation</th>
<th>Story retelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Realized plot elements</td>
<td>r=0.894</td>
<td>p=0.016</td>
</tr>
<tr>
<td>% ungrammatical utterances</td>
<td>r=0.936</td>
<td>p=0.006</td>
</tr>
<tr>
<td></td>
<td>r=0.987</td>
<td>p=0.000</td>
</tr>
<tr>
<td></td>
<td>r=0.962</td>
<td>p=0.002</td>
</tr>
</tbody>
</table>

As shown in table 2.2., the correlation coefficient between the two examiners is above r=0 and all correlations are significant. The same results are obtained when the correlations are analyzed separately for the SLI and TD group (r > 0.866). The results show that the inter-rater reliability is high and therefore further analysis can be done, whereby it is assumed that the reliability for other variables is high enough as well.

2.5 Statistical analysis

The results were checked for normality by computing the mean, median and mode as well as making histograms and appeared to be not normally distributed. Therefore non parametric tests were used for the statistical analysis. In this paragraph the analyses used for answering the 3 research questions will be discussed, starting with the group comparison followed by task effects and Relative Strength of Form.
2.5.1 Group comparison SLI and TD
Because the data turned out to be not normally distributed Mann Whitney U Tests were used for each variable to compute whether there are significant differences between the SLI and TD group. Because it is not known on forehand which effect can be expected, 2-tail testing was used. The p-level was set as lower than 0.05, but the p value was adjusted in case of multiple comparisons to 0.02. For each variable the mean and standard deviation per group were computed as well to interpret the results.

2.5.2 Task effects
To answer the question whether there are task effects for story generation in comparison to story retelling the scores on the variables for both narratives were compared. Because the data was not normally distributed, Wilcoxon Sign Ranked Tests were used to compare the related samples. Firstly, the data were pooled for the 2 groups to see whether there are overall task effects. Secondly, the data were separated for SLI and TD to measure task effects in the 2 groups separately. Again for a significant task effect p must be lower than 0.02, due to the multiple comparisons.

2.5.3 Relative Strength of Form
For repeating the study of Colozzo et al (2011) a measure of Relative Strength of Form (RSF) was computed by dividing the mean number of errors per utterance by the total number of plot elements plus the mean number of errors per utterance. For this measure rankings were used instead of raw scores, due to very different distributions for the variables (mean errors per utterance is likely to be a smaller value than number of plot elements). For the mean number of errors a ranking was made whereby the smallest value gets the highest rank and for the number of story elements it was done vice versa. This because of a high number of errors is negative, but a high number of realized plot elements is positive. The RSF outcomes were grouped in a low, balanced and high group. A RSF falling at or below 0.4 was called ‘low’, which means that form was relatively less strong than content. A RSF falling between 0.4 and 0.6 was called ‘balanced’; form and content are equal. A RSF falling at or higher than 0.6 was called ‘high’ and a child with such score seems relatively stronger in the domain of content than in the domain of form. Note that the RSF does not give information about whether the form or content are well developed or not, it is only a proportion.
With a Mann Whitney U Test is was checked whether the RSF differ significantly between the SLI and TD group.
3. Results

To give an answer to the research questions formulated in paragraph 1.5 several statistical analyses were done. In this chapter the results will be shown. Firstly, the results for group comparison of SLI and TD will be discussed followed by a comparison of the results for story generation and story retelling whereby task effects will be examined. Finally, the analysis of Colozzo et al (2011) will be repeated for the data of this study. In this chapter only the most informative data will be discussed, tables in which all data are combined can be found in appendix C.

3.1. Group comparison SLI and TD

3.1.1. Story generation

To compare the results of the SLI group with the TD group Mann Whitney U Tests were performed. In table 3.1 the results are given, together with the results per group for each variable. The significant differences are marked with grey, the 3 types of variables (content, morphosyntactic complexity and morphosyntactic correctness) are separated with double lines.

<table>
<thead>
<tr>
<th></th>
<th>SLI</th>
<th>TD</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTR Content Words</td>
<td>0.41(0.07)</td>
<td>0.46(0.11)</td>
<td>-1.466</td>
<td>0.143</td>
</tr>
<tr>
<td>% Realized plot elements</td>
<td>20.62(11.97)</td>
<td>16.78(11.88)</td>
<td>-0.635</td>
<td>0.525</td>
</tr>
<tr>
<td>MLU</td>
<td>5.92(0.92)</td>
<td>6.13(0.97)</td>
<td>-0.604</td>
<td>0.546</td>
</tr>
<tr>
<td>% Linguistic complex sentences</td>
<td>9.32(7.22)</td>
<td>10.71(11.40)</td>
<td>-0.070</td>
<td>0.944</td>
</tr>
<tr>
<td>% Ungrammatical utterances</td>
<td>54.11(15.60)</td>
<td>30.79(11.54)</td>
<td>-3.577</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean number of errors per utterance</td>
<td>0.82(0.36)</td>
<td>0.41(0.18)</td>
<td>-3.460</td>
<td>0.001</td>
</tr>
<tr>
<td>% Gender errors</td>
<td>29.75(17.33)</td>
<td>9.46(7.71)</td>
<td>-3.205</td>
<td>0.001</td>
</tr>
<tr>
<td>% Subject verb agreement errors</td>
<td>6.92(7.01)</td>
<td>1.25(1.94)</td>
<td>-2.874</td>
<td>0.004</td>
</tr>
<tr>
<td>% Total past tense errors</td>
<td>2.90(4.75)</td>
<td>11.73(8.91)</td>
<td>-2.923</td>
<td>0.003</td>
</tr>
<tr>
<td>% Omitted arguments</td>
<td>9.33(7.35)</td>
<td>3.30(3.40)</td>
<td>-2.339</td>
<td>0.019</td>
</tr>
</tbody>
</table>

