JSLHR

Research Article

Development of Morphosyntactic Accuracy and Grammatical Complexity in Dutch School-Age Children With SLI

Rob Zwitserlood,[a,c] Marjolijn van Weerdenburg,[b] Ludo Verhoeven,[b] and Frank Wijnen[c]

Purpose: The purpose of this study was to identify the development of morphosyntactic accuracy and grammatical complexity in Dutch school-age children with specific language impairment (SLI).
Method: Morphosyntactic accuracy, the use of dummy auxiliaries, and complex syntax were assessed using a narrative task that was administered at three points in time (T1, T2, T3) with 12-month intervals during a 2-year period. Participants were 30 monolingual Dutch children with SLI, age 6;5 (years;months) at T1; 30 typically developing peers, age 6;6 at T1; and 30 typically developing language-matched children, age 4;7 at T1.
Results: On the morphosyntactic accuracy measures, the group with SLI performed more poorly than both control groups. Error rates in the group with SLI were much higher than expected on the basis of mean length of T-units and scores on standardized language tests. Percentages of dummy auxiliaries remained high over time. No group differences were found for grammatical complexity, except at T3, when the group with SLI used fewer relative clauses than the typically developing peer group.
Conclusions: The narrative analysis demonstrates different developmental trajectories for morphosyntactic accuracy and grammatical complexity in children with SLI and typically developing peer and language-matched children. In the group with SLI, grammatical skills continue to develop.

A prominent question in research on specific language impairment (SLI) is whether the disorder affects various different language competencies in equal measure. In the past, this issue figured in the delay versus deviance debate. Leonard (1998) pointed out that children with SLI can differ from typically developing (TD) children in various ways, and he proposed to differentiate among five qualitatively different patterns of atypical language development. One of these, he suggested, is characterized by an *uneven profile*, which implies that the developmental delay for certain elements of morphosyntax is markedly larger than for others. Later classifications of the developmental trajectories in SLI, such as those proposed by Law, Tomblin, and Zhang (2008) and Thomas et al. (2009), overlap in part with Leonard's framework, but these seem to pay much less attention to the concept of an uneven developmental profile.

Nonetheless, the proposals by Leonard (1998), Law et al. (2008), and Thomas et al. (2009) clearly converge on one important point: To adequately characterize SLI, and its phenotypical variants, and thus to better understand the nature of the disorder, it is critical to carefully investigate the developmental trajectories of affected individuals. This would seem relevant to the issue of uneven development as well. Single-time-point cross-sectional studies may suggest that one particular component of language is more affected than another, but only a longitudinal design will tell us if such a difference is persistent. Of course, persistence is one of the key characteristics of developmental disorders.

Longitudinal Studies of Language Development in SLI

In order to investigate developmental patterns in children with SLI, a three-group design using TD language

[a]Royal Dutch Auris Group, Gouda, the Netherlands
[b]Behavioural Science Institute, Radboud University Nijmegen, the Netherlands
[c]Utrecht Institute of Linguistics OTS, Utrecht University, the Netherlands
Correspondence to Rob Zwitserlood: r.zwitserlood@gmail.com

Editor: Rhea Paul
Associate Editor: Sandra Gillam
Received January 17, 2014
Revision received January 13, 2015
Accepted February 26, 2015
DOI: 10.1044/2015_JSLHR-L-14-0015

Disclosure: The authors have declared that no competing interests existed at the time of publication.

age–matched (LA) and chronological age–matched (CA) control groups is the most appropriate choice. Moreover, to reveal different and possibly unique developmental patterns of various different components of language, a longitudinal approach examining a wide range of measures is necessary (Paradis, 2010). To date, the number of studies that have taken such an extensive and longitudinal approach is limited.

Curtiss, Katz, and Tallal (1992) longitudinally examined the order of acquisition of a set of linguistic structures using comprehension and production tasks. Participants were children with SLI age 4;4 (years; months) and LA children age 2;9. Picture-pointing tasks (comprehension) and sentence-completion tasks (production) of an omnibus language test were administered repeatedly over a 5-year period. The results indicated that the developmental trajectories of children with SLI were similar to those of LA controls and that, therefore, grammatical development in SLI should be seen as (merely) delayed. No uneven profiles were reported. However, a few points may have affected the validity of this study. First, data loss of 20% was reported during test rounds, partly in the higher-level subtests of the children with SLI. Second, the large difference in age range between the LA group (1;7 years) and the SLI group (0;9 years) may have influenced the results. Third, a sentence-completion task might not be the most suitable way to study expressive language because of the highly structured nature of such an elicitation task. Moreover, the exclusive use of a standardized language test, rather than combining it with an analysis of conversational or narrative samples, may have limited the likelihood of finding subtle or individual deviations from the normal (expected) developmental trajectory.

Rice, Wexler, and Hershberger (1998) used a longitudinal three-group design with LA and CA control groups to investigate the development of verb and noun morphology in 5-year-old children with SLI over a 3-year period. The children with SLI performed poorly compared with LA controls on the correct production of tense and agreement, but they had age-appropriate scores on noun morphology (regular *-s* plurals). With respect to verb morphology, the children with SLI showed no signs of catching up over time. Thus, tense and agreement lagged further behind in these children than other elements of morphosyntactic development, which is a clear example of an uneven developmental pattern, and which suggests that certain aspects of grammar (in this case, the verb inflection system) are more vulnerable than others.

Rice (2003) described another selective delay pattern. Vocabulary development of children with SLI was similar to that of LA children and could be regarded as delayed. By contrast, the SLI group performed much more poorly than the LA control group on tense and agreement tasks. This unexpected delay in grammatical morphology was regarded as deviant compared with the development of other language skills. Rice (2003) described this unevenness in language development as a *delay-within-delay pattern*, which falls within the boundaries of the *uneven profile* in Leonard's (1998) taxonomy.

The studies by Rice and colleagues (Rice, 2003; Rice et al., 1998) compared acquisition of verb morphology with development in vocabulary and noun morphology. As yet, we do not know whether the conclusions drawn by Rice and colleagues extend to other domains. Specifically, evidence is mounting that SLI specifically targets verb-related morphosyntax, but it is unclear whether other components of grammar, such as complex grammar, or functional categories, such as determiners, pronouns, and prepositions, are affected as well in children with evident difficulties in verb-related morphosyntax. This study addresses this question. It aims to determine whether a defective development of verb-related morphosyntax implies defective development in other domains of grammar such as the syntax of complex sentences and of functional elements outside the verb-related domain.

