KU LEUVEN

Citation	Ellen Vandewalle, Bart Boets, Tinne Boons, Pol Ghesquière, Inge Zink, (2012), **Oral language and narrative skills in children with specific language impairment with and without literacy delay: a three-year longitudinal study,** Research in Developmental Disabilities, 33(6),1857-1870
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	http://dx.doi.org/10.1016/j.ridd.2012.05.004
Journal homepage	http://www.journals.elsevier.com/research-in-developmental-disabilities/
Author contact	vdwalle.ellen@gmail.com *+ 32 (0)16 330479*
IR	https://lirias.kuleuven.be/handle/123456789/346750

(article begins on next page)

Oral language and narrative skills in children with specific language impairment with and without literacy delay: a three-year longitudinal study

Ellen Vandewalle[a], *Bart Boets*[b,c], *Tinne Boons*[a,d], *Pol Ghesquière*[b], *Inge Zink*[a,e]

[a] ExpORL, Department of Neurosciences, Faculty of Medicine, KU Leuven, Belgium
[b] Parenting and Special Education Research, Faculty of Psychology and Educational Sciences, KU Leuven, Belgium.
[c] Child and Adolescent Psychiatry, Department of Neurosciences, Faculty of Medicine, KU Leuven, Belgium.
[d] Institute of Allied Health Sciences, Fontys University of Applied Sciences, Eindhoven, the Netherlands.
[e] Department of ENT, Head & Neck Surgery, MUCLA, University Hospitals Leuven, Belgium.

Published version: http://dx.doi.org/10.1016/j.ridd.2012.05.004
Journal homepage: http://www.journals.elsevier.com/research-in-developmental-disabilities/

KU LEUVEN

Abstract

This longitudinal study compared the development of oral language and more specifically narrative skills (storytelling and story retelling) in children with specific language impairment (SLI) with and without literacy delay. Therefore, 18 children with SLI and 18 matched controls were followed from the last year of kindergarten (mean age = 5 years 5 months) until the beginning of grade 3 (mean age = 8 years 1 month). Oral language tests measuring vocabulary, morphology, sentence and text comprehension and narrative skills were administered yearly. Based on first and third grade reading and spelling achievement, both groups were divided into a group with and a group without literacy problems. Results showed that the children with SLI and literacy delay had persistent oral language problems across all assessed language domains. The children with SLI and normal literacy skills scored also persistently low on vocabulary, morphology and story retelling skills. Only on listening comprehension and storytelling, they evolved towards the level of the control group. In conclusion, oral language skills in children with SLI and normal literacy skills remained in general poor, despite their intact literacy development during the first years of literacy instruction. Only for listening comprehension and storytelling, they improved, probably as a result of more print exposure.

Key words

specific language impairment; literacy delay; narrative skills; oral language; reading comprehension

1 Introduction

While it is well established that there is a considerable comorbidity between specific language impairment (SLI) and dyslexia (e.g., Catts, Adlof, Hogan, & Weismer, 2005; McArthur, Hogben, Edwards, Heath, & Mengler, 2000; Snowling, Bishop, & Stothard, 2000), the underlying causes for this comorbidity remain unclear. Phonological impairments are widely studied as a potential cause for dyslexia and its comorbidity with SLI. In contrast, other language skills such as vocabulary, morphology, syntax and narrative skills, and their role in the overlap between SLI and dyslexia have been studied less frequently.

1.1 Narrative skills

Telling a story involves a number of higher-level language and cognitive skills. All language components come together to form a cohesive, well-formulated, meaningful story (Seiger-Gardner, 2009). The analysis of narratives provides information about grammatical skills and the ability of children to formulate sentences, to use cohesive devices relating meanings across sentences, and to organize the story content in a meaningful way.

It has been observed that typically developing children are generally able to comprehend and retell stories by the age of six years (Merritt & Liles, 1987). By the age of seven, stories with multiple episodes begin to emerge, and by age nine or ten children use considerable detail and complete their episodes (for an overview, see Crais & Lorch, 1994). As children mature, the number of complete episodes consisting of initiating events, attempts and consequences increases with age (Muñoz, Gillam, Peña, & Gulley-Faehnle, 2003).

There are different elicitation methods to evaluate narrative skills (for a review, see Liles, 1993). Two frequently used methods are story generation tasks (e.g., Frog, where are you? Mayer, 1969), in which children are instructed to compose stories from sequencing cards or wordless books, and story retelling tasks (e.g., the Bus Story Test; Renfrew, 1997), in which children listen to stories narrated by the experimenter and are asked to retell the stories back to the experimenter (Seiger-Gardner, 2009). Story retelling tasks result in longer stories with more story grammar components and complete episode structures than story generation tasks (Merritt & Liles, 1989; Westerveld & Gillon, 2010). Story retelling is more clinically useful with older children (9 to 11 years old) for the assessment of the global organization of content than story generation. In contrast, a story generation task allows the speaker to access a larger range of structural variation and content. It is also more exigent than retelling and generally offers a better indication of vocabulary ability or potential word-finding difficulties. Additionally, it is more representative of spontaneous communication, reflects more accurately the pragmatic characteristics of the narrative and the subject is less influenced by the context because no story was told in advance to the child (Liles, 1993). However, Merritt and Liles (1989) concluded that story retelling and story generation are both effective measures of narrative ability that both activate a cognitive organization consistent with story schema. McCabe, Bliss, Barra and Bennett (2008) compared a fictional story generation task based on pictures (Frog, where are you?) with personal narratives without pictures in 7 years to 9 years 9 months old children with language impairment. They observed a significant longer story in the fictional generation task than in the personal narrative. However, more personal narratives met minimal narrative criteria. In the fictional generation task, the children often treated the pictures in isolation from each other rather than a series of events. Hence, the use of pictures might have a negative influence on narrative criteria.

Assessment of narrative skills offers a rich source of information about the higher-level language abilities of young children (Paul & Smith, 1993). Liles, Duffy, Merritt and Purcell (1995) demonstrated that a large range of narrative production variables could be reliably represented by two dominant factors: a factor measuring global organization of content (i.e., episode structure), currently assigned as *macrostructure*, and a factor measuring within- and across-sentence structure (i.e., grammatical sentence structure, within subordinate clause productivity, and textual cohesion), also called *microstructure*. In contrast, a study of Westerveld and Gillon (2010) demonstrated that microstructure measures of oral narrative performance did not represent a single construct. Mean length of utterance in morphemes and grammatical accuracy loaded on different factors. Therefore, it is important to analyze oral narratives on a range of measures.

1.2 Oral language and narrative skills in children with SLI

It is well known that children with SLI show oral language problems across many different language domains (lexical and semantic, morphological, syntactic, phonological, and pragmatic deficits) that can be measured through standardized oral language tests (Schwartz, 2009). As a result, children with SLI may have difficulties with the linguistic abilities necessary to narrate stories.

A number of *cross-sectional studies* demonstrated a poorer *macrostructure* in the stories of children with (a history of) SLI compared to typically developing children. For instance, they obtain a lower information score (4 years) (Paul & Smith, 1993), fewer complete story episodes and more incomplete episodes (9 years to 11 years 4 months) (Merritt & Liles, 1987), a lower narrative stage (6 to 7 years) (Paul, Hernandez, Taylor, & Johnson, 1996) and a lower macrostructure composite score (8 to 9 years) (Manhardt & Rescorla, 2002). Soodla and Kikas (2010) observed that SLI and control children between 6 and 8 years were equally capable to produce structurally complete stories (no group difference for the presence of the story grammar components such as setting, initiating events, internal response), but the SLI group scored significantly lower for the quantity of story information units (i.e., the amount of relevant information included in the story) in comparison to typically developing children. Also the *microstructure* of stories is poorer in children with SLI in comparison with typically developing children, as evidenced by fewer sentences (7 years 6 months to 10 years 6 months) (Liles, 1985), fewer main and subordinate clauses (9 years to 11 years 4 months) (Merritt & Liles, 1987), fewer complete cohesive ties (4 years, 6 years and 7 years 6 months to 10 years 6 months) (Liles, 1985; Paul et al., 1996; Paul & Smith, 1993), fewer complex sentences and more tense errors (6 to 10 years) (Norbury & Bishop, 2003), a lower mean length of utterances (4 to 12 years) (Kit-Sum To, Stokes, Cheung, & T'sou, 2010; Paul & Smith, 1993), a lower syntax composite score (8 and 9 years) (Manhardt & Rescorla, 2002), referential problems and the use of less sophisticated vocabulary (4 years 10 months to 12 years 1 month) (Kit-Sum To et al., 2010), and a smaller number of different word roots (4 to 6 years) (Paul et al., 1996; Paul & Smith, 1993).