For the content variables no significant differences were found. However, it is remarkable that in both groups much variation is found in realizing plot elements as shown by the remarkable high standard deviations in comparison to the means. For morphosyntactic complexity significant differences between the SLI and TD group can be found neither. But as for the content variables, great individual variation was found for percentage of linguistic complex sentences. Where the variables for morphosyntactic complexity focus on the ability to use complex and long sentences, the variables for morphosyntactical correctness focus on errors that were made by the children. These 2 types of variables can be related when children make complex sentences, but make many errors in these sentences. On the other side, children may use less complex sentences and make therefore less errors in their sentences. Of course the combination of less complex sentences and still making
errors is possible as well. The Mann Whitney U Tests show that there are differences between the children with SLI and TD children for morphosyntactic correctness. As shown in table 3.1, the children with SLI produced significantly more ungrammatical utterances than the TD children and make significantly more errors per utterance. The analysis shows as well that children with SLI make more errors with the use of gender, subject verb agreement and realizing obligatory arguments. It is striking that the TD children make more errors in the correct realization of past tense than the children with SLI. However, table 3.2 shows that the TD children use significantly more past tense forms than the children with SLI. And when the correct past tense forms are divided by all finite verbs they use significantly more correct past tense forms as well. Therefore, the significant difference between SLI and TD for past tense errors seem to be an example of using more complex forms and therefore making more errors as well.

Table 3.2: Results of the Mann Whitney U Test for past tense use during story generation

<table>
<thead>
<tr>
<th></th>
<th>SLI</th>
<th>TD</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Past tense</td>
<td>23.94 (29.34)</td>
<td>74.19 (28.61)</td>
<td>-3.236</td>
<td>0.001</td>
</tr>
<tr>
<td>% Correct past tense</td>
<td>22.71 (28.13)</td>
<td>64.23 (24.83)</td>
<td>-3.143</td>
<td>0.002</td>
</tr>
</tbody>
</table>

3.1.2. Story retelling

For Story Retelling (measured with the Bus Story) the same variables were analyzed. The results of the Mann Whitney U Tests are shown in table 3.3.

Table 3.3: Results of the Mann Whitney U Test for story retelling

<table>
<thead>
<tr>
<th></th>
<th>SLI</th>
<th>TD</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTR Content Words</td>
<td>0.60 (0.08)</td>
<td>0.67 (0.07)</td>
<td>-2.393</td>
<td>0.017</td>
</tr>
<tr>
<td>% Realized plot elements</td>
<td>24.46 (15.12)</td>
<td>25.25 (18.11)</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>MLU</td>
<td>6.00 (0.60)</td>
<td>6.40 (1.20)</td>
<td>-2.229</td>
<td>0.026</td>
</tr>
<tr>
<td>% Linguistic complex sentences</td>
<td>9.98 (8.04)</td>
<td>14.44 (9.90)</td>
<td>-1.097</td>
<td>0.273</td>
</tr>
<tr>
<td>% Ungrammatical utterances</td>
<td>52.73 (15.76)</td>
<td>30.01 (22.94)</td>
<td>-2.903</td>
<td>0.004</td>
</tr>
<tr>
<td>Mean number of errors per utterance</td>
<td>0.75 (0.40)</td>
<td>0.41 (0.33)</td>
<td>-2.579</td>
<td>0.010</td>
</tr>
<tr>
<td>% Gender errors</td>
<td>13.79 (7.87)</td>
<td>11.24 (12.37)</td>
<td>-1.241</td>
<td>0.214</td>
</tr>
<tr>
<td>% Subject verb agreement errors</td>
<td>2.76 (5.58)</td>
<td>1.18 (2.18)</td>
<td>-0.183</td>
<td>0.885</td>
</tr>
<tr>
<td>% Total past tense errors</td>
<td>3.55 (4.62)</td>
<td>4.97 (7.38)</td>
<td>-0.154</td>
<td>0.877</td>
</tr>
<tr>
<td>% Omitted arguments</td>
<td>14.26 (16.78)</td>
<td>8.48 (21.10)</td>
<td>-2.376</td>
<td>0.017</td>
</tr>
</tbody>
</table>

The content variables for story retelling show that, instead of for story generation, there is a significant difference between SLI and TD for TTR of Content Words. The TD children had higher TTR’s than the children with SLI. The percentage of realized plot elements seem to be almost the same in children with SLI and younger TD children.
For morphosyntactic complexity during story retelling the results of the children with SLI are comparable to those of the TD group, because no significant differences on these variables were found.

Children with SLI produced during story retelling overall more ungrammatical utterances than the TD children matched for passive vocabulary and made more errors per utterance. But, for the separate categories of morphological correctness only a significant difference between SLI and TD was found for omission of obligatory arguments. For story retelling no differences between the groups for past tense errors were found, however there were, just as for story generation, significant differences for the use of past tense forms. The children with SLI use less (47.88% of all finite verbs) past tense forms than the TD children (76.26% of all finite verbs) as well as less correct past tense forms (SLI: 45.56% of all finite verbs, TD: 72.68%). This illustrates that the TD children use more complex tense forms than the children with SLI.

In sum, for both story generation and story retelling children with SLI seem to differ mainly from TD children matched on passive vocabulary in morphosyntactic correctness. The morphosyntactic complexity and content seem for most of the variables comparable, with exception of the TTR of Content Words (nouns, verbs) for which significant differences were found for story retelling, but not for story generation. In the next paragraph it is explored whether there are significant differences between the two tasks.