SLI in Dutch

The present study concerns the acquisition of Dutch, which is a language that, in comparison to English, has some enriched morphosyntactic features that may be useful for a more detailed characterization of developmental trajectories in SLI. First, Dutch is an S(ubject)-O(bject)-V(erb) + verb-second language, which means that the inflected (i.e., finite) verb takes second position in main clauses and final position in subordinate clauses. In main clauses, only one constituent can precede the inflected verb. Nonfinite verbs (infinitives, participles) always appear in clause final position. First, it has been documented that TD Dutch children go through a stage in which they use dummy auxiliaries as placeholders for inflected lexical verbs in second sentence position (Wijnen & Verrips, 1998). This stage occurs prior to mastery of the rules for verb inflection. Several studies have noted that this dummy auxiliary stage is markedly longer in Dutch children with SLI than in TD controls. It has been suggested that inserting an inflected dummy auxiliary is less demanding than computing the inflected form of a lexical verb, and this may explain the behavior of children with SLI (de Jong, 1999; de Jong, Blom, & Orgassa, 2013; Orgassa, 2009). Second, the Dutch inflectional paradigm for verbs is somewhat richer than the English system, and this has its consequences for the symptomatology of SLI. English-speaking children with SLI predominantly omit inflectional morphemes of lexical verbs, whereas Dutch-speaking children with SLI not only omit these morphemes, but they also often make substitution errors (Blom, Vasić, & de Jong, 2014; de Jong, 1999). Third, Dutch has a more elaborate determiner system than English, with two classes of definite determiners: common gender *de, deze, die* [the, this, that] and neuter gender *het, dit, dat* [the, this, that]. For singular nouns, there has to be gender agreement between determiner and noun. English-speaking children with SLI are known to omit determiners, but analogous to errors in verb morphology, Dutch-speaking children with SLI also make substitution errors, notably replacing neuter gender elements by common gender morphemes (e.g., substituting definite article *de* for *het*; Keij, Cornips, van Hout, Hulk, & van Emmerik, 2012). A fourth characteristic is related to this gender agreement between determiner and noun. In relative clauses, the relative pronoun *die* or *dat* [who or that] has to agree with the gender of the relativized noun in the main clause. Finally, a feature of Dutch known to be very problematic for children with SLI (Bol & Kuiken, 1990) is the correct use of pronominal/adverbial *er* [there]. Dutch children with SLI often omit this word in obligatory contexts. The use of *er* can influence word order in sentences, and five different functions are generally distinguished (Shetter & Ham, 2007): (a) expletive and (b) existential *er,* both inserted for grammatical reasons; (c) *er* as the unstressed form of *daar* [there], expressing location; (d) quantitative *er,* referring to a number; and (e) prepositional *er,* which is used in combination with a preposition, replacing certain personal and demonstrative pronouns. An examination of these specific characteristics of Dutch longitudinally in children with SLI may offer a window into different developmental trajectories for verb-related and non–verb-related morphosyntax and complex syntax.

Aims and Approach

This study aims to contribute to our knowledge of grammatical development in children with SLI. We address the question of whether different components of grammar are affected in equal measure (as compared with the profile of TD children) by comparing the development of verb-related and non–verb-related morphosyntax with grammatical complexity (complex syntax). We do so, not by using standardized tests (Curtiss et al., 1992) or spontaneous speech samples combined with elicitation (Rice et al., 1998), but by using a narrative task. Narratives have been shown to provide rich information on the development of vocabulary, morphosyntax, and discourse, particularly in older children with SLI (Reilly, Losh, Bellugi, & Wulfeck, 2004). Moreover, a narrative task may lead to more and different errors, because children are prompted to express semantic relations that may be difficult for them and that they may avoid in spontaneous speech. For this reason, a narrative task may more effectively reveal the relative strengths and weaknesses in the language performance of children with SLI (Wetherell, Botting, & Conti-Ramsden, 2007).

An investigation of morphosyntactic accuracy and grammatical complexity can reveal which types of grammatical errors and which difficulties in complex syntax are most prominent for Dutch-speaking children with SLI. Therefore, the following questions will be addressed:

1. To what extent do patterns of development in morphosyntactic accuracy and grammatical complexity of children with SLI differ from those observed in CA and LA control groups?
2. To what extent are morphosyntactic accuracy (i.e., verb-related and non–verb-related errors) and syntactic complexity in children with SLI equally affected?

Method

Participants

Participants were 30 monolingual Dutch children with SLI, age 6;5 at first measurement (T1); 30 CA control children, age 6;6 at T1; and 30 LA control children, age 4;7 at T1. Informed consent was obtained from all parents. The initial selection criterion for the LA group was age; the LA children were on average 2 years younger than the children with SLI and had no history of speech or language impairment. The 2-year difference was motivated by the fact that numerous other studies on SLI have used LA control groups that were 2 years younger than the SLI groups. Moreover, studies that matched SLI and LA groups on mean length of utterance also showed that the LA group was 2 years younger than the SLI group (Paradis, 2010).

The data from the SLI group originated from a previous study by Van Weerdenburg, Verhoeven, and van Balkom (2006). The children with SLI (23 boys, 7 girls) were diagnosed by a multidisciplinary team of specialists consisting of a physician, a psychologist, an educationalist, and a speech therapist. Diagnosis was based on standardized language tests, and it was established that the language impairment was not the direct result of global intellectual, sensory, motor, emotional, or physical impairments. The children with SLI were all enrolled in special education for children with severe speech and language impairments.

The children in the LA group (18 boys, 12 girls) and the CA group (16 boys, 14 girls) were recruited from four different primary schools in the central part of the Netherlands. As can be seen in Table 1, all children had nonverbal IQ scores within the normal range, as determined with the Raven's Coloured Progressive Matrices (Raven CPM; van Bon, 1986). One-way analysis of variance (ANOVA) showed that there were no significant differences between the groups on the Raven CPM, $F(2, 87) = 0.57$, $p = .569$. The children's language proficiency was evaluated using the Taaltoets Alle Kinderen (Dutch Language Proficiency Test for All Children [LPT]; Verhoeven & Vermeer, 2001), which is a standardized discrete-point test for the assessment of 4- to 10-year-old children consisting of 10 subtests. All of the subtests have been shown to be reliable, with Cronbach's alphas ranging between $\alpha = .90$ and $\alpha = .97$. Norm scores for Dutch-speaking children were based on a nationwide sample of 727 children (Verhoeven & Vermeer, 2006).

One-way ANOVAs yielded significant differences between groups for all LPT subtests: receptive vocabulary: $F(2, 87) = 40.80$, $p < .001$; sentence comprehension 1: $F(2, 87) = 14.19$, $p < .001$; sentence comprehension 2: $F(2, 87) = 31.55$, $p < .001$; morphology: $F(2, 87) = 11.99$, $p < .001$. Post hoc tests revealed that the CA group performed significantly higher on all LPT subtests compared with both the SLI and the LA groups ($p \leq .001$ for all comparisons). There were no significant differences between the SLI group and the LA group on any of the LPT subtests ($p > .05$ for all comparisons).