Although several studies observed group differences on both macro- and microstructure qualities while comparing children with SLI and typically developing controls, Norbury and Bishop (2003) observed no group differences for the macrostructure of stories of children between 6 and 10 years old performing a telling task. In this regard, Liles et al. (1995) observed that the macrostructure factor did not distinguish children with SLI from typically developing children after microstructure was accounted for. In contrast, Manhardt and Rescorla (2002) demonstrated that late talking children did show weaknesses in macrostructure skills independent of the variance accounted for by their weaker general language abilities. It is still a matter of debate whether the macrostructure ability is independently disabled in children with SLI. In contrast, microstructure abilities seem to be more systematically disabled in children with SLI and the

constituting variables are more effective in distinguishing children with and without SLI. Kit-Sum To et al. (2010) demonstrated that syntactic complexity, measured by the mean length of utterances, was the single best predictor to differentiate between children with and without SLI.

The observed problems with macrostructure narrative ability in children with SLI do not result from difficulties perceiving the characters' mental state (Befi-Lopes, Bento, & Perissinoto, 2008; Norbury & Bishop, 2003) or from a poorer understanding of the factual details of the retold stories (Merritt & Liles, 1987), because children with SLI did not differ from typically developing children on these aspects. However, they did differ in their comprehension of relationships linking the critical parts of the stories together (Merritt & Liles, 1987).

Several studies observed deficits on narrative abilities in children with SLI on both story generation and story retelling tasks (e.g., Befi-Lopes et al., 2008). Merritt and Liles (1989) mention a strong overall homogeneity between story generation and story retelling, although for some aspects only group differences in a retelling and not in a generation task were observed and vice versa. For example in this study of Merritt and Liles, the children with SLI produced a similar number of main clauses as children with normal language in the story generation task, but fewer main clauses when retelling a story.

Moreover, there are inconsistencies across studies concerning the number of utterances, the number of words (for an overview, see Crais & Lorch, 1994) and the mean length of utterances over the ages in children with SLI. In early preschool years, children with SLI have a lower mean length of utterances compared to typically developing children (e.g., Paul & Smith, 1993). However, from the late preschool years onwards, results of different studies are contradictory. Paul et al. (1996) observed no group differences at 6, 7 and 8 years, while Kit-Sum To et al. (2010) did observe significant group differences in children between 4 years 10 months to 12 years 1 month.

Crais and Lorch (1994) suggested that these conflicting findings are caused by differences in the type of task applied, in motivation of the children, in subject selection and in method of transcription and analysis (e.g., whether or not repetitions were included in the analyses). In general, it can be concluded that certainly the narrative microstructure is poorer in stories of children with SLI as compared to typically developing children. In addition, the narrative macrostructure may be impaired too, probably as a secondary effect because of problems with the microstructure or with the comprehension of relationships in the stories.

Only few *longitudinal studies* investigated the development of narrative abilities in children with SLI. In a study of Paul et al. (1996), children with language impairment and controls performed a narrative task yearly from kindergarten to grade 2. At each time point, the children were repartitioned into three groups: a group with a past history of expressive language delay, a group with persistent expressive language delay and a typically developing group. Results demonstrated that in kindergarten, typically developing children scored significantly higher than both language delayed groups on lexical diversity and narrative stage and they had a higher proportion of complete cohesive ties than the children with persistent expressive language delay. There were no significant differences for the information score and the mean length of utterances. Group differences for macro- and microstructure narrative scores were more pronounced in kindergarten than in grade 2. The authors hence stated that problems with narrative skills tend to disappear in children with a history of expressive language delay. However, the narrative deficits disappear at a slower rate than the expressive syntactic deficits. At earlier ages, children with a history of language impairment have a deficit in many oral language aspects. When children mature, fewer oral language aspects are impaired.

In another longitudinal study (Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004) different conclusions were drawn. Children were divided into four groups based on kindergarten language and intelligence

scores, that is a typically developing group, an SLI group, a group with nonspecific language impairment and a group with low nonverbal IQ. The children performed a narrative task in grade 2 and 4. The stories of the typically developing group had higher micro- and macrostructure narrative scores (more different words, more grammatical complexity, fewer errors and more overall quality) than both language impaired groups at either grade. Moreover, children with language impairment (specific and nonspecific) were divided by the end of grade 2 in a group with and without language impairment at that moment measured using a standardized language test. The group without language impairment in grade 2 also appeared to have recovered in storytelling abilities at that point comparable to the study of Paul et al. (1996). However, by the fourth grade, the study of Fey et al. demonstrated that the stories of these children were less like the typically developing children and more like those of children with persistent language impairment. Therefore, the authors argued that the narrative deficits do not disappear but rather increase from the second to the fourth grade in children with a history of language impairment. Moreover, they warned that children with typical language had reached a ceiling in grade 2 on the measure of episode structure in the study of Paul et al.

In a study of Finestack, Fey and Catts (2006), pronominal reference skills were analyzed in these same four groups from grade 2 to grade 4. The children with SLI could not be differentiated from typically developing children based on the rate of pronominal referencing. Moreover, this rate of pronominal referencing did not significantly increase from grade 2 to grade 4 in children with SLI, which was comparable to typically developing children. However, in a similar study of Strong and Shaver (1991) concerning all types of cohesion, the opposite was observed for both observations.

To summarize, narrative tasks are more sensitive to persistent oral language problems than other standardized oral language tests. They have a good sensitivity to detect oral language difficulties (Kit-Sum To et al., 2010; Pankratz, Plante, Vance, & Insalaco, 2007) and can therefore play a significant role in the evaluation of children with developmental language impairment (Fey et al., 2004). They may also constitute a reliable indicator of future language and literacy performance in children with SLI (Bishop & Edmundson, 1987; Pankratz et al., 2007; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). Yet, the specificity of these narrative tasks seems to be rather poor (Kit-Sum To et al., 2010; Pankratz et al., 2007). Even typically developing children show a wide variation in narrative performance (measured through a telling task) between 6 and 10 years (Norbury & Bishop, 2003). A possible solution is the combination of a narrative task with other language measures to identify language disorders in school-age children (Kit-Sum To et al., 2010). Redmond, Thompson and Goldstein (2011) observed a high diagnostic accuracy in discriminating children with SLI from typically developing children when a combined measure comprising story understanding, macrostructure and microstructure narrative ability was used.

1.3 Links with reading comprehension and accuracy

Reading comprehension is a complex cognitive linguistic process that involves multiple levels of processing, particularly when texts become longer. All components of language (phonology, morphology, syntax, semantics, discourse structure, and pragmatics) are involved, as well as attention, executive function, and memory skills (Hook & Haynes, 2009). Different studies have shown that oral language (receptive vocabulary, listening comprehension, expressive vocabulary, text comprehension and semantic skills) at the age of eight predict future reading comprehension in typically developing children (Nation & Snowling, 2004; Spear-Swerling, 2006). In children with SLI, sentence comprehension and the information score of the Bus Story Test at seven years were good predictors of reading comprehension at 11 years

(Botting, Simkin, & Conti-Ramsden, 2006). Also in children with reading disabilities, the predictive value of storytelling for reading comprehension was frequently demonstrated (Feagans & Appelbaum, 1986; Gilmore, Klecan-Aker, & Owen, 1999). In general, narrative skills might be more important predictors for reading comprehension than traditional language skills such as vocabulary and syntax (Feagans & Appelbaum, 1986).