3.2. Task effects

The second research question was about task effects: Is there a significant task effect between story retelling and story generation in children with SLI and language matched younger TD children? Task effects were tested for all measures except number of utterances and number of words. This because a task effect is evident for these productivity measures, due to differences in number of pictures (24 pictures for the Frog Story in comparison with 16 for the Bus Story). As there are more pictures involved in the Frog Story more utterances and words are elicited than in the Bus Story.

3.2.1 Task effects for the pooled data

Task effects were computed with 2-tailed Wilcoxon Sign Ranked Tests, a non-parametric T-test for related samples. As shown in table 3.4 task effects were found for both content variables, whereby the TTR of Content Words was significant higher for story retelling than for story generation and more plot elements were realized during story retelling in comparison to story generation.
Table 3.4: Results of Wilcoxon Sign Ranked Tests pooled data

<table>
<thead>
<tr>
<th></th>
<th>Story generation (s.d.)</th>
<th>Story retelling (s.d.)</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTR Content Words</td>
<td>0.44 (0.10)</td>
<td>0.64 (0.08)</td>
<td>-4.603</td>
<td>0.000</td>
</tr>
<tr>
<td>% Realized plot elements</td>
<td>18.42 (11.85)</td>
<td>24.91 (16.60)</td>
<td>-2.528</td>
<td>0.011</td>
</tr>
<tr>
<td>MLU</td>
<td>6.04 (0.93)</td>
<td>6.23 (0.99)</td>
<td>-0.524</td>
<td>0.600</td>
</tr>
<tr>
<td>% Linguistic complex sentences</td>
<td>10.11 (9.69)</td>
<td>12.53 (9.26)</td>
<td>-0.745</td>
<td>0.456</td>
</tr>
<tr>
<td>% Ungrammatical utterances</td>
<td>40.71 (17.64)</td>
<td>39.79 (22.88)</td>
<td>-1.070</td>
<td>0.284</td>
</tr>
<tr>
<td>Mean number of errors per utterance</td>
<td>0.59 (0.34)</td>
<td>0.56 (0.39)</td>
<td>-1.014</td>
<td>0.311</td>
</tr>
<tr>
<td>% Gender errors</td>
<td>18.16 (16.12)</td>
<td>12.33 (10.58)</td>
<td>-1.891</td>
<td>0.059</td>
</tr>
<tr>
<td>% Subject verb agreement errors</td>
<td>3.68 (5.50)</td>
<td>1.85 (3.99)</td>
<td>-1.381</td>
<td>0.167</td>
</tr>
<tr>
<td>% Total past tense errors</td>
<td>7.94 (8.55)</td>
<td>4.36 (6.28)</td>
<td>-2.030</td>
<td>0.042</td>
</tr>
<tr>
<td>% Omitted arguments</td>
<td>5.88 (6.14)</td>
<td>10.96 (19.25)</td>
<td>-1.521</td>
<td>0.128</td>
</tr>
</tbody>
</table>

For morphosyntactic complexity no task effects were found. However, great variety in linguistic complex sentences was found again for each task.

For the pooled data no significant differences were found for morphosyntactic correctness between story generation and story retelling. However, a significantly higher percentage of correct past tense forms was found for story retelling (table 3.5) since the past tense verb forms are included in the input of the model story.

Table 3.5: Results of Wilcoxon Sign Ranked Tests pooled data for past tense

<table>
<thead>
<tr>
<th></th>
<th>Story generation (s.d.)</th>
<th>Story retelling (s.d.)</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Past tense</td>
<td>52.66 (38.03)</td>
<td>64.10 (35.44)</td>
<td>-1.951</td>
<td>0.051</td>
</tr>
<tr>
<td>% Correct past tense</td>
<td>46.43 (33.21)</td>
<td>61.06 (34.55)</td>
<td>-2.527</td>
<td>0.012</td>
</tr>
</tbody>
</table>

3.2.2. Task effects for the SLI group

In the preceded paragraph the task effects for the pooled data were given, this paragraph focuses on the task effects for the SLI group whereby it will be discussed whether the effects for the SLI group resemble the findings for the pooled data or not.

For the SLI group can be found in table 3.6 that, as for the pooled data, the TTR for Content Words is significantly higher for story retelling than for story generation. However, the SLI group does not show a significant task effect for percentage of realized plot elements.

A task effect for morphosyntactic correctness was not found for the SLI group and therefore these results resemble the picture of the pooled data.

Children with SLI did make significantly more gender errors during story generation than during story retelling, whereby the percentage of gender errors is more than twice as high for story generation than for story retelling. This task effect of morphosyntactic correctness was not found for the pooled data. For the SLI group task effects for the use of past tense forms were computed as well and, in contrast to the pooled data, no significant task effects were found as shown in table 3.7.
### Table 3.6: Results of Wilcoxon Sign Ranked Tests for the SLI group