Table 1. Age (*SD*, in months) and scores on Raven Coloured Progressive Matrices (CPM) and Dutch Language Proficiency Test for All Children (LPT) subtests in the specific language impaired (SLI), language age–matched (LA), and chronological age–matched (CA) groups (*n* = 30 in each group) at first measurement.

Measure	Max. score[a]	SLI *M (SD)*	LA *M (SD)*	CA *M (SD)*
Age (years;months)		6;5 (1.5)	4;7 (2.6)	6;6 (2.4)
Raven CPM[b]		5.92 (1.92)	5.89 (1.50)	6.32 (1.79)
LPT receptive vocabulary[c]	96	56.93 (13.40)	53.63 (11.70)	77.33 (6.85)
LPT sentence comprehension 1[c]	42	32.67 (4.48)	31.07 (5.73)	36.90 (2.17)
LPT sentence comprehension 2[c]	42	30.80 (4.15)	28.27 (3.71)	35.40 (2.51)
LPT morphology[c]	24	11.03 (4.49)	12.77 (3.00)	16.57 (5.56)

Note. LPT sentence comprehension 1 assesses understanding of function words; LPT sentence comprehension 2 assesses understanding of syntactic patterns, and LPT morphology measures production of noun plurals and past participles.

[a]Maximum raw score. [b]Standard score (–1 *SD* to +1 *SD*, range = 3.0–7.0). [c]Raw score.

Materials

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

Procedure

All participants were tested at three points in time separated by 12-month intervals, during a period of 2 years. The narratives were orthographically transcribed in accordance

Figure 1. Dutch Language Proficiency Test for All Children (LPT) storytelling task 1. © 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.

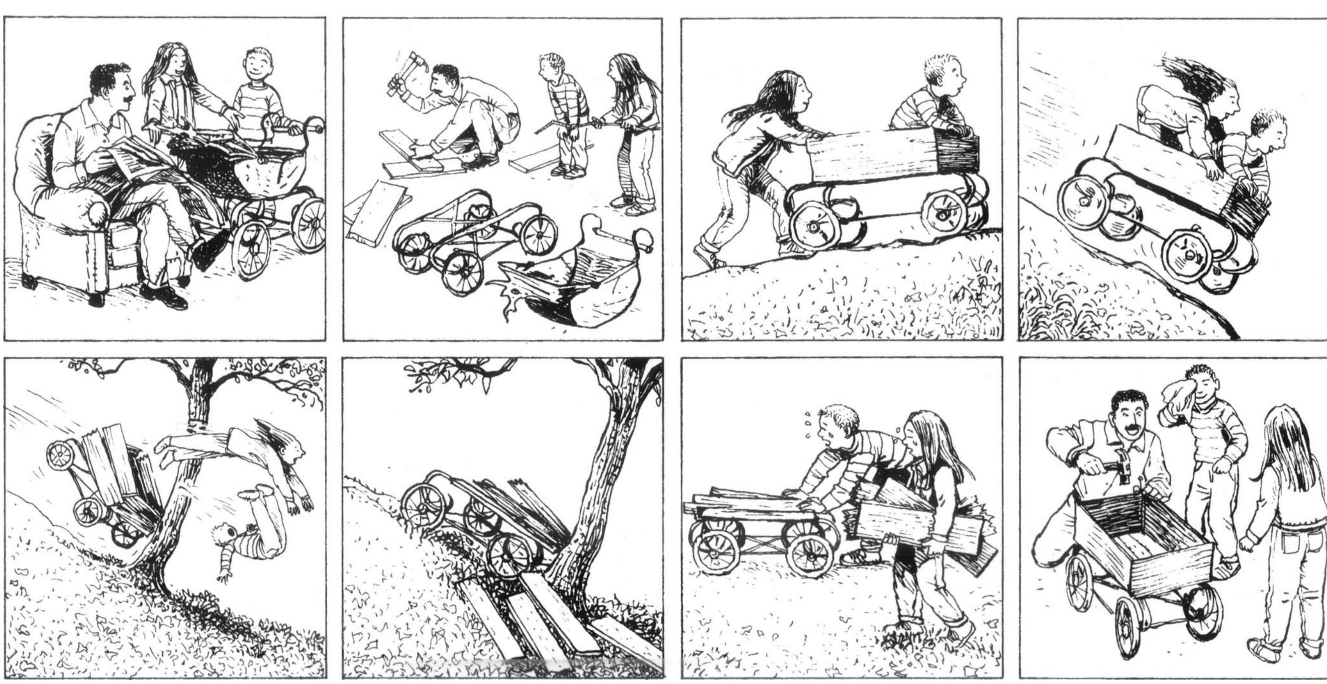
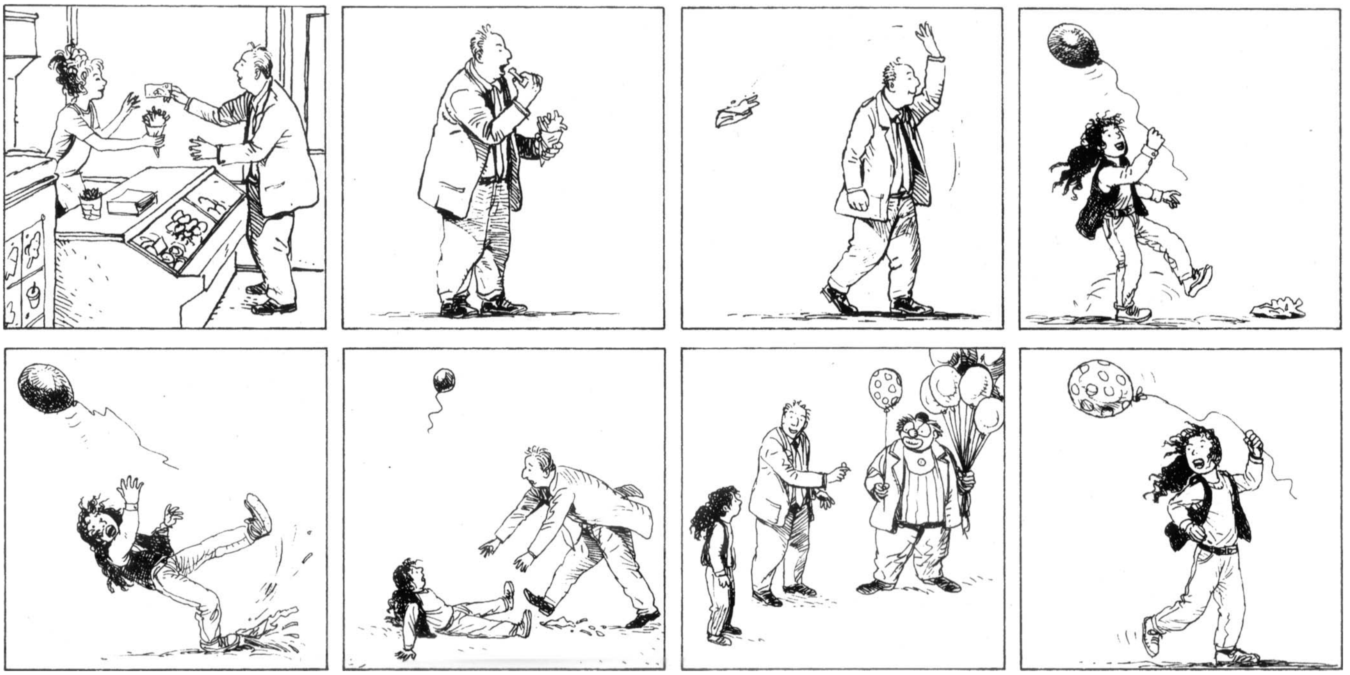