Surprisingly, in some studies, not only reading comprehension, but also word reading accuracy could be predicted by nonphonological oral language skills. Nation and Snowling (2004) showed that expressive vocabulary, text comprehension and semantic skills at eight years predicted reading accuracy at the age of 13 in typically developing children. In a similar vein, Botting et al. (2006) showed that sentence comprehension at seven years uniquely predicted reading accuracy at the age of 11 in children with SLI, even after controlling for early single word reading level at 7 years. Also Westerveld, Gillon and Moran (2008) demonstrated that children with problems on word reading, listening comprehension and reading comprehension had a lower performance on narrative comprehension, narrative macrostructure, grammatical complexity and semantic diversity in comparison to typically developing children. Compared to reading-matched controls, they had significantly lower scores on narrative comprehension, suggesting a specific oral narrative comprehension deficit. In a study of Gillam and Carlile (1997), children with SLI and control children matched on the basis of single-word reading had to read and retell stories one grade level above their reading level. Remarkably, results showed that story retellings in the two groups were similar in terms of percentages of recalled vocabulary, story elements, and problem-resolution pairs, suggesting that the level of reading is associated with the narrative level. However, holistic analysis showed that fewer retellings in the SLI group were complete, and more of their retellings were confusing. Based on these studies, we can conclude that there is a link between oral language ability (including narrative skills) and reading accuracy and comprehension.

1.4 Present study

Previous studies have shown that children with SLI exhibit problems with narrative skills. Yet, since most of these studies have a cross-sectional design, little is known about the development of these abilities. Narrative skills are closely related to reading comprehension and reading accuracy. Together with other oral language skills, these skills can play a role in the overlap between SLI and dyslexia, which is of special interest in our study. In the present longitudinal study we compare three groups of children on various oral language and narrative skills: a group with SLI and literacy delay, a group with SLI and normal literacy and a typically developing group. Oral language measures were assessed in kindergarten, grade 1 and grade 2, literacy skills were evaluated in grade 1 and grade 3 and reading comprehension skills were measured in grade 3. In two previous reports (Vandewalle, Boets, Ghesquière, & Zink, 2010; Vandewalle, Boets, Ghesquière, & Zink, in press), we described the development of phonological and literacy skills, and the interrelation between these abilities. The present study focuses on narrative and other oral language skills beyond phonology (i.e., vocabulary, morphology and syntax), and reading comprehension skills in these same groups of children.

2 Method

2.1 Participants

Eighteen children (three girls and 15 boys) with SLI participated in the study. All children were born in 2000, attended the last year of regular kindergarten at the start of the study, were monolingual Dutch speakers, had severe and persistent oral language problems and received language therapy. Oral language problems were determined by a score below percentile 3 before the start of language therapy and a score below percentile 10 after 4 years 4 months of age on at least one subtest of three standardized Dutch language tests: *Reynell Taalontwikkelingsschalen* [Reynell Developmental Language Scales] (Schaerlaekens, Zink, & van Ommeslaeghe, 2003), *Taaltests voor Kinderen* [Language Tests for Children] (van Bon & Hoekstra, 1982) or *Schlichting Test voor Taalproductie* [Schlichting Test for Language Production] (Schlichting, van Eldik, Spelberg, van der Meulen, & van der Meulen, 2003). In addition to these language problems, some children had concurrent articulation problems. Exclusion criteria were low intelligence (non-verbal, verbal and total IQ below 85), hearing loss and any other known problem (e.g., neurological disorder) that could explain their language impairment.

For every individual child of the SLI group, the best matching control child was selected based on five criteria: (1) educational environment, i.e., same class, (2) gender, (3) age, (4) non-verbal intelligence, and (5) parental educational level. Non-verbal intelligence was assessed collectively in the class with an adapted version of the Raven Coloured Progressive Matrices (Raven, Court, & Raven, 1984) and parental educational level was obtained by a questionnaire completed by the parents (based on the International Standard Classification of Education scale, UNESCO, 1997). All control children were monolingual Dutch speakers, had normal hearing and did not show any speech or language difficulties as confirmed by their parents and teachers.

Initially 26 children with SLI were selected, but only 18 children are included in the present report because one child left the study after kindergarten and seven children did not move on to grade 1 of a regular primary school after the last year of kindergarten. Consequently, the eight corresponding control children were also left out of the present report.

Both the SLI group (n = 18) and the control group (n = 18) were divided into two subgroups based on word reading and spelling achievement. The 'literacy-delayed' group comprised all children scoring 1 *SD* below population mean on a standardized word reading (van den Bos, Spelberg, Scheepstra, & De Vries, 1994) or spelling test (Dudal, 2006), in both first and third grade. This criterion took into account both the severity and the persistence of a child's literacy problem. By applying this criterion, we identified eight children with literacy delay in the SLI group and four children in the control group. In the following analyses only three groups of children will be compared: (1) children with SLI and literacy delay (SLI-LD; n = 8), (2) children with SLI and normal literacy (SLI-NL; n = 10), and (3) control children with normal literacy (Control-NL; n = 14).

Mixed Model Analyses showed that the three groups had equivalent age (p = .90) and did not differ significantly on non-verbal intelligence (p = .09). In contrast, Fisher Exact Test revealed significant group differences for educational level of the mother (p = .03). Therefore, in all group comparisons, we will control for maternal educational level as this can have an influence on language and reading comprehension skills. A more detailed description of the three groups can be found in Vandewalle et al. (in press).

Published version: http://dx.doi.org/10.1016/j.ridd.2012.05.004
Journal homepage: http://www.journals.elsevier.com/research-in-developmental-disabilities/

2.2 Measures

2.2.1 Oral language skills

The oral language test battery included tests that have not been used in the selection procedure or diagnostic assessment to avoid learning effect.

Vocabulary. Vocabulary was measured through three different tests. In the productive vocabulary test from the *Taaltoets Allochtone Kinderen* [Language test for foreign children] (Verhoeven & Vermeer, 1986) the child had to name the correct word depicted by a picture. The receptive vocabulary test and the word description test from the *Taaltoets Alle Kinderen* [Language test for all children] (Verhoeven & Vermeer, 2001) measured respectively whether the child could assign the correct picture to a given word (out of four alternatives) and could describe the word by giving a definition or a correct example.

Word fluency. To measure word fluency a newly designed test was administered. This word fluency test measured the number of words belonging to one semantic category (e.g., beverages, fruit, animals) that the child was able to produce in 20 seconds (Vandewalle et al., 2010).

Productive morphology. Morphology was measured through a productive morphological test from the *Taaltoets Alle Kinderen* [Language test for all children] (Verhoeven & Vermeer, 2001): the child had to make 12 plurals and 12 past participles, both regular as well as irregular forms.

***Sentence listening comprehension*.** This was measured through the sentence comprehension test from the *Taaltoets Alle Kinderen* [Language test for all children] (Verhoeven & Vermeer, 2001). In this task, the child had to depict the picture that represents the sentence read by the instructor out of three alternatives.

Text listening comprehension. This receptive task is also a part of the *Taaltoets Alle Kinderen* [Language test for all children] (Verhoeven & Vermeer, 2001). Different oral language skills had to be combined to perform this task. Six short stories were told to the child and immediately after each story, four questions were asked. Each correct answer resulted in one point.

2.2.2 Narrative skills

Story retelling. A Dutch translation (Jansonius-Schultheiss, Borgers, De Bruin, & Stumpel, 2007) of the Bus Story Test from the Renfrew Language Scales (Renfrew, 1997) was presented to the child and was accompanied with twelve pictures. Immediately after listening to the story, the child had to retell the story as accurately as possible while looking at the pictures. This story was mainly told in the past tense with long sentences and a lot of subordinate clauses which make the morphology and syntax of the story complex.

Storytelling. This storytelling task is part of the *Taaltoets Alle Kinderen* [Language test for all children] (Verhoeven & Vermeer, 2001). The child first had to look at eight pictures representing one sequence of events of a story. Then the child had to give his/her own interpretation of the story by telling the story while looking at the pictures.

Scoring of the story(re)telling. Firstly, all stories were transcribed literally. Afterwards different types of scores were awarded to each transcription. The *information score* was calculated as described in the manuals of both tasks. For the Bus Story Test, this score indicates the number of relevant pieces of information the child included in the story. Two points were given for each item that Renfrew designated as 'essential' and one point for each item designated as 'subsidiary'. For each element, the reference to the subject had to be correct too. If not, one point was subtracted from the information score. In the storytelling task, nine information elements and seven relations between different elements were scored by one point

each. Furthermore, the transcribed stories were divided into utterances as described in the manual of the Bus Story Test. A new utterance was identified after a coordinating conjunction (e.g., and, or, but), unless there was a coreferential subject deletion in the second clause. The *mean length of the five longest correct utterances (MLU5)* and the *mean length of all utterances* were calculated for each story. All utterances with syntactic errors in the sentence construction (e.g., incomplete sentences, wrong ordering of words) were excluded from these scores. Finally, different *syntactic and morphological measures* were calculated based on both stories: the total number of all correct and incorrect utterances, all words, the proportion of sentences with syntactic errors, the proportion verbs with a disagreement between subject and verb, the proportion errors in the used definite articles and the proportion errors in verb conjugations. In Dutch language, male and female words have the definite article /de/ and neuter words have the definite article /het/. In early language development, children often switch both articles. Moreover, since the Bus Story Test was told with a lot of subordinate clauses and mainly in the past tense, the proportion of subordinate clauses and the proportion verbs in the past tense were calculated as well.