<table>
<thead>
<tr>
<th></th>
<th>Story generation</th>
<th>Story retelling</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(s.d.)</td>
<td>(s.d.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTR Content Words</td>
<td>0.41 (0.07)</td>
<td>0.60 (0.08)</td>
<td>-3.065</td>
<td>0.002</td>
</tr>
<tr>
<td>% Realized plot elements</td>
<td>20.62 (11.97)</td>
<td>24.46 (15.12)</td>
<td>-1.255</td>
<td>0.209</td>
</tr>
<tr>
<td>MLU</td>
<td>5.92 (0.92)</td>
<td>6.00 (0.60)</td>
<td>-0.235</td>
<td>0.814</td>
</tr>
<tr>
<td>% Linguistic complex sentences</td>
<td>9.32 (7.22)</td>
<td>9.98 (8.04)</td>
<td>-0.314</td>
<td>0.754</td>
</tr>
<tr>
<td>% Ungrammatical utterances</td>
<td>54.11 (15.60)</td>
<td>52.73 (15.76)</td>
<td>-0.667</td>
<td>0.505</td>
</tr>
<tr>
<td>Mean number of errors per utterance</td>
<td>0.82 (0.36)</td>
<td>0.75 (0.40)</td>
<td>-1.099</td>
<td>0.272</td>
</tr>
<tr>
<td>% Gender errors</td>
<td>29.75 (17.33)</td>
<td>13.79 (7.87)</td>
<td>-2.746</td>
<td>0.006</td>
</tr>
<tr>
<td>% Subject verb agreement errors</td>
<td>6.92 (7.01)</td>
<td>2.76 (5.58)</td>
<td>-1.779</td>
<td>0.075</td>
</tr>
<tr>
<td>% Total past tense errors</td>
<td>2.90 (4.75)</td>
<td>3.55 (4.62)</td>
<td>-0.254</td>
<td>0.799</td>
</tr>
<tr>
<td>% Omitted arguments</td>
<td>9.33 (7.35)</td>
<td>14.26 (16.78)</td>
<td>-1.569</td>
<td>0.117</td>
</tr>
</tbody>
</table>

### Table 3.7: Results of Wilcoxon Sign Ranked Tests for past tense for the SLI group

<table>
<thead>
<tr>
<th></th>
<th>Story generation</th>
<th>Story retelling</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(s.d.)</td>
<td>(s.d.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Past tense</td>
<td>23.94 (29.34)</td>
<td>47.88 (35.32)</td>
<td>-1.988</td>
<td>0.047</td>
</tr>
<tr>
<td>% Correct past tense</td>
<td>22.71 (28.13)</td>
<td>45.56 (33.94)</td>
<td>-1.988</td>
<td>0.047</td>
</tr>
</tbody>
</table>

3.2.3. Task effects for the TD group

Wilcoxon Sign Ranked Tests were used to compare the results for story generation and story retelling in the TD group as well. These results are shown in table 3.8.

### Table 3.8: Results of Wilcoxon Sign Ranked Tests for the TD group

<table>
<thead>
<tr>
<th></th>
<th>Story generation</th>
<th>Story retelling</th>
<th>Z-score</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(s.d.)</td>
<td>(s.d.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTR Content Words</td>
<td>0.46 (0.11)</td>
<td>0.67 (0.07)</td>
<td>-3.466</td>
<td>0.001</td>
</tr>
<tr>
<td>% Realized plot elements</td>
<td>16.78 (11.88)</td>
<td>25.25 (18.11)</td>
<td>-2.173</td>
<td>0.030</td>
</tr>
<tr>
<td>MLU</td>
<td>6.13 (0.97)</td>
<td>6.40 (1.20)</td>
<td>-0.621</td>
<td>0.535</td>
</tr>
<tr>
<td>% Linguistic complex sentences</td>
<td>10.71 (11.40)</td>
<td>14.44 (9.90)</td>
<td>-0.852</td>
<td>0.394</td>
</tr>
<tr>
<td>% Ungrammatical utterances</td>
<td>30.79 (11.54)</td>
<td>30.01 (22.94)</td>
<td>-0.724</td>
<td>0.469</td>
</tr>
<tr>
<td>Mean number of errors per utterance</td>
<td>0.41 (0.18)</td>
<td>0.41 (0.33)</td>
<td>-0.440</td>
<td>0.660</td>
</tr>
<tr>
<td>% Gender errors</td>
<td>9.46 (7.71)</td>
<td>11.24 (12.37)</td>
<td>-0.622</td>
<td>0.534</td>
</tr>
<tr>
<td>% Subject verb agreement errors</td>
<td>1.25 (1.94)</td>
<td>1.18 (2.18)</td>
<td>-0.059</td>
<td>0.953</td>
</tr>
<tr>
<td>% Total past tense errors</td>
<td>11.73 (8.91)</td>
<td>4.97 (7.38)</td>
<td>-2.130</td>
<td>0.033</td>
</tr>
<tr>
<td>% Omitted arguments</td>
<td>3.30 (3.40)</td>
<td>8.48 (21.10)</td>
<td>-0.533</td>
<td>0.594</td>
</tr>
</tbody>
</table>

Task effects for content variables of the TD group resemble those of the SLI group. For story retelling the TTR for Content Words is again higher than for story generation. For the percentage of realized plot elements a significant difference was found when a significance level of 0.05 was used, however due to the great number of related tests in this study a p-value of <0.02 is called significant. Therefore no significant task effect is found for plot elements.
For both morphosyntactic complexity and correctness no task effects were found in the TD group in contrast with the pooled data and the data of the SLI group. Neither there were differences in the use of past tense (table 3.9) during story generation and story retelling for the TD group.

Table 3.9: Results of Wilcoxon Sign Ranked Tests for past tense for the TD group

<table>
<thead>
<tr>
<th></th>
<th>Story generation (s.d.)</th>
<th>Story retelling (s.d.)</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Past tense</td>
<td>74.19 (28.61)</td>
<td>76.26 (31.26)</td>
<td>-0.909</td>
<td>0.363</td>
</tr>
<tr>
<td>% Correct past tense</td>
<td>64.23 (24.83)</td>
<td>72.68 (31.10)</td>
<td>-1.655</td>
<td>0.098</td>
</tr>
</tbody>
</table>

To summarize all the task effects discussed for the pooled data and the separate data for SLI and TD table 3.10 gives an overview, whereby the task effects were marked with grey. Here it is shown that children with SLI show task effects for gender errors in contrary to younger TD children. Task effects for the TTR of Content Words were found in both groups. Differences in realized plot elements were only found in the pooled data, but not in the separate groups.