Figure 2. Dutch Language Proficiency Test for All Children (LPT) Storytelling task 2. © 2001 by Cito, Arnhem, the Netherlands. Reprinted with permission.



with Codes for the Human Analysis of Transcripts conventions, the coding system of the Child Language Data Exchange System (MacWhinney, 2000). Each transcript contained storytelling tasks 1 and 2 told in succession in this fixed order. The basic unit of analysis was the *T-unit*, defined as a single main clause plus any subordinate clause or nonclausal structure that is attached to it or embedded in it (Hunt, 1970). Coordinate clauses were transcribed and counted as separate utterances, unless there was conjunction reduction (e.g., *The man buys a new balloon and gives it to the girl.*). Sentences with quoted speech, in which the quotation forms a full clause containing a subject and a verb, were also transcribed as separate utterances (e.g., *The children ask their father: "Can you repair our go-kart again?"*). In addition to the transcriptions on the main tier, additional dependent tiers were created in the Codes for the Human Analysis of Transcripts files to code grammatical complexity and morphosyntactic errors.

The reliability of the transcriptions was checked by retranscribing 10% of the files of each group by either trained speech-language pathologists (nine transcripts of the children with SLI) or trained master students in linguistics (nine transcripts of each control group). The point-to-point reliability at word level reached 90%. Coding agreement reached 96% for grammatical complexity and 93% for morphosyntactic accuracy.

Analysis

Both morphosyntactic accuracy and grammatical complexity were investigated. The operationalizations of the dependent measures are presented in Table 2. Because the narratives varied in length, percentages were computed for most variables, in order to make the dependent measures comparable across children and time points. Verb-related morphosyntactic errors were regarded as indicators of SLI and thus acted as benchmarks to which to relate the development of non–verb-related morphosyntactic errors as well as the development of grammatical complexity.

First of all, two general measures related to narrative length, namely, number of T-units and total number of words (TNW), were computed. These two measures can be regarded to reflect general narrative skills. Second, morphosyntactic accuracy in the narratives was examined. As an overall measure of grammatical accuracy, the percentage of correct T-units (i.e., error-free T-units) was used. However, a T-unit can contain several grammatical errors, and, therefore, all individual grammatical errors in the narratives were tallied as well. In order to examine possible changes in error patterns over time, the grammatical errors were further specified into a number of different error types, which were subsequently arranged into two broad categories: (a) percentage of verb-related errors and (b) percentage of non–verb-related errors. The first composite measure, percentage of verb-related errors, contained all observed errors in verb morphology (e.g., subject–verb agreement, tense, auxiliaries, and participles), verb placement, and verb argument structure (see Appendix). The second composite measure, percentage of non–verb-related errors, captured all remaining grammatical errors. This composite measure was not further subdivided, because many different error types were counted, and percentages of individual error types

Table 2. Operationalizations of grammatical accuracy and grammatical complexity.

Measure	Definition (errors are underlined)
Grammatical accuracy	
T-units correct	Number of error-free T-units/total number of T-units.
Verb-related errors	Sum of all errors related to verbs/number of clauses (examples for the main verb-related error categories are presented below).
Verb morphology errors	Subject–verb agreement: **hij loop* [*he walk], correct: *hij loopt* [he walks]. Verb-second placement: **dan de man komt*$_{3Sg}$ [then the man comes], correct: *dan komt*$_{3Sg}$ *de man* [then comes the man]. Tense (present or past tense adverb with an incorrectly tensed verb): **toen*$_{PAST}$ *valt*$_{PRESENT}$ *het meisje* [*then falls the girl], correct: *toen*$_{PAST}$ *viel*$_{PAST}$ *het meisje* [then fell the girl]. Root infinitives: **hij de ballon geven* [*he the balloon give], correct: *hij geeft de ballon* [he gives the balloon]. Past tense verb overregularizations: **hij brengde* [*he bringed], correct: *hij bracht* [he brought]. Omissions (Ø) and substitutions of auxiliaries with a past participle: **toen Ø ze naar de clown gelopen* [*then Ø they to the clown walked], correct: *toen waren*$_{AUX}$ *ze naar de clown gelopen* [then they had walked to the clown]. Past participle (deletion or substitution of a prefix and/or suffix): **hij heeft het meisje pak* [*he has the girl take], correct: *hij heeft het meisje gepakt* [he has taken the girl].
Verb argument structure	Subject and object omissions/number of clauses in which a subject or object was expected and obligatory (instances of allowed subject drop or object drop were not counted as errors).
Non–verb-related errors	Sum of all non–verb-related errors/number of clauses: all errors in word order, deletion of nouns, determiners (substitutions and omissions), prepositions, pronouns (case, gender, and number), conjunctions, omission of pronominal/adverbial *'er'* [there], and adjectival inflection.
Dummy auxiliaries	Number of dummy auxiliaries/number of clauses.
Grammatical complexity	
Mean length of T-units	Mean length of T-units in words (all retracings were excluded).
Complex sentences	Total number of complex sentences (e.g., sum of all sentences containing subordinate, infinitival, and reduced clauses; coordination with reduction and direct speech)/total number of T-units.
Subordinate clauses	Total number of subordinate clauses/total number of T-units.
Relative clauses	Total number of relative clauses/total number of T-units.

could therefore be too low for a meaningful quantitative analysis. Because a prolonged and frequent use of dummy verbs can be argued to reflect an immature stage of verb morphology mastery, the percentages of dummy auxiliaries were included in the analysis. Using a dummy auxiliary does not render a sentence ungrammatical; therefore, counts of dummy auxiliaries were not included in the grammatical accuracy measures. The final step in the analysis was the computation of four measures of grammatical complexity. (a) Mean length of T-units (MLTU) in words was used as a general measure of grammatical complexity. Grammatical complexity was further operationalized by the composite measure (b) percentage of complex sentences. All sentences with subordinate clauses, coordinated sentences with conjunction reduction, direct speech, and infinitival clauses were counted as complex sentences. Subordinate clauses included all forms of adverbial, nominal, and relative clauses. Complex sentences were subdivided further by computing (c) percentages of subordinate clauses and (d) percentages of relative clauses. This subdivision was motivated by the fact that subordinate clauses, and relative clauses in particular, are known to pose difficulties for school-age children with SLI (for an overview, see Jensen de López, Sundahl Olsen, & Chondrogianni, 2014).