2.2.3 Literacy skills

Reading. The One-Minute Reading test (van den Bos et al., 1994) was used as a standardized measure of single word identification. This test combines speed and accuracy into one index score. The child had to read a list of words of increasing difficulty as correctly and quickly as possible. The score on the test is the number of words read correctly within one minute. For diagnostic purposes, this score was transformed to an age-adjusted standard score relative to population average.

Spelling. Spelling was evaluated by a standardized spelling achievement test (Dudal, 2006). Children had to spell single words presented in isolation, single words presented in sentence context, and short sentences. The maximum score on the test was 60. For diagnostic purposes this score was transformed to an age-adjusted standard score relative to population average. Grade-appropriate versions of the test were used in first and third grade.

Reading comprehension. The reading comprehension task is part of the Dutch reading comprehension test for children at the end of grade 2 (Krom, van Berkel, & Jongen, 2006). After four practice items, the child had to read one story with four sentences in a wrong order and the child had to encircle the first sentence of the story. Afterwards, the child had to read silently seven short texts and immediately after reading each text, the child had to encircle the correct answer for multiple choice questions with four alternatives. Only the number of correct answers is counted.

2.3 Design and procedure

The administration of the tests was spread over a period of three years (see Table 1). Each matched pair of children from the SLI and control group was tested in a quiet room at their school under identical circumstances and in the same period. All tests were administered individually, only the tests for spelling and reading comprehension were administered collectively. Data collection was carried out by trained Master students in Speech-Language Pathology and Audiology. The data of one control child on the Bus Story retelling Test and one child with SLI on the telling task in kindergarten were removed because of their unreliable performance. Moreover, the reading level of one child with SLI was too low to administer the reading comprehension task. Hence, for these three tasks, the measurements of one child are missing.

Groups were compared using Mixed Model Analyses (MMA) with pair as a random variable and participant group (SLI-LD, SLI-NL and Control-NL group) as the fixed between-subject variable. Maternal educational level was included as fixed covariable. To investigate developmental trends, repeated measures MMA were calculated with time (kindergarten, grade 1 and grade 2) as within-subject variable, participant group as between-subjects variable and maternal educational level as covariate. Post hoc analyses were corrected for multiple comparisons using the Tukey procedure, and standardized mean differences were calculated as a measure of effect size by dividing the difference between the least-squares means by the pooled standard deviation. Given the rather small sample sizes, the statistical power to reveal significant group differences is relatively small and interpretation of the results should also be guided by inspection of these effect sizes. Kruskal Wallis nonparametric ANOVA was used to compare the three groups on the non-normally distributed proportion scores. The Bonferroni procedure for multiple comparisons was used for the post hoc analyses of the Kruskal Wallis tests by using an adjusted significance level.

Table 1
Occasions of Test Administrations

	Kindergarten		**Grade 1**	**Grade 1**	**Grade 2**	**Grade 3**
	Middle		**Middle**	**End**	**Middle**	**Begin**
years;months	5;5	5;9	6;8	6;10	7;6	8;1
Oral language skills						
Productive vocabulary	x		x		x	
Word description		x	x		x	
Word fluency		x	x		x	
Productive morphology	x		x		x	
Receptive vocabulary		x	x		x	
Sentence comprehension	x		x		x	
Text comprehension		x	x		x	
Narrative skills		x	x		x	
Literacy						
Word reading				x		x
Spelling				x		x
Reading comprehension						x

2.4 Reliability

Master students in Speech-Language Pathology and Audiology were intensively trained before administering the tests, performing the transcriptions and scoring all items based on clearly specified coding rules. The same person did the coding of all narrative tests of all children administered in the same grade to avoid effects within a grade. To check the quality and consistency of scoring over the years, results on all items were individually verified by the first author, who has extended experience with these tasks. Moreover, 25% randomly selected items of the information score of the retelling task in grade 1 and 2 were independently rescored, which resulted in a strong inter-rater reliability (r_S = .98).

Moreover, in a pilot study, the inter-rater reliability between 3 students Speech-Language Pathology was compared for the retelling task. After a training, they independently transcribed the stories and calculated

the information score and the mean length of the five longest sentences of 12 randomly selected children (9 typically developing children and 3 children with a cochlear implant). Spearman correlations revealed again a strong inter-rater reliability for the information score (r_S = .94) and an acceptable inter-rater reliabiltity for the mean length of the five longest sentences (r_S = .75).

3 Results

3.1 Development of oral language skills

The development of the different language skills (i.e. productive and receptive vocabulary, word description, word fluency, morphology, sentence and text listening comprehension) is depicted in Table 2 and Figure 1. Repeated measures MMA showed main effects of group (*p*s ≤ .0001) and grade (*p*s ≤ .0004) for all tasks, but no group by grade interactions (*p*s ≥ .09), except for the sentence listening comprehension (*p* = .005) and text listening comprehension task (*p* = .003). Both SLI groups scored significantly lower than the Control-NL group on all tasks in all grades, except for the sentence listening comprehension task in grade 1 and the text listening comprehension task in grade 1 and 2, where there were no significant differences between the SLI-NL group and the Control-NL group. There were no significant differences between both SLI groups on any of the oral language tasks across the three years, except for text listening comprehension in grade 1 and 2. It is remarkable that the SLI-NL group evolved towards the Control-NL group in grade 1 and grade 2 for text listening comprehension, while scores of the SLI-LD group remained significantly lower. A similar pattern was apparent for sentence listening comprehension from kindergarten to grade 1.

Published version: http://dx.doi.org/10.1016/j.ridd.2012.05.004
Journal homepage: http://www.journals.elsevier.com/research-in-developmental-disabilities/

Table 2
Oral Language Measures for the Three Participant Groups: Maximum (*Max*), Mean (*M*) and Standard Deviation (*SD*); Effect Size (*ES*)

			SLI-LD (1) *n* = 8		SLI-NL (2) *n* = 10		Control-NL (3) *n* = 14		Group comparisons[a] 1-2	1-3	2-3
	Max		*M*	*SD*	*M*	*SD*	*M*	*SD*	*ES*	*ES*	*ES*
Productive vocabulary	60	KG	19.6	5.0	25.2	7.1	34.1	7.8	-1.16	-2.38***	-1.23*
		G1	31.1	3.2	32.6	4.7	41.9	2.9	-0.19	-3.33***	-2.48***
		G2	35.3	1.5	37.2	5.0	43.1	3.1	-0.57	-3.04***	-1.46**
Receptive vocabulary	96	KG	56.4	5.4	51.4	10.6	68.9	7.6	0.70	-1.57*	-1.90***
		G1	64.3	7.1	65.0	7.0	76.5	5.0	-0.27	-2.06**	-1.71**
		G2	69.8	4.7	70.7	5.1	80.7	5.4	-0.10	-1.70**	-1.56**
Word description	45	KG	8.5	3.4	10.4	2.3	17.5	3.1	-0.79	-2.80***	-2.51***
		G1	12.9	4.0	12.9	3.0	20.7	4.0	0.05	-1.86**	-2.10***
		G2	10.8	3.1	15.9	5.9	22.7	4.2	-0.98	-3.13***	-1.47**
Word fluency	-	KG	3.8	0.7	4.1	0.3	5.3	0.8	-0.40	-1.85**	-1.84**
		G1	4.5	0.7	5.2	1.0	6.0	0.6	-0.73	-2.70***	-1.25*
		G2	5.1	0.9	5.6	0.5	6.5	0.7	-0.69	-1.71**	-1.32*
Productive morphology	24	KG	9.4	3.6	11.4	2.7	15.6	3.4	-0.22	-1.37*	-1.29*
		G1	13.0	3.9	15.6	2.4	19.4	2.0	-0.50	-1.83**	-1.66*
		G2	15.6	3.1	18.3	2.9	22.1	2.0	-0.78	-2.29***	-1.31*
Sentence listening comprehension	42	KG	26.3	3.4	25.0	5.1	32.5	2.8	0.35	-2.03***	-1.97*
		G1	30.0	3.9	33.0	3.3	35.1	2.3	-0.99	-1.93**	-0.78
		G2	34.0	2.1	34.4	2.5	38.0	1.5	-0.46	-2.60***	-1.74**
Text listening comprehension	24	KG	14.6	2.3	15.6	3.2	19.4	2.5	-0.55	-2.05**	-1.20*
		G1	14.0	4.6	19.1	1.4	19.9	1.8	-2.02***	-2.16***	-0.09
		G2	15.5	2.9	20.1	2.4	20.1	2.4	-2.11***	-1.94**	0.20