Table 3.10: Summary of task effects found in the pooled data and the data for SLI and TD

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>SLI</th>
<th>TD</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTR Content Words</td>
<td>SG &lt; SR</td>
<td>SG &lt; SR</td>
<td>SG &lt; SR</td>
<td>SR elicited higher TTR than SG in both groups</td>
</tr>
<tr>
<td>% Realized plot elements</td>
<td>SG &lt; SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>A Higher percentage of plot elements was found for SR in the pooled data</td>
</tr>
<tr>
<td>MLU</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
<tr>
<td>% Linguistic complex sentences</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
<tr>
<td>% Ungrammatical utterances</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
<tr>
<td>Mean number of errors per utterance</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
<tr>
<td>% Gender errors</td>
<td>SG = SR</td>
<td>SG &gt; SR</td>
<td>SG = SR</td>
<td>For SG a higher percentage of gender errors was found for SLI</td>
</tr>
<tr>
<td>% Subject verb agreement errors</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
<tr>
<td>% Total past tense errors</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
<tr>
<td>% Omitted arguments</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>SG = SR</td>
<td>No differences</td>
</tr>
</tbody>
</table>

SR= story retelling, SG= story generation

3.3 Relative Strength of Form

Following Colozzo et al (2011) the Relative Strength of Form (RSF) was computed for both story generation and story retelling (see appendix C.3.). Table 3.11 and 3.12 show the results grouped for low, balanced and high RSF’s. Note that as described in the method a RSF ≤0.4 was categorized as low, a RSF between 0.4 and 0.6 as balanced and a RSF ≥0.6 as high.
Table 3.11: Results Relative Strength of Form Story Generation

<table>
<thead>
<tr>
<th>Relative Strength of Form Story Generation</th>
<th>(Low)</th>
<th>Balanced</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLI</td>
<td>9 (75%)</td>
<td>2 (17%)</td>
<td>1 (8%)</td>
</tr>
<tr>
<td>TD</td>
<td>1 (6%)</td>
<td>8 (50%)</td>
<td>7 (44%)</td>
</tr>
</tbody>
</table>

Table 3.12: Results Relative Strength of Form Story Retelling

<table>
<thead>
<tr>
<th>Relative Strength of Form Story Retelling</th>
<th>(Low)</th>
<th>Balanced</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLI</td>
<td>8 (67%)</td>
<td>4 (33%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>TD</td>
<td>2 (13%)</td>
<td>9 (56%)</td>
<td>5 (31%)</td>
</tr>
</tbody>
</table>

In table 3.11 and table 3.12 it is shown that the distribution of low, balanced and high RSF’s are not equal between the SLI and TD group. It is shown that a higher percentage of SLI children have a relatively low form and high content (RSF is low) than TD children. On the other hand TD children show more often a relatively high content and low form (RSF is high) than the children with SLI. TD children have more often a relatively balanced RSF for both story generation and story retelling. A Mann Whitney U-test was done to see whether these differences are significant (table 3.13).

Table 3.13: Results of the Mann Whitney U Test for Relative Strength of Form

<table>
<thead>
<tr>
<th></th>
<th>SLI</th>
<th>TD</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story generation</td>
<td>0.39</td>
<td>0.62</td>
<td>-2.507</td>
<td>0.012</td>
</tr>
<tr>
<td>Story retelling</td>
<td>0.33</td>
<td>0.62</td>
<td>-3.482</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Out of table 3.13 it becomes clear that there is a significant difference in RSF between SLI and TD for both story generation and story retelling. For both types of narration the children with SLI score lower than the TD children, which means that their form is relatively less strong than their content.

Colozzo et al (2011) claim that the dissociations Low Form – High Content and High Form – Low Content (scores low and high out of table 3.11 and 3.12) are more or less equally divided in the SLI population. However, the pie charts in figure 3.1 show that this is not the case in this study. Instead, in the SLI group the dissociation Low Form – High Content occurs much more often than the dissociation High Form – Low Content. It is remarkable that the TD children show an opposite pattern, whereby more children show the dissociation High Form – Low Content than the dissociation Low Form – High Content. Beside this, it becomes clear that for the SLI group form and content are less often balanced than for TD children, thus SLI children show more often a dissociation than the younger, language matched TD children.
Figure 3.1: Pie Charts for Relative Strength of Form for SLI and TD
4. Discussion & Conclusion

This study focuses on narrative ability of children with specific language impairment and younger typically developing children matched on passive vocabulary. Two different types of tasks were conducted and the narratives were analyzed for several variables which measure content, morphosyntactic complexity and morphosyntactic correctness. In this chapter the results that were found in chapter 3 will be discussed. Firstly the children with SLI will be compared with the younger TD children, whereby also Relative Strength of Form is taken into account. Secondly the 2 different tasks will be discussed. Thirdly, it will be discussed what the contribution is of this study to the investigation of narrative ability in children with SLI and young TD children. This chapter will be rounded up with some considerations about clinical implications of the results found in this study.

4.1. Children with SLI compared with younger TD children

It was hypothesized that the children with SLI and younger TD children matched on passive vocabulary will not be different for content measures, but differences were expected for form measures. Furthermore it was hypothesized that the children with SLI will show more Low Form – High Content dissociations than High Form – Low Content dissociations (Colozzo et al, 2011, see paragraph 1.5).