Mixed-model ANOVAs were used to examine differences between groups and at different time points using time (T1, T2, and T3) as the within-subject factor and group (SLI, LA, and CA) as the between-subjects factor. Significance level was set at .05. In order to analyze differences between groups, subsequent one-way ANOVAs with post hoc Bonferroni correction were used. Generalized linear modeling repeated measures (RM) ANOVA was used to test differences over time within the separate groups. The assumption of sphericity was checked for all variables with Mauchly's test of sphericity. Whenever this assumption was violated, the Greenhouse–Geisser corrected values were reported.

Results

The chronological ages of the three groups, the numbers of T-units, and the total numbers of words used by the three groups in the narratives at the three measurement points are presented in Table 3. The numbers of T-units give an indication of story length, and total number of words is a broad measure of lexical diversity. These background measures of the narratives are necessary for a comparison of narratives of TD children and children with SLI.

Background Measures

Number of T-Units

For number of T-units, no significant interaction was found between time and group, $F(4, 174) = 1.23$, $p = .301$, $\eta_p^2 = .027$. However, there was a significant main effect of group, $F(2, 87) = 16.18$, $p < .001$, $\eta_p^2 = .271$, as well as of time, $F(2, 174) = 16.48$, $p < .001$, $\eta_p^2 = .159$. One-way ANOVAs yielded significant differences between groups at T1, $F(2, 87) = 3.27$, $p = .043$; T2, $F(2, 87) = 7.52$, $p = .001$;

Table 3. Age (*SD*, in months), means, standard deviations (*SD*), and ranges of numbers of T-units and total number of words in the narratives.

	SLI			LA			CA		
Measure	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
Age (years;months)									
M (*SD*)	6;5 (1.5)	7;4 (2.1)	8;5 (2.0)	4;7 (2.6)	5;7 (2.6)	6;7 (2.6)	6;6 (2.4)	7;6 (2.4)	8;6 (2.4)
Number of T-units									
M (*SD*)	22.10 (6.89)	25.33 (6.54)	28.03 (6.78)	18.50 (4.86)	20.37 (3.48)	20.73 (3.58)	20.37 (4.26)	24.47 (5.42)	25.00 (6.90)
Range	7–35	18–43	19–51	9–34	15–28	15–29	11–29	16–35	14–41
Total number of words									
M (*SD*)	129.27 (42.46)	151.73 (40.74)	177.47 (49.36)	108.20 (34.60)	128.5 (33.80)	135.39 (25.65)	132.27 (37.83)	164.13 (40.50)	172.50 (52.45)
Range	48–203	88–252	102–329	53–182	79–193	85–184	65–248	107–273	93–285

and T3, $F(2, 87) = 11.52$, $p < .001$. Post hoc tests revealed that the SLI group produced more T-units than the LA group at T1 ($p = .037$), T2 ($p = .001$). and T3 ($p = < .001$). The CA group outperformed the LA group at T2 ($p = .011$) and T3 ($p = .019$). Between the SLI and CA groups, no differences were found. RM ANOVA for the separate groups revealed that in the SLI group, the number of T-units increased significantly between T1 and T2, $F(1, 29) = 4.43$, $p = .046$, $\eta_p^2 = .130$, and between T1 and T3, $F(1, 29) = 12.50$, $p = .001$, $\eta_p^2 = .301$. In the LA control group, the number of T-units increased between T1 and T3, $F(1, 29) = 4.51$, $p < .042$, $\eta_p^2 = .135$. In the CA group, number of T-units increased significantly between T1 and T2, $F(1, 29) = 11.28$, $p = .002$, $\eta_p^2 = .280$, and between T1 and T3, $F(1, 29) = 10.75$, $p = .003$, $\eta_p^2 = .270$.

Total Number of Words

For total number of words (TNW), no significant interaction was found between time and group, $F(4, 174) = 1.06$, $p = .377$, $\eta_p^2 = .024$. There was a significant main effect of group, $F(2, 87) = 11.58$, $p < .001$, $\eta_p^2 = .210$, as well as of time, $F(2, 174) = 27.85$, $p < .001$, $\eta_p^2 = .242$. One-way ANOVAs returned significant differences between groups at T1, $F(2, 87) = 3.49$, $p = .035$; T2, $F(2, 87) = 6.60$, $p = .002$; and T3, $F(2, 87) = 8.18$, $p = .001$. Post hoc tests yielded no significant differences between the SLI and CA groups at all time points. At T2, the LA group used fewer words than the CA group ($p = .002$). At T3, both the SLI and the CA groups used more words than the LA group ($p = .001$ and $p = .005$, respectively). RM ANOVA for the separate groups revealed that in the SLI group, TNW increased significantly between T1 and T2, $F(1, 29) = 4.96$, $p = .034$, $\eta_p^2 = .146$; T2 and T3, $F(1, 29) = 7.29$, $p = .011$, $\eta_p^2 = .201$; and T1 and T3, $F(1, 29) = 17.35$, $p < .001$, $\eta_p^2 = .347$. The LA control group showed a significant increase in TNW between T1 and T2, $F(1, 29) = 7.98$, $p < .008$, $\eta_p^2 = .216$, and T1 and T3, $F(1, 29) = 13.61$, $p = .001$, $\eta_p^2 = .319$. In the CA group, TNW increased between T1 and T2, $F(1, 29) = 11.78$, $p = .002$, $\eta_p^2 = .289$, and T1 and T3, $F(1, 29) = 15.80$, $p < .001$, $\eta_p^2 = .358$.

Grammatical Accuracy and the Use of Dummy Auxiliaries

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

T-Units Correct

The percentages of T-units correct (i.e., counts of error-free T-units divided by the sum of T-units) were regarded as a general measure of grammatical accuracy. There was no significant interaction between group and time, $F(3.39, 147.23) = 0.99$, $p = .405$, $\eta_p^2 = .022$. However, there were significant main effects of time, $F(1.69, 147.23) = 43.97$, $p < .001$, $\eta_p^2 = .336$, and group, $F(2, 87) = 109.92$, $p < .001$, $\eta_p^2 = .716$. Figure 3a depicts the development in the groups over time. One-way ANOVAs returned significant differences between groups at T1, $F(2, 87) = 48.59$, $p < .001$; T2, $F(2, 87) = 73.76$, $p < .001$; and T3, $F(2, 87) = 86.95$, $p < .001$. Post hoc tests revealed that at all time points, the SLI group had fewer correct T-units than the LA and CA control groups ($p < .001$ for all comparisons). The LA group had fewer correct T-units than the CA group at T1 ($p = .009$) and T2 ($p = .001$). At T3, this difference did not reach significance ($p = .093$). RM ANOVAs indicated that percentages of T-units correct increased significantly in the SLI group between T1 and T2, $F(1, 29) = 8.68$, $p = .006$, $\eta_p^2 = .230$, and between T2 and T3, $F(1, 29) = 7.16$, $p = .012$, $\eta_p^2 = .198$. In the LA group, percentages of T-units correct increased significantly between T1 and T2, $F(1, 29) = 7.65$, $p = .010$, $\eta_p^2 = .209$, and between T2 and T3, $F(1, 29) = 15.30$, $p = .001$, $\eta_p^2 = .345$. In the CA group, the difference between T1 and T2 was significant, $F(1, 29) = 14.38$, $p < .001$, $\eta_p^2 = .332$, but did not reach significance between T2 and T3, $F(1, 29) = 0.99$, $p = .328$, $\eta_p^2 = .033$.