Note. [a] MMA with maternal educational level as covariate; * $p < .05$; ** $p < .01$; *** $p < .001$
KG = kindergarten; G = grade

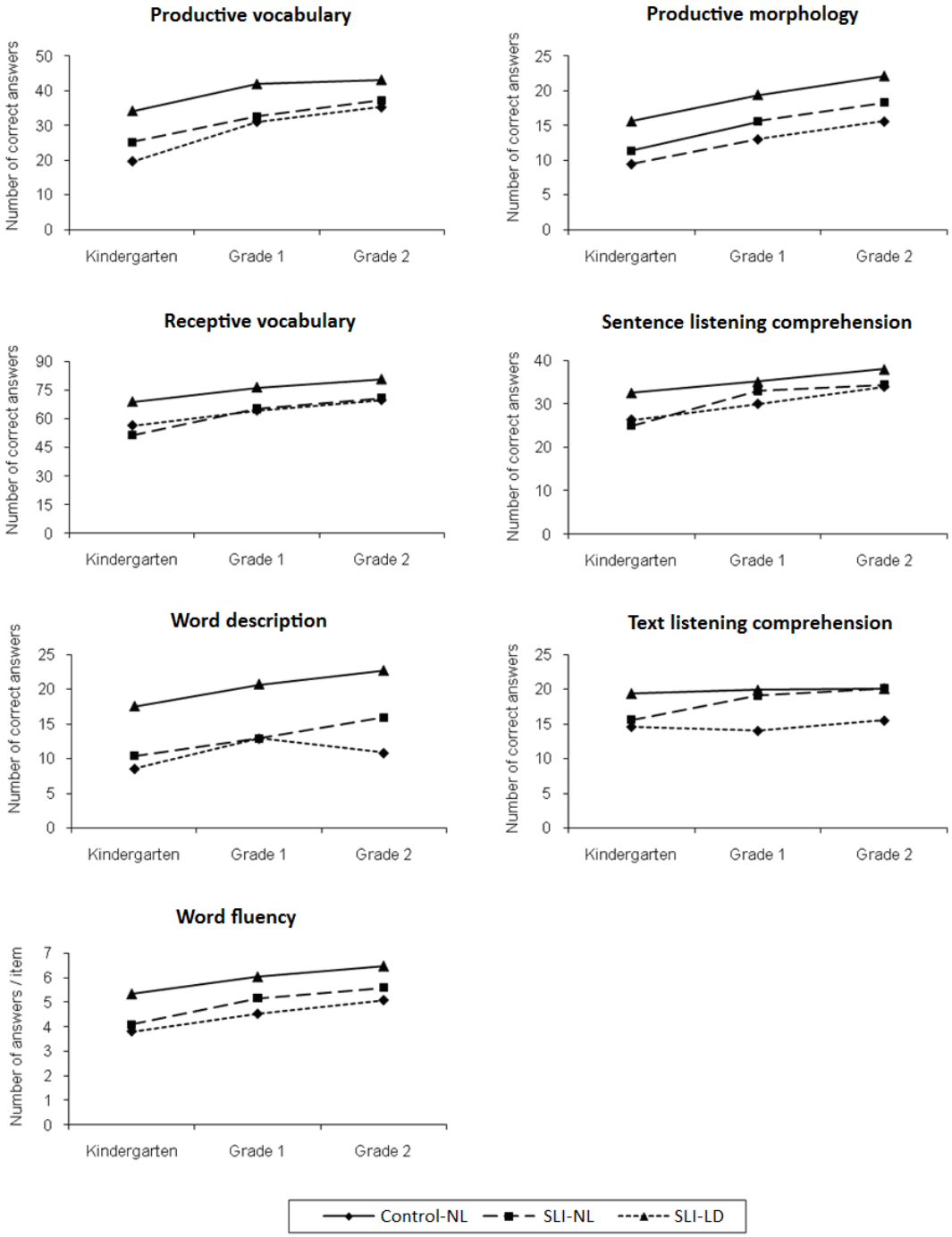

Figure 1 *Development of Oral Language Measures for the SLI-LD, SLI-NL and Control-NL Group.*

Published version: http://dx.doi.org/10.1016/j.ridd.2012.05.004
Journal homepage: http://www.journals.elsevier.com/research-in-developmental-disabilities/

3.2 Development of narrative skills

The development of the information score and MLU5 on the story retelling and storytelling tasks can be observed in Table 3 and Figure 2. Repeated measures MMA showed main effects of group ($p \leq .0008$) and grade ($p < .005$) for the information score and MLU5 for the retelling and the telling task, but no significant interaction effects ($p \geq .60$).

In general, the SLI-LD group scored significantly lower than the Control-NL group across almost all grades. However, the SLI-NL group showed a differential pattern on both narrative tasks. On the story retelling task, the SLI-NL group scored rather similar to the SLI-LD group and significantly lower than the Control-NL group. In contrast, on the storytelling task, the SLI-NL group scored in between the Control-NL and the SLI-LD group, and did not differ from these two groups in kindergarten. Moreover, across time the SLI-NL group evolved in the direction of the Control-NL group on the storytelling task, while the SLI-LD group continued lagging behind. Group differences on the mean length of all utterances index and MLU5 were very similar for both the retelling and telling task. Hence, only MLU5 scores were presented.

In the retelling task, the children were implicitly encouraged to make longer sentences as the story was told in long sentences. This resulted in a significantly higher MLU5 on the retelling task as compared to the telling task over the three groups and years (ps $< .0001$).

Repeated measures MMA showed that there were no group differences in the number of sentences used (correct and incorrect) in the retelling ($p = .09$) and the telling task ($p = .52$). The average number of words that were used in the retelling and telling task did differ between the three groups for both tasks (ps $< .04$): the two SLI groups used significantly fewer words than the Control-NL group. Kruskal Wallis tests revealed also that the SLI-LD group made significantly more errors in sentence constructions (e.g., incomplete sentences, wrong ordering of words) than the Control-NL group on the retelling task in grade 1 and on the telling task in each of the three grades (see Appendix).

Different morphological skills were scored for the three groups as well (see Appendix). The SLI-LD group made significantly more errors in verb conjugations than the Control-NL group on the retelling task in grade 2. In other grades and on the telling task, there were no significant differences between the three groups. For subject-verb agreement, Kruskal Wallis tests showed that the SLI-LD group made significantly more errors than the Control-NL group on both the retelling and the telling task in kindergarten. In grade 1, both SLI groups made more subject-verb agreement errors than the Control-NL group on both tasks. In grade 2, no group differences were retained. Although hardly any errors in the use of definite articles were made in any of the groups, children of the SLI-LD group made significantly more errors than children of the Control-NL group on the retelling task in grade 2. Kruskal Wallis test showed that children in the SLI-LD group used significantly fewer subordinate clauses than the Control-NL group in the retelling task in all grades. There was a large variability within the SLI-NL group in grade 1 and 2. On the telling task, however, no group differences were retained in the number of subordinate clauses. Analyzing the use of the past tense in the retelling task, Kruskal Wallis tests revealed that both SLI groups used the past tense less frequently in kindergarten compared to the Control-NL group. In grade 1, the SLI-NL group still used the past tense less frequently than the Control-NL group. However, in grade 2, the SLI-NL group evolved towards the level of the Control-NL group, whereas the SLI-LD group made significantly more errors than the Control-NL group.