For the content of narratives the hypothesis was not confirmed for story retelling, because the children with SLI had a significantly lower TTR, which implicates that they varied less in the use of nouns and verbs during retelling. For SLI it is known that verbal short term memory as part of the working memory is less developed than in age matched and language matched TD children (Gathercole & Baddeley, 1990; Ellis Weismer, 1996; Norbury & Bishop, 2002; Dodwell & Bavin, 2008; Duinmeijer, 2010). This would be an explanation for a lower TTR in comparison to the TD children, because they may remember less specific content words out of the model story.

The second content variable, percentage of realized plot elements, did not show significant differences for both story generation and story retelling between the children with SLI and younger TD children. Therefore the hypothesis is confirmed for this specific variable and this study resembles the findings of Olley (1989) for story generation and of Gillam & Carlile (1997) and Dodwell & Bavin (2008) for story retelling. However, Dodwell & Bavin (2008) and Pearce et al (2010) found for story generation that the SLI children, age 5 and 6, outperform the younger TD children, age 3 and 4. The difference between the results of this study and those of Pearce et al (2010) may be explained by differences in age of the TD group, because in the study of Pearce et al the children had a mean age of 3;1 years old. At this age narrative development is just started and children are not able to tell a goal-oriented narrative yet and still need help from adults (Hudson & Shapiro, 1991; Beals & Snow,
In the study of Dodwell & Bavin (2008) the age of the children was comparable to this study, but they used another story for which is not clear whether it is as complicated as the Frog Story.

In this study both groups show a great variety in realized plot elements and the TD group did not reach the level of realized plot elements found by Roelofs (1998) and Blankenstijn & Scheper (2003) for story generation of around 40% (Roelofs, 1998) and 47% (Blankenstijn & Scheper, 2003). More detailed analyses of the results show that 3 TD children and 1 child with SLI did not realize any of the plot elements for story generation correctly, which may implicate that this task was beyond their limits of knowledge of narrative structure (Berman & Slobin, 1994). These children name pictures, but do not add time and the alternation of characters or do not relate the pictures to the overall theme of searching for the frog. However, when these 3 TD children were not included in the analysis the mean percentage of plot elements is 20.46% (s.d.= 9.47) which is far less than 41% or 47% found by Roelofs (1998) and Blankenstijn & Scheper (2003) for children of this age. This difference may be due to the limited number of participants in each study (16 TD children in this study and 30 in the study of Roelofs and Blankenstijn & Scheper), especially when it is taken into account that children of this age perform heterogeneously on narrative tasks (Berman & Slobin, 1994). On the other hand, the TD children in this study often used deictic references whereby the referent was not clear. Possibly, this can be explained by the suggestion that it was not clear enough for the children that the examiner could not see the pictures, because they were sitting at the same table. When the 4 children who did not realize plot elements correctly were not included in this study, the significant difference for TTR in story retelling is not found anymore (Appendix C, table C.2).

Morphosyntactic complexity appeared to be not different in the 2 research groups of this study, therefore the hypothesis that the groups will differ has to be rejected. However, qualitative analysis show that children with SLI use more quoted speech in both story generation and story retelling than the TD children. On the other hand, more embedding was found in the TD group for both story generation and story retelling. Compound sentences were used more by the SLI group during story generation and more during story retelling by the TD group. Although both groups are comparable for complexity of morphosyntax, children with SLI make, as hypothesized, significantly more errors than the TD children. This is in line with the results of Olley (1989) who found that the SLI and language matched TD group did not differ in number of complex sentences, but do differ in number of appropriate complex sentences. The differences in morphosyntactic correctness are for story generation somewhat more prominent than for story retelling. Especially when the children who did not realize plot elements were excluded, because significant differences were not found anymore then for mean number of errors per utterance and omission of arguments during story
retelling (Appendix C, table C.2). The combination found in the SLI group of both more ungrammatical utterances as errors per utterance for story generation, make their narratives less comprehensible. The type of errors found in this study, omission of arguments, agreement errors and gender errors, are similar to other studies on SLI in Dutch (De Jong, 1999; Orgassa & Weerman, 2008; Van Gils, 2010; Verhoeven et al, 2011).

For the young TD children it was shown that they still make errors and that only 10 to 40% of their utterances are linguistically complex utterances. This resembles the findings of Berman & Slobin (1994), Botting (2002) and Blankenstijn & Scheper (2003) that young TD children still make errors and do not use many complex sentences. Thereby this study shows more errors in past tense forms for the TD group than for the SLI group. On the other hand, they realized significantly more correct past tense forms as well.

Colozzo et al (2011) claimed that children with SLI show dissociations between content and form during narration, whereby TD children of the same age perform more balanced. In this study these findings were resembled; around 50% of the younger TD children score balanced in comparison to 17% (story generation) and 33% (story retelling) of the children with SLI. However, differences were found for the distribution between the categories ‘Low Form – High Content’ and ‘High Form – Low Content’. No equal distribution was found for the SLI group, instead, the majority belongs in the category Low Form – High Content. This may be in line with Colozzo et al, because they found differences of around 10% and 15% as well, but the differences between both dissociations for SLI in this study were for both story generation and story retelling 67%, which is much higher. The method of Colozzo et al (2011) and this study differ in respect to the scoring of plot elements. Where this study used prescribed formats of elements which have to be present in the narrative based on a story grammar and the causal network model (Labov & Waletsky, 1967, Trabasso & Rodkin, 1994), Colozzo et al just counted for each utterance whether it contains a story element like an initiating event, internal response, attempt, direct consequence or a reaction. Therefore it seems to be the case that their method does not take into account the coherent relation between events or a hierarchical organized narrative as done in this study. One may conclude out of this that Colozzo et al look specifically to semantics (what is the message of an utterance in terms of story elements) and that this study looked more to semantic-pragmatic features like coherent relations between utterances and events in the narrative. It then may be the case that children with SLI only show both dissociations in form and content when only semantics are taken into account, but not when pragmatics are taken into account as well.