Verb-Related Errors

Development of verb-related errors (i.e., all verb-related errors divided by the total number of T-units) is presented in Figure 3b. There was a significant interaction between group and time, $F(3.10, 134.76) = 3.81$, $p = .011$, $\eta_p^2 = .081$. The main effects of group and time were significant as well, group: $F(2, 87) = 57.49$, $p < .001$, $\eta_p^2 = .569$; time: $F(1.55, 134.76) = 29.57$, $p < .001$, $\eta_p^2 = .081$. One-way ANOVAs yielded significant differences between groups at T1, $F(2, 87) = 26.63$, $p < .001$; T2, $F(2, 87) = 37.74$, $p < .001$; and T3, $F(2, 87) = 48.59$, $p < .001$. Post hoc tests revealed that at all time points, the SLI group made more verb-related errors than both control groups ($p < .001$ for all comparisons). At T1, the difference between LA and CA groups was not significant ($p = .115$). At T2, the difference between LA and CA groups did reach significance ($p = .022$), but at T3, this difference was not significant ($p = 1.00$). RM ANOVAs for the separate groups revealed that in the SLI group, verb-related errors decreased between T1 and T2, $F(1, 29) = 20.00$, $p < .001$, $\eta_p^2 = .408$), but not between T2 and T3, $F(1, 29) = 3.03$, $p = .092$, $\eta_p^2 = .095$. In the LA group, verb-related errors did not decrease between T1 and T2, $F(1, 29) = 4.02$, $p = .054$, $\eta_p^2 = .012$, but a significant decrease was found between T2 and T3, $F(1, 29) = 13.82$, $p = .001$, $\eta_p^2 = .323$. In the CA group, the decrease was significant between T1 and T2, $F(1, 29) = 11.30$, $p = .002$, $\eta_p^2 = .280$, but not between T2 and T3, $F(1, 29) = 0.56$, $p = .461$, $\eta_p^2 = .019$.

Non–Verb-Related Errors

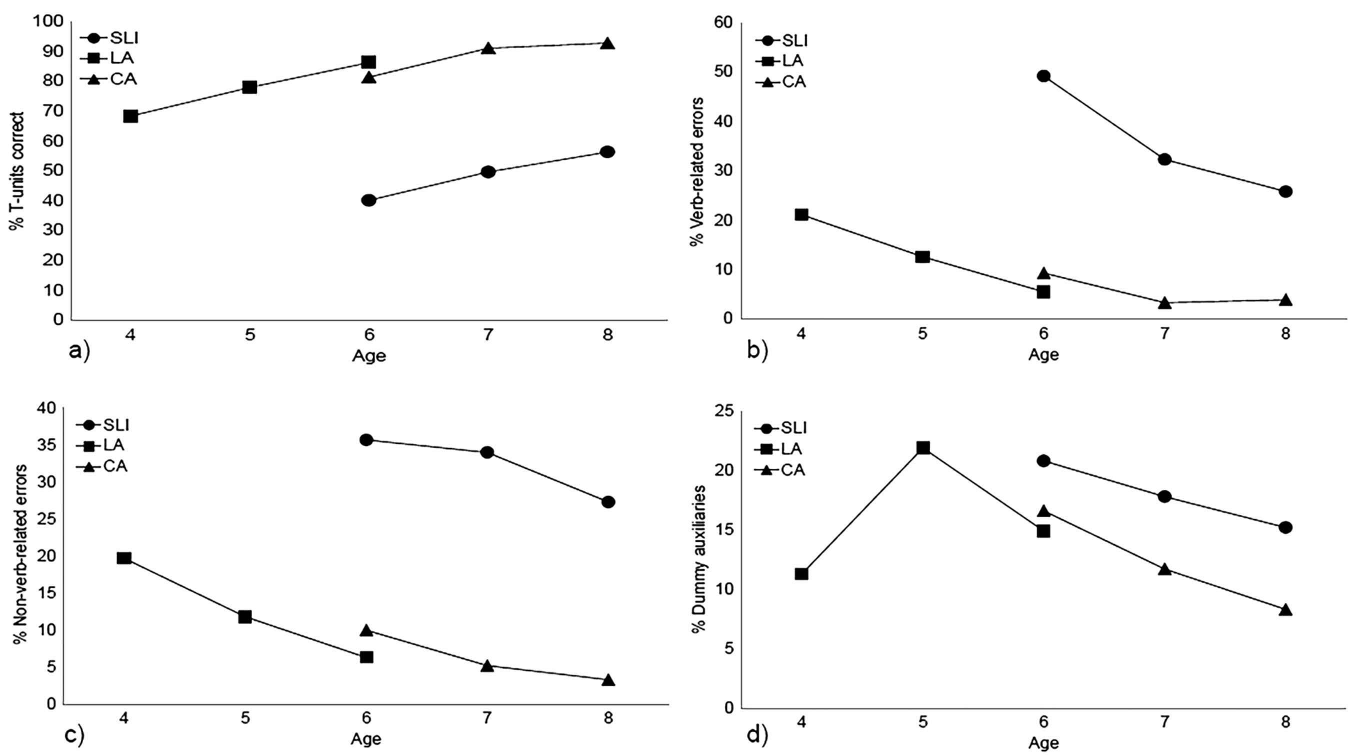
For this measure (see Figure 3c), the sum of all non–verb-related errors was divided by the total number of T-units. There was no significant interaction between group and time, $F(4, 174) = 1.75$, $p = .142$, $\eta_p^2 = .039$. However, there was a significant main effect of group, $F(2, 87) = 101.84$, $p < .001$, $\eta_p^2 = .701$, as well as of time, $F(2, 174) = 25.11$, $p < .001$, $\eta_p^2 = .224$. One-way ANOVAs revealed significant differences between groups at T1, $F(2, 87) = 39.70$, $p < .001$; at T2, $F(2, 87) = 53.94$, $p < .001$; and at T3,

Table 4. Grammatical accuracy measures: mean percentages and standard deviations (*SD*), total raw numbers/total number of T-units, and ranges of numbers of correct T-units and raw numbers of verb-related errors, non–verb-related errors, and dummy auxiliaries per group at the three time points.