Table 3
Narrative and Literacy Measures for the Three Participant Groups: Maximum (*Max*), Mean (*M*) and Standard Deviation (*SD*); Effect Size (*ES*)

			SLI-LD (1)		SLI-NL (2)		Control-NL (3)		Group comparisons[a]		
			n = 8		*n* = 10		*n* = 14		1-2	1-3	2-3
	Max		*M*	*SD*	*M*	*SD*	*M*	*SD*	*ES*	*ES*	*ES*
Story retelling											
Information score	53	KG	15.3	3.7	18.7	7.6	29.4	8.0	-0.79	-2.24**	-1.32*
		G1	22.4	7.0	23.8	7.2	32.9	6.0	-0.29	-1.71*	-1.35*
		G2	24.8	5.5	27.9	7.5	34.3	4.6	-0.49	-1.91*	-1.03*
MLU5	-	KG	6.8	0.8	7.5	0.8	9.0	1.7	-0.87	-1.40*	-0.94
		G1	7.5	1.8	7.6	1.6	9.3	1.6	-0.10	-1.25**	-1.18**
		G2	8.4	0.7	8.8	2.0	10.4	1.4	-0.45	-2.01*	-1.02*
Storytelling											
Information score	16	KG	6.8	1.4	8.8	2.3	10.3	1.8	-1.11	-2.51**	-0.95
		G1	8.9	2.8	10.6	1.8	11.1	1.9	-1.24*	-1.35*	-0.12
		G2	9.1	3.1	10.8	1.9	11.0	2.4	-0.40	-0.59	-0.25
MLU5	-	KG	5.4	0.6	6.0	1.4	6.9	1.3	-0.82	-1.24	-0.50
		G1	6.2	1.2	6.9	1.7	7.8	1.9	-0.58	-1.60*	-0.69
		G2	6.2	1.0	7.9	1.4	8.6	1.6	-1.30	-1.71**	-0.49
Word reading[b]	-	G1	91.3	7.9	100.5	6.9	107.9	15.0	-0.79	-0.79	-0.37
		G3	78.1	19.3	102.0	11.8	107.9	14.4	-1.2	-1.45*	-0.36
Spelling[b]	-	G1	74.3	8.8	92.4	10.3	100.3	14.0	-1.59	-1.90**	-0.65
		G3	67.1	10.7	92.1	12.6	107.3	14.5	-1.82*	-2.51***	-0.86
Reading comprehension	25	G3	7.9	2.2	13.5	3.7	16.9	3.8	-1.80*	-2.39***	-0.60

Note. [a]MMA with maternal educational level as covariate; [b]standard scores with mean = 100, SD = 15; * $p < .05$, ** $p < .01$, *** $p < .001$

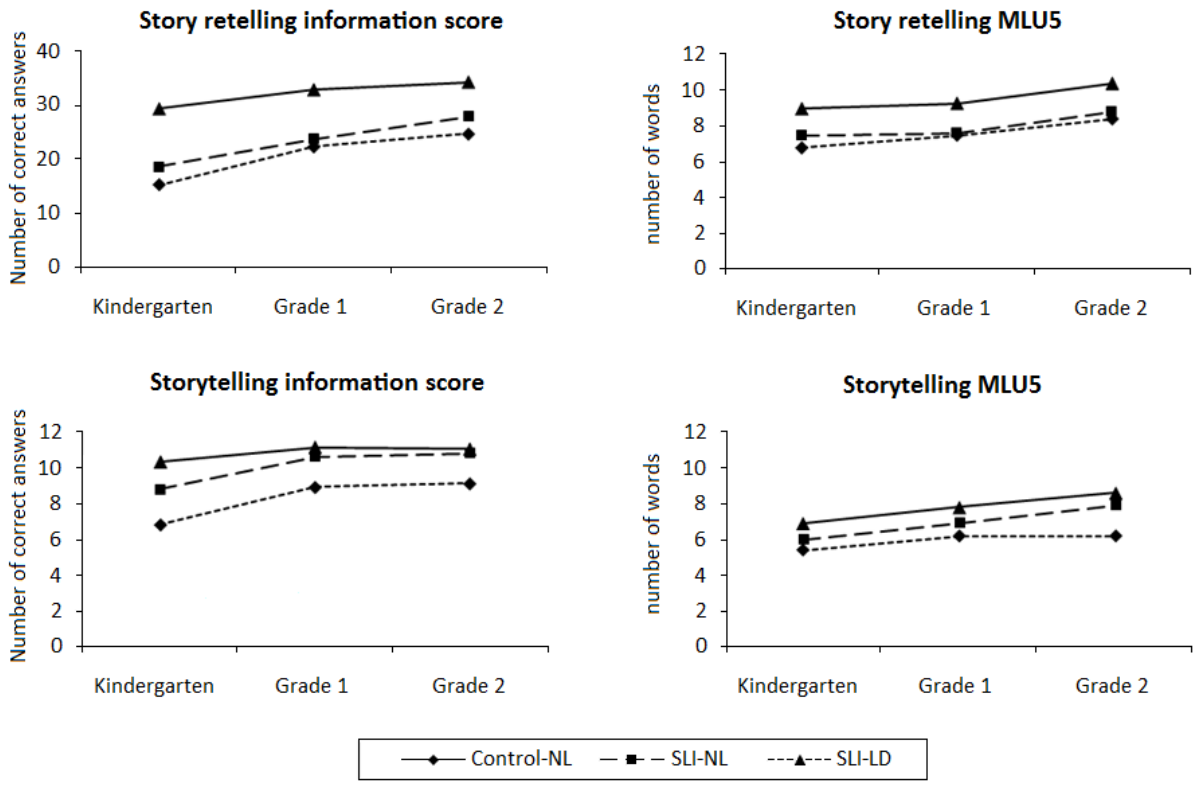

Figure 2 *Development of Narrative Measures for the SLI-LD, SLI-NL and Control-NL Group.*

3.3 Reading comprehension skills

Reading comprehension was only measured in grade 3. Table 3 shows that the SLI-LD group scored significantly poorer than the SLI-NL and the Control-NL group. The SLI-NL group had similar scores as the Control-NL group.

4 Discussion

From our study we can conclude that *children with SLI and literacy delay* showed persistent language problems across all the assessed language domains: vocabulary, morphology, syntax and narrative skills (storytelling and retelling). They experienced productive as well as receptive oral language problems from kindergarten until at least grade 2. Hence, in spite of the fact that all children with SLI had started oral language therapy in kindergarten, they did not show any recovery of their language problems. Our observation that children with SLI and literacy delay (and to a minor extent also those with normal literacy development) showed persistent narrative problems until grade 2, contrasts with the findings of Paul et al. (1996) but confirms the conclusions of Fey et al. (2004). We observed that these children made more errors than control children in sentence construction, verb conjugation, subject-verb agreement and the use of definite articles. In the Bus Story retelling Test, they used subordinate clauses and the past tense less frequently. These observations are in accordance with other studies (Liles et al., 1995; Merritt & Liles, 1987; Norbury & Bishop, 2003), although we did not observe any group difference in the number of

sentences used in the narrative tasks (in contrast to the study of Liles, 1985). In accordance with other studies (Bishop & Adams, 1990; Conti-Ramsden, Botting, Simkin, & Knox, 2001), the children with SLI and literacy delay also showed problems with reading comprehension in the beginning of grade 3.

The c*hildren with SLI and normal literacy* also scored persistently low from kindergarten till grade 2 on vocabulary, morphology and story retelling skills. However, for some oral language tasks, this group caught up with the control group.

First, for text listening comprehension and partly for sentence listening comprehension, the children with SLI and normal literacy evolved towards the control group while the SLI group with literacy delay continued lagging behind. In these tasks, children had to combine different receptive oral language skills in order to complete the task. Surprisingly, the children with SLI and normal literacy experienced problems with the isolated skills (e.g., receptive vocabulary), but not in combining the skills in an integrated manner to comprehend a sentence or text. Because most improvement was accomplished from kindergarten to grade 1, we postulate that this improvement might have been provoked by the exposure to written text during literacy instruction. The SLI group with normal literacy received more print exposure as a result of their better reading than the SLI group with literacy delay.