For the young TD children it was found that their ability to produce plot elements (content) is relatively less developed than their ability to use the appropriate morphosyntactic utterances (form)
during narration, which may be explained by the observation that they are just at the start of narrative development but the morphosyntax is already developing a couple of years. Therefore in narratives semantic-pragmatic skills may be less sophisticated than morphosyntactic skills.

4.2. Story generation versus story retelling

As explained in chapter 1 story retelling seems to be cognitively demanding due to the appeal on verbal short memory (as part of working memory) and auditory attention. Story generation seems to be, beside cognitively demanding, more linguistically demanding, because of the length of the story, number of events and the hierarchical causal structure. For both groups it was argued that they will have explicitly difficulties with linguistic demanding tasks, and therefore task effects in favor of story retelling were expected. In this study not only task effects were computed for the separated groups, but for the pooled data as well in order to see whether the task effects are only due to differences in groups (SLI and TD) or also to the type of task.

The hypothesis for the group of SLI children, that they will show an advantage in favor of story retelling, is confirmed for the content variables, because task effects were found in this domain. For both the SLI and the TD group story retelling yielded higher TTR’s which reflected that more variety was shown in story retelling than story generation for both verbs, nouns and conjunctions (appendix C.2). An explanation for the differences in variation of verbs and conjunctions can be that for the Bus Story (story retelling) the example given by the examiner has a TTR for verbs of 0.78 and for conjunctions of 0.38. The children may benefit from this example and use more varied verbs and conjunctions themselves. However, for nouns the TTR of the Bus Story told by the examiner is 0.43, which is much lower than the mean TTR for nouns in this study (pooled data 0.72, SLI 0.67, TD 0.76). This may be explained by the observations that the children in this study do not reintroduce the bus often in contrast with the example, which makes that they vary more in the use of nouns. Beside this, it seems possible that the Frog Story yields less variation in nouns because the main characters must be reintroduced more often than in de Bus Story, due to description of different actions of the different main characters.

The realization of plot elements showed a task effect in the pooled data only. However, when the children who did not realize plot elements during story generation were excluded no significant task effect was found. Therefore it seems to be the case that both children with SLI and the younger TD children still have more difficulties with the realization of plot elements in both stories as older TD children have. However, the plot structure of the Frog Story seem to be more difficult to realize during narration for both SLI and TD children, because they have to structure it by themselves instead of having a model as for the Bus Story. Thereby the Frog Story contains more protagonists.
(the boy, the dog and the frog), which actions must be described instead of one protagonist (the bus) in the Bus Story.

For morphosyntactic complexity no task effects were found for the combined measure of linguistic complex sentences, but when separated into different categories of complex sentences task effects were found (Appendix C.2.). A task effect in favor for story generation was found for compound sentences in the pooled data and SLI group and in favor for story retelling a task effect was found for embedding in the pooled data. The task effect for embedding could mean that the children with SLI and the TD children benefit from the embedded clauses that were modeled in the example of the examiner (appendix A). For compound sentences an explanation for more of these sentences in the Frog Story could be that at most pictures both the boy and the dog are involved in different events, which may elicited more compound sentences. On the other hand the percentage of compound sentences is still very small for story generation (pooled data 2.55% and SLI 2.82%) and much variation is found, therefore it is doubtable whether this gives a reliable picture of the abilities of the children.

In the domain of morphosyntactic correctness the only task effect was found in the SLI group for gender errors. This type of errors occurred more during story generation than during story retelling. This seem to be a specific effect for the Frog Story, because this story has much more neuter nouns which requires in Dutch the marked determiner ‘het’ instead of the more common determiner ‘de’, especially because the children use often diminutives which are assigned neuter gender (little frog, little boy, little dog, little hole etc). The content nouns (bus, busdriver, train, cow, police officer etc) of the Bus Story are assigned common gender much more often (Duinmeijer, 2010; Van Gils, 2010).

4.3. Contributions of this study to the investigation of narrative ability in children with SLI and young TD children

Duinmeijer (2010) and Van Gils (2010) established that children with SLI perform less on content (realized plot elements) and form (morphosyntactic correctness for both story generation and story retelling and embedding for story retelling) in comparison to age matched TD children. This study adds that these children with SLI do not differ for content (with exception of TTR in story retelling) and morphosyntactic complexity from language matched TD children, but that they still make more morphosyntactic errors. Several scholars have proposed a processing deficit as underlying cause of SLI (e.g. Joanisse & Seindeberg, 1998; Chiat, 2001; Jakubowicz, 2003; Marinis, 2011). Jakubowicz (2003) and Marinis (2011) present, for example, the Computational Complexity Hypothesis (CCH), which claims “that SLI is caused by difficulties with accessing and integration of information at the interfaces and to relate language to other cognitive systems” (Marinis, 2011: 467). Narratives seem
to be pre-eminently the discourse form in which integration of information, both cognitive and linguistic, is essential. Therefore it seems plausible that the difficulties of children with SLI with narratives can be explained by such a processing deficit as supposed as well by scholars who studied narratives in SLI (Gillam & Carlile, 1997; Dodwell & Bavin, 2008; Andreu et al, 2011; Colozzo et al, 2011). Colozzo et al (2011) suggest that the dissociations in form and content can be explained by the great demand on processing skills which cause that form and content could not be both meet the criteria for a good narrative. The results of this study adds that especially realizing adequate morphosyntactic sentences seem to be vulnerable during narration in SLI. This may implicate that a processing deficit results in a delayed development for morphosyntactic complexity and content and a more delayed or even deviant development for morphosyntactic correctness.