	SLI			LA			CA		
Measure	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
T-units correct									
M % (*SD*)	40.1 (16.1)	49.6 (18.0)	56.3 (14.7)	68.3 (19.8)	77.9 (13.2)	86.3 (10.2)	81.4 (13.1)	91.1 (7.1)	92.8 (8.6)
Raw/T-units	272/663	378/760	483/841	383/555	480/611	358/662	497/611	667/734	692/750
Range	1–23	1–25	5–31	2–26	6–25	10–25	11–28	14–31	14–38
Verb-related errors									
M % (*SD*)	49.2 (29.8)	32.3 (20.0)	25.8 (14.6)	21.1 (21.3)	12.6 (10.4)	5.5 (6.2)	9.3 (8.8)	3.3 (3.9)	3.9 (4.9)
Raw/T-units	322/663	253/760	211/841	109/555	75/611	33/622	55/611	24/734	30/750
Range	1–33	0–24	0–18	0–12	0–8	0–3	0–7	0–4	0–5
Non–verb-related errors									
M % (*SD*)	35.7 (14.1)	34.0 (14.8)	27.3 (11.0)	19.7 (10.4)	11.8 (11.2)	6.3 (8.0)	10.0 (8.6)	5.2 (5.9)	3.3 (5.8)
Raw/T-units	235/663	252/760	220/841	106/555	69/611	38/622	61/611	39/734	27/750
Range	1–21	3–19	3–12	1–7	0–8	0–6	0–5	0–8	0–7
Dummy auxiliaries									
M % (*SD*)	20.8 (14.4)	17.8 (11.4)	15.2 (12.5)	11.3 (11.4)	21.9 (16.2)	14.9 (8.9)	16.6 (13.4)	11.7 (10.9)	8.3 (8.2)
Raw/clauses	145/679	138/791	132/841	65/571	134/628	94/643	105/637	83/782	62/806
Range	0–13	0–12	0–12	0–8	0–11	0–8	0–12	0–14	0–5

$F(2, 87) = 70.92$, $p < .001$. Post hoc tests revealed that at all time points, the SLI group produced more non–verb-related errors than both control groups ($p < .001$ for all comparisons). At T1, the LA group had more errors than the CA group ($p = .004$). At T2 and T3, differences between LA and CA children were not significant ($p = .077$ and $p = .541$, respectively). RM ANOVAs for the separate groups revealed that percentages of non–verb-related errors decreased significantly in the SLI group between T2 and T3, $F(1, 29) = 6.32$, $p = .018$, $\eta_p^2 = .179$, but not between T1 and T2,

Figure 3. Development of markers of grammatical accuracy at ages 4 to 8 years in the three groups, specific language impaired (SLI), language age–matched (LA), and chronological age–matched (CA): (a) percentages of T-units correct, (b) percentages of total verb-related errors, (c) percentages of non–verb-related errors, and (d) percentages of dummy auxiliaries.



$F(1, 29) = 0.23$, $p = .633$, $\eta_p^2 = .008$. In the LA group, percentages of non–verb-related errors decreased significantly between T1 and T2, $F(1, 29) = 12.49$, $p = .001$, $\eta_p^2 = .301$, and between T2 and T3, $F(1, 29) = 9.44$, $p = .005$, $\eta_p^2 = .246$. In the CA control group, the decrease was significant between T1 and T2, $F(1, 29) = 9.76$, $p = .004$, $\eta_p^2 = .252$, but not between T2 and T3, $F(1, 29) = 2.05$, $p = .163$, $\eta_p^2 = .066$.

Dummy Auxiliaries

To compute percentages of dummy auxiliaries, total numbers of dummy auxiliaries were divided by the total number of clauses. Developmental trajectories are presented in Figure 3d. There was a significant interaction between group and time, $F(4, 174) = 5.82$, $p < .001$, $\eta_p^2 = .118$; a significant main effect of group, $F(2, 87) = 3.17$, $p = .047$, $\eta_p^2 = .068$; and a significant main effect of time, $F(2, 174) = 4.76$, $p = .010$, $\eta_p^2 = .052$. One-way ANOVAs yielded significant differences between groups at T1, $F(2, 87) = 4.00$, $p = .022$; at T2, $F(2, 87) = 4.66$, $p = .012$; and at T3, $F(2, 87) = 4.55$, $p = .013$. Post hoc tests revealed that at T1, the SLI group used more dummy auxiliaries than the LA group ($p = .018$). The difference between the SLI and the CA groups was not significant ($p = .664$). At T2, differences between the SLI group and the control groups were not significant. However, the LA group used more dummy auxiliaries than the CA group ($p = .009$). At T3, the SLI group used more dummy auxiliaries than the CA group ($p = .028$), and the LA group also used more dummies than the CA group ($p = .037$) but did not differ from the SLI group. RM ANOVAs for the separate groups revealed that the use of dummy auxiliaries did not change significantly in the SLI group over time. In the LA group, numbers of dummy verbs increased significantly between T1 and T2, $F(1, 29) = 16.07$, $p < .001$, $\eta_p^2 = .356$, followed by a significant decrease between T2 and T3, $F(1, 29) = 5.99$, $p = .021$, $\eta_p^2 = .171$. In the CA group, the use of dummy verbs only decreased significantly between T1 and T3, $F(1, 29) = 10.33$, $p = .003$, $\eta_p^2 = .263$. Inspection of the data revealed that at T2, three LA children were responsible for the spike in use of dummy auxiliaries. These three children produced 11 dummy auxiliaries each (in 17, 19, and 21 clauses, respectively). In comparison, at T2, two children with SLI produced more than 10 dummies each (11/31 and 12/42), and one child in the CA group produced a high number (14/31).

Summary of Grammatical Accuracy Analysis and Dummy Auxiliaries

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

Grammatical Complexity

In order to relate the development of morphosyntactic accuracy to complex syntax, grammatical complexity in the narratives of the three groups at the three time points was investigated. The descriptives of the grammatical complexity measures at the different measurement points are presented in Table 5.