Second, on storytelling, the SLI group with normal literacy had in general better scores than the SLI group with literacy delay, both on the information score and on the mean length of the five longest sentences. Already in kindergarten, this group performed at a similar level as the control group, which suggests that this good performance level resulted from spontaneous development or compensation. Certainly, it could not be the effect of literacy instruction or print exposure because formal literacy instruction started one year later. In grades 1 and 2, the group evolved further towards the control group. Remarkably, both on the information score (measuring the content – an aspect of narrative macrostructure) and on the mean length of the five longest sentences (measuring the linguistic form – an aspect of narrative microstructure) group differences showed a similar pattern (see Figure 2). However, there was an effect of the kind of task, i.e. telling versus retelling. The SLI group with normal literacy scored consistently poorer on story retelling (both content and linguistic form) compared to the control group, but scored nearly identically to the control group for storytelling (both content and linguistic form). An effect of task was also observed by Merritt and Liles (1987), but not by Befi-Lopes et al. (2008). Storytelling is a complex task in which different oral language skills have to be combined. Therefore, we expected that the deficits would have been more persistent compared to those for the individual language skills (Paul et al., 1996), but this was not the case. Another influencing effect might be the use of pictures. McCabe et al. (2008) observed that a lot of children with language impairment treated the pictures in isolation from each other rather than a series of events. In contrast, personal narratives (without presenting pictures) were often more coherent stories. In our study, pictures were used in both tasks, but we suppose that this effect was larger in the storytelling than in the story retelling task because in the storytelling task, the children could only get information based on the pictures. However, this effect did not seem to cause more differentiation between the three groups. Another potential explanation of the larger group differences in the retelling task might be the fact that control children benefit more from the told story (the auditive information) in contrast to the two SLI groups. However, based on the improved scores for listening comprehension in the SLI-NL group in grades 1 and 2, we would expect improved retelling scores in this group as well, which was not the case. The different pattern of performance on story retelling versus storytelling might have been provoked by the fact that children could freely choose the difficulty level of expression (e.g., the length of the sentences) in the storytelling task. In contrast, the difficulty level was somehow imposed in the retelling task, where children were implicitly encouraged to make longer and more complex sentences. This was evidenced by the observation that children produced shorter sentences in the telling task as compared to the retelling task.

As a consequence, these shorter and easier sentences and stories in the telling task might have been less sensitive to differentiate between the groups.

There is a narrow link between oral language and reading comprehension skills, which is also evidenced by several studies showing that a high proportion of children with SLI develop reading comprehension problems (Bishop & Adams, 1990; Conti-Ramsden et al., 2001). Hence we expected the entire SLI group to be at risk for reading comprehension problems. However, we observed that only the SLI group with literacy delay developed reading comprehension problems in grade 3, whereas the SLI group with normal literacy did not present reading comprehension problems despite their oral language problems. This pattern of findings suggests that reading comprehension performance at this age was mainly determined by word reading skills and not by oral language skills. It should be noted that the texts of this particular reading comprehension task were short and easy. Possibly at a later age when texts become longer and use more complex language, the SLI group with normal literacy may develop reading comprehension problems as a consequence of their remaining oral language problems.

To investigate developmental trends in longitudinal studies, it is important to compare scores for the same tests measured at different time instants, as was done in this study for the narrative and other oral language skills. However, then the scores might be influenced by unwanted learning effects of the specific task. In order to minimize these learning effects, in our study the interval between the consecutive test administrations was kept large (i.e., yearly).

More longitudinal studies with larger groups, in different languages, spanning a longer period and using a wider range of oral language and narrative measures are needed. In particular, a more extensive comparison of children with SLI with and without literacy delay is desirable to generalize our results. Because a good narrative ability involves many different skills such as working memory, vocabulary, pragmatic skill, and meta-linguistic knowledge (Botting, 2010), further research is needed to analyze the underlying mechanisms that play a role in the development of narrative skills in both SLI groups. Early diagnosis and effective therapy will ultimately be based on a solid understanding of the underlying relations over time.

Acknowledgments

The authors wish to thank Intse Boey, Ellen De Wever, Joke Lauwers and Sarah Saey for their contribution in the data collection and Ine Deckx and Marijke Vanelven for their help with the transcription and scoring of the stories. Moreover, the authors wish to acknowledge the rehabilitation centers, speech-language therapists, and the *Vlaamse Vereniging voor Logopedisten (VVL)* [Flemish Union of Speech-Language Therapists] for assistance with the selection of the children with SLI, and all the children, their parents, schools, and teachers for their commitment to the study.

Appendix Median (Med) and Range of Percentages of Different Scores in the Storytelling and Story Retelling Task for the Three Participant Groups

		Story retelling task			Storytelling task		
		SLI-LD	**SLI-NL**	**Control-NL**	**SLI-LD**	**SLI-NL**	**Control-NL**
		$n = 8$	$n = 10$	$n = 14$	$n = 8$	$n = 10$	$n = 14$
		Med (Range)	*Med (Range)*	*Med (Range)*	*Med (Range)*	*Med (Range)*	*Med (Range)*
% erroneous sentences	KG	14 (5-36)	10 (0-36)	7 (0-31)	23 (9-50)$_a$	10 (0-40)$_{ab}$	11 (0-18)$_b$
	G1	20 (5-72)$_a$	16 (0-45)$_{ab}$	5 (0-24)$_b$	26 (9-42)$_a$	13 (0-42)$_{ab}$	4 (0-25)$_b$
	G2	21 (0-31)	6 (0-21)	5 (0-20)	18 (0-44)$_a$	9 (0-25)$_{ab}$	0 (0-30)$_b$
% errors in verb conjugations	KG	6 (0-14)	4 (0-14)	4 (0-24)	0 (0-0)	0 (0-11)	0 (0-20)
	G1	7 (0-58)	5 (0-13)	2 (0-14)	3 (0-20)	0 (0-5)	0 (0-18)
	G2	7 (0-29)$_a$	7 (0-16)$_{ab}$	4 (0-16)$_b$	11 (0-36)	7 (0-12)	0 (0-18)
% wrong subject-verb agreement	KG	6 (0-18)$_a$	6 (0-20)$_{ab}$	0 (0-31)$_b$	13 (0-43)$_a$	8 (0-33)$_{ab}$	0 (0-10)$_b$
	G1	2 (0-16)$_a$	0 (0-7)$_a$	0 (0-0)$_b$	10 (0-31)$_a$	5 (0-20)$_a$	0 (0-9)$_b$
	G2	0 (0-11)	0 (0-4)	0 (0-4)	0 (0-7)	4 (0-13)	0 (0-9)
% wrong articles	KG	0 (0-0)	0 (0-13)	0 (0-6)	0 (0-40)	8 (0-67)	0 (0-33)
	G1	0 (0-6)	0 (0-10)	0 (0-0)	0 (0-50)	0 (0-29)	0 (0-13)
	G2	3 (0-8)$_a$	0 (0-10)$_{ab}$	0 (0-0)$_b$	0 (0-17)	0 (0-17)	0 (0-0)
% subordinate clauses	KG	0 (0-22)$_a$	5 (0-13)$_{ab}$	13 (0-35)$_b$	0 (0-0)	0 (0-9)	3 (0-13)
	G1	6 (0-14)$_a$	5 (0-33)$_{ab}$	15 (0-29)$_b$	0 (0-13)	0 (0-10)	0 (0-9)
	G2	6 (0-16)$_a$	8 (0-35)$_{ab}$	22 (00-54)$_b$	0 (0-11)	0 (0-33)	3 (0-25)
% past tense	KG	40 (12-76)$_a$	61 (14-86)$_a$	85 (52-100)$_b$			
	G1	75 (16-95)$_{ab}$	73 (14-100)$_a$	90 (50-96)$_b$			
	G2	64 (21-83)$_a$	83 (44-94)$_{ab}$	87 (74-96)$_b$			

Note. Different subscripts designate significant differences ($p < .017$) using Kruskal Wallis tests with Bonferroni correction

REFERENCES

Befi-Lopes, D. M., Bento, A. C. P., & Perissinoto, J. (2008). Narration of stories by children with specific language impairment. *Pró-Fono Revista de Actualização Científica, 20*(2), 93-98.

Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, 31*(7), 1027-1050.

Bishop, D. V. M., & Edmundson, A. (1987). Language-impaired 4-year-olds: distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156-173.

Botting, N. (2010). 'It 's not (just) what you do, but the way that you do it': factors that determine narrative ability in atypical language learners. *Developmental Medicine and Child Neurology, 52*, 885-890.