For TD children no processing deficit is expected, which explains the relatively absence of dissociations in form and content in the young TD children in this study. For this group, who are at the start of narrative development, form and content seem to be equally difficult as shown by less realized plot elements, less complex sentences and more morphosyntactic errors in comparison to older children (Roelofs, 1998; Blankenstijn & Scheper, 2003; Duinmeijer, 2010; Van Gils, 2010).

Furthermore, the results of this study adds that story generation is more demanding for children with SLI in comparison to story retelling, because story generation does appeal more on the linguistic skills of the children besides the cognitive appeal of both story retelling and story generation. However, not all task effects found by Duinmeijer (2010) are resembled in this study, which may be due to different ages; the children of Duinmeijer (2010) were between 6 and 9 years old and in this study only the 6 and 7 year olds are taken into account. This study adds for the children with SLI of 6 and 7 years and younger TD children of 4 and 5 years old, that story generation is especially more demanding than story retelling for content. The higher TTRs and more plot elements during story retelling suggest that both groups benefit from the model story. This may implicate that the children with SLI are able, unless their difficulties with executive functions, to pay attention to the information in a narrative told by an adult and that they store at least some of the elements in their memory. It could be the case that the visual scaffold of pictures is a facilitating factor. For morphosyntactic complexity and correctness no task effects were found, which may implicate that for young TD children and children with SLI cognitive and linguistic demand is equally difficult in the case of narratives elicited with wordless picture books.

4.4. Clinical implications
The results of this study indicate that narratives are a useful tool for the diagnosis of children with SLI in Dutch as they show significant group differences for morphosyntactic correctness between SLI and younger TD children and as shown by Duinmeijer et al (2010) for both content and form when the
children with SLI are compared with age matched TD children. Because children with SLI and younger TD children differ especially in the domain morphosyntactic accuracy, it could be helpful for clinical linguists and speech language therapists to realize that more morphosyntactic errors in comparison to language matched TD children may indicate that a child has SLI instead of a more general language delay. Furthermore, the distinction between content and form may give clinicians information about the relative strength of a child during narration and especially a dissociation of Low Form – High Content may suggest that a given child has SLI. Thereby this distinction may give more specific information to conduct a treatment strategy.

The results of this study subscribe, together with the results of Duinmeijer (2010), the claim of Hudson & Shapiro (1991:127) that “narrative skill cannot be considered as a general skill applicable to all narrative genres and elicitation contexts”. Different tasks seem to tap different skills and therefore it is important that a narrative task is chosen consciously, for example based on the choice whether a cognitive or linguistic demanding task give the information needed for a particular child. Thereby, the narratives of both SLI and TD children reflect the information in the narrative or given by the examiner. For example a story like the Frog Story elicited more neuter nouns due to the participants of the story and therefore more errors could be made as well. And with the example of embedded clauses and past tense forms in a model story, more embedded clauses and past tense forms may be elicited as well in a narrative of a child. These differences are important to realize and must be described in studies to make comparisons between studies more accurate.

This paragraph is rounded up with the statement of Merrit & Liles (1987), confirmed by the results of this study and the study of Duinmeijer (2010), that examining both story generation and story retelling give a more complete picture of the narrative abilities of children and may give clinicians more insight in the particular problems of a child and the treatment that is needed.
5. Suggestions for future research

As it is important for narratives as a clinical tool to make a conscious decision about which narrative task will be used, it is important as well for scientific uses of narrative tasks. In order to make comparisons between studies possible, scholars have to document and argue why they use a specific narratives task.

In this study children with SLI were matched with typically developing children based on passive vocabulary. This is a measure for lexical-semantic skills which may cause that less differences between SLI and TD for narrative content were found. Therefore it would be interesting to conduct a study in which children with SLI are matched for form variables, like mean length of utterances or linguistic complexity in spontaneous language, in order to see whether this will elicited other differences between SLI and TD. Most ideal it would be to match children both on content and form variables, with for example a standardized language test as the CELF-4-NL (Kort et al, 2008). Beside this, especially the TD children in this study were quite young and therefore at the start of their narrative development. It would be useful to included older TD children in such a study (and therefore older SLI as well when matching on language ability) to see whether the differences still exceed in such a population. It could be the case that the TD children of this study were not capable yet to tell a narrative (as shown by 3 0-scores for plot elements during story generation) and that this results in differences which are not due to differences in speech and language development.

For another study it would be necessary as well to make it more explicitly clear to the children that the examiner does not know the story which is generated and could not see the pictures, in order to avoid to get a wrong impression of the child’s ability to use anaphoric in a story. If such a story is conducted this may give an answer on the question why the TD children of this study scores much lower for realized plot elements than children of the same age in other Dutch studies (Roelofs, 1998; Blankenstijn & Scheper, 2003)

In this study it was suggested that narratives are pre-eminently a difficult task for children with SLI, due to the great demand on processing and integration of information. To verify this suggestion it is useful to conduct a study in which content and form for different tasks, like a narrative, spontaneous language but also tasks that look more specifically to one language domain (like subtests of a standardized language test), will be compared.

For children with SLI a task effect was found for gender errors, whereby story generation elicited more gender errors than story retelling. As already argued, this may due to the type of nouns in the Frog Story. Therefore it would be useful to separate in a next study the words which get assigned common gender and neuter gender and then examine whether gender differences still exist.
References