Botting, N., Simkin, Z., & Conti-Ramsden, G. (2006). Associated reading skills in children with a history of Specific Language Impairment. *Reading and Writing, 19*, 77-98.

Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48*, 1378-1396.

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: outcomes at 11 years of age. *International Journal of Language and Communication Disorders, 36*(2), 207-219.

Crais, E. R., & Lorch, N. (1994). Oral narratives in school-age children. *Topics in Language Disorders, 14*(3), 13-28.

Dudal, P. (2006). *Leerlingvolgsysteem VCLB. Spelling: toetsen 1-6. Basisboek en kopieerbundel.* Leuven, Belgium: Garant.

Feagans, L., & Appelbaum, M. I. (1986). Validation of language subtypes in learning disabled children. *Journal of Educational Psychology, 78*(5), 358-364.

Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research, 47*, 1301-1318.

Finestack, L. H., Fey, M. E., & Catts, H. W. (2006). Pronominal reference skills of second and fourth grade children with language impairment. *Journal of Communication Disorders, 39*, 232-248.

Gillam, R. B., & Carlile, R. M. (1997). Oral reading and story retelling of students with specific language impairment. *Language, Speech, and Hearing Services in Schools, 28*, 30-42.

Gilmore, S. E., Klecan-Aker, J. S., & Owen, W. L. (1999). The relationship of storytelling ability to reading comprehension in children with learning disability. *Journal of Speech-Language Pathology and Audiology, 23*(3), 142-151.

Hook, P. E., & Haynes, C. W. (2009). Reading and writing in child language disorders. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 424-444). New York, U.S.: Psychology Press.

Jansonius-Schultheiss, K., Borgers, M., De Bruin, B. C., & Stumpel, H. (2007). Renfrew's taalschalen Nederlandse aanpassing (RTNA): diagnostisch onderzoek en therapieopzet bij jonge taalgestoorde kinderen. *Logopedie en Foniatrie, 11*, 352-359.

Kit-Sum To, C., Stokes, S. F., Cheung, H.-T., & T'sou, B. (2010). Narrative assessment for Cantonese-speaking children. *Journal of Speech, Language, and Hearing Research, 53*, 648-669.

Krom, R., van Berkel, S., & Jongen, I. (2006). *Begrijpend lezen groep 4: handleiding.* Arnhem, the Netherlands: Cito B.V. Arnhem.

Liles, B. Z. (1985). Cohesion in the narratives of normal and language disordered children. *Journal of Speech, Language, and Hearing Research, 28*, 123-133.

Liles, B. Z. (1993). Narrative discourse in children with language disorders and children with normal language: a critical review of literature. *Journal of Speech and Hearing Research, 36*(5), 868-882.

Liles, B. Z., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech, Language, and Hearing Research, 38*, 415-425.

Manhardt, J., & Rescorla, L. (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics, 23*, 1-21.

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

McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry, 41*(7), 869-874.

McCabe, A., Bliss, L., Barra, G., & Bennett, M. (2008). Comparison of personal versus fictional narratives of children with language impairment. *American Journal of Speech-Language Pathology, 17*, 194-206.

Merritt, D. D., & Liles, B. Z. (1989). Narrative analysis: clinical applications of story generation and story retelling. *Journal of Speech and Hearing Disorders, 54*(3), 429-438.

Merritt, D. D., & Liles, B. Z. (1987). Story grammar ability in children with and without language disorder: story generation, story retelling, and story comprehension. *Journal of Speech, Language, and Hearing Research, 30*, 539-552.

Muñoz, M. L., Gillam, R. B., Peña, E. D., & Gulley-Faehnle, A. (2003). Measures of language development in fictional narratives of Latino children. *Language, Speech, and Hearing Services in Schools, 34*, 332-342.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*, 342-356.

Norbury, C. F., & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders, 38*(3), 287-313.

Pankratz, M. E., Plante, E., Vance, R., & Insalaco, D. M. (2007). The diagnostic and predictive validity of the Renfrew Bus Story. *Language, Speech, and Hearing Services in Schools, 38*, 390-399.

Paul, R., Hernandez, R., Taylor, L., & Johnson, K. (1996). Narrative development in late talkers: early school age. *Journal of Speech, Language, and Hearing Research, 39*, 1295-1303.

Paul, R., & Smith, R. L. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech, Language, and Hearing Research, 36*, 592-598.

Raven, J. C., Court, J. H., & Raven, J. (1984). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* London, U.K.: Lewis.

Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research, 54*, 99-117.

Renfrew, C. E. (1997). *The Renfrew Language Scales: Bus Story Test: a test of narrative speech* (4th ed.). Bicester, U.K.: Winslow Press Limited.

Schaerlaekens, A., Zink, I., & van Ommeslaeghe, K. (2003). *Reynell Taalontwikkelingsschalen: handleiding tweede versie.* Lisse, the Netherlands: Swets & Zeitlinger.

Schlichting, J. E. P. T., van Eldik, M. C. M., Spelberg, H. C. L., van der Meulen, Sj., & van der Meulen, B. F. (2003). *Schlichting Test voor Taalproductie: handleiding, derde gewijzigde druk.* Lisse, the Netherlands: Swets & Zeitlinger.

Schwartz, R. G. (2009). Specific Language Impairment. In R. G. Schwartz (Ed.), *Handbook of Child Language Disorders* (pp. 3-43). New York, U.S.: Psychology Press.

Seiger-Gardner, L. (2009). Language production approaches to child language disorders. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 465-487). New York, U.S.: Psychology Press.

Snowling, M., Bishop, D. V. M., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence. *Journal of Child Psychology and Psychiatry, 41*(5), 587-600.

Soodla, P., & Kikas, E. (2010). Macrostructure in the narratives of Estonian children with typical development and language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 1321-1333.

Spear-Swerling, L. (2006). Children's reading comprehension and oral reading fluency. *Reading and Writing, 19*, 199-220.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: a follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*, 407-418.

Strong, C. J., & Shaver, J. P. (1991). Stability of cohesion in the spoken narratives of language-impaired and normally developing school-aged children. *Journal of Speech and Hearing Research, 34*, 95-111.

UNESCO (1997). *International Standard Classification of Education*.

van Bon, W. H. J., & Hoekstra, J. G. (1982). *Taaltest voor kinderen: handleiding.* Lisse, the Netherlands: Swets & Zeitlinger.

van den Bos, K. P., Spelberg, H. C. L., Scheepstra, A. J. M., & De Vries, J. R. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling.* Nijmegen, the Netherlands: Berkhout.

Vandewalle, E., Boets, B., Ghesquière, P., & Zink, I. (2010). Who is at risk for dyslexia? Phonological processing in five- to seven-year-old Dutch-speaking children with SLI. *Scientific Studies of Reading, 14*(1), 58-84.

Vandewalle, E., Boets, B., Ghesquière, P., & Zink, I. (in press). Development of phonological processing and literacy skills in children with specific language impairment: a three-year longitudinal study. *Journal of Speech, Language, and Hearing Research*.

Verhoeven, L., & Vermeer, A. (1986). *Taaltoets Allochtone Kinderen: diagnostische toets voor de mondelinge vaardigheid Nederlands bij allochtone kinderen van 5-9 jaar (handleiding, toetsboeken, scoreboeken).* Tilburg, the Netherlands: Zwijsen b.v.

Verhoeven, L., & Vermeer, A. (2001). *Taaltoets Alle Kinderen: diagnostische toets voor de mondelinge taalvaardigheid Nederlands bij kinderen van groep 1 tot en met 4 (handleiding, toetsboeken, scoreboeken).* Arnhem, the Netherlands: Cito.

Westerveld, M. F., & Gillon, G. T. (2010). Oral narrative context effects on poor readers' spoken language performance: story retelling, story generation, and personal narratives. *International Journal of Speech Language Pathology, 12*(2), 132-141.

Westerveld, M. F., Gillon, G. T., & Moran, C. (2008). A longitudinal investigation of oral narrative skills in children with mixed reading disability. *International Journal of Speech Language Pathology, 10*(3), 132-145.

Research Coordination Office
Huis Bethlehem
Schapenstraat 34
B-3000 Leuven
Belgium
Tel.: +32 16 32 40 65
Fax: +32 16 32 41 98
onderzoek@kuleuven.be
www.kuleuven.be