MLTU

The development in MLTU in the groups is presented in Figure 4a. For MLTU, no significant interaction was found between time and group, $F(4, 174) = 1.87$, $p = .118$, $\eta_p^2 = .041$. However, there was a significant main effect of group, $F(2, 87) = 9.66$, $p < .001$, $\eta_p^2 = .182$, and of time, $F(2, 174) = 24.96$, $p < .001$, $\eta_p^2 = .223$. One-way ANOVAs yielded significant differences between groups at T1, $F(2, 87) = 7.99$, $p = .001$; T2, $F(2, 87) = 4.26$, $p = .017$; and T3, $F(2, 87) = 6.97$, $p = .002$. Post hoc tests revealed that the SLI group had a significantly lower MLTU than the CA group at T1 ($p = .015$), at T2 ($p = .014$), and at T3 ($p < .001$). The SLI and the LA groups did not differ significantly at any of the time points. The LA group had a lower MLTU than the CA group at T1 ($p = .001$). However, this difference was not significant at T2 and T3. RM ANOVA for the separate groups revealed that in the SLI group, MLTU increased significantly only between T1 and T3, $F(1, 29) = 9.55$, $p = .004$, $\eta_p^2 = .248$. The LA control group showed a significant increase in MLTU between T1 and T2, $F(1, 29) = 19.91$, $p < .001$, $\eta_p^2 = .407$, but no reliable difference between T2 and T3, $F(1, 29) = 3.97$, $p = .056$, $\eta_p^2 = .120$. Between T1 and T3, MLTU increased significantly, $F(1, 29) = 31.74$, $p < .001$, $\eta_p^2 = .523$. In the CA group, MLTU increased significantly between T1 and T3, $F(1, 29) = 9.13$, $p < .005$, $\eta_p^2 = .239$.

Complex Sentences

This composite measure represents the number of complex sentences divided by the number of T-units in the narratives. No significant interaction was found between time and group, $F(4, 174) = 0.13$, $p = .971$, $\eta_p^2 = .003$. However, there was a significant main effect of time, $F(2, 174) = 12.91$, $p < .001$, $\eta_p^2 = .129$, and of group, $F(2, 87) = 5.70$, $p = .005$, $\eta_p^2 = .116$. The use of complex sentences increased steadily over time in all three groups, as can be seen in Figure 4b. One-way ANOVAs returned only one significant difference between groups: at T3, $F(2, 87) = 3.27$,

Table 5. Grammatical complexity: means and standard deviations (*SD*) of mean length of T-units, percentages of complex sentences, subordinate clauses, and relative clauses with total raw numbers/total number of T-units, and ranges of raw numbers per group at the three time points.

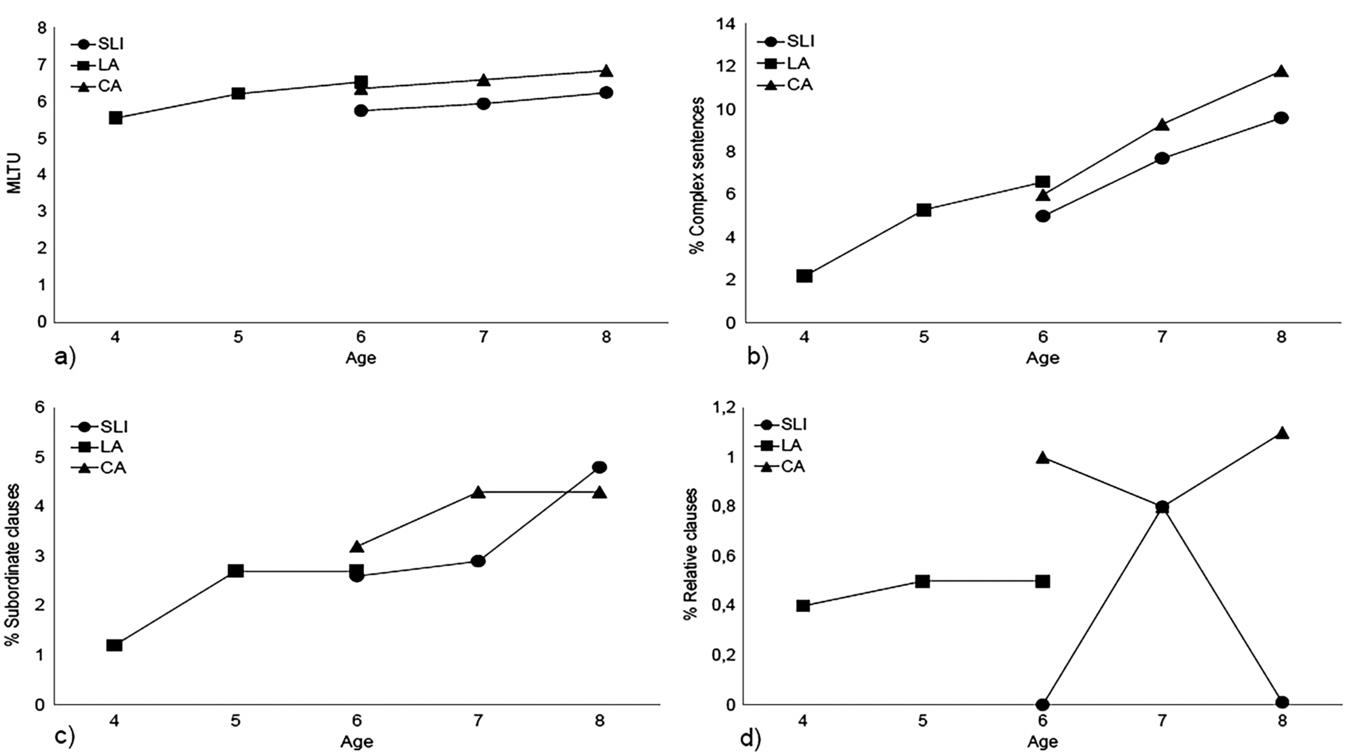
	SLI			LA			CA		
Measure	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
Mean length of T-units in words									
M (*SD*)	5.75 (0.77)	5.94 (0.80)	6.24 (0.60)	5.55 (0.85)	6.21 (0.89)	6.53 (0.53)	6.35 (0.78)	6.59 (0.86)	6.84 (0.74)
Complex sentences									
M % (*SD*)	5.0 (6.9)	7.7 (6.7)	9.6 (7.8)	2.2 (4.7)	5.3 (6.7)	6.6 (5.9)	6.0 (6.8)	9.3 (9.3)	11.8 (9.7)
Raw/T-units	34/663	60/760	80/841	13/555	34/611	43/622	41/611	73/734	95/750
Range	0–7	0–6	0–9	0–4	0–5	0–4	0–9	0–10	0–9
Subordinate clauses									
M % (*SD*)	2.6 (4.6)	2.9 (4.0)	4.8 (6.8)	1.2 (3.4)	2.7 (3.8)	2.7 (4.0)	3.2 (4.2)	4.3 (5.5)	4.3 (5.5)
Raw/T-units	16/663	22/760	40/841	8/555	17/611	17/622	22/611	33/734	35/750
Range	0–5	0–3	0–9	0–4	0–3	0–3	0–4	0–6	0–5
Relative clauses									
M % (*SD*)	0.0 (0.0)	0.8 (2.0)	0.1 (0.5)	0.4 (1.4)	0.5 (1.4)	0.5 (1.6)	1.0 (2.6)	0.8 (1.7)	1.1 (2.0)
Raw/T-units	0/663	5/760	1/841	2/555	3/611	3/622	7/611	7/734	10/750
Range	0–0	0–2	0–1	0–1	0–1	0–1	0–2	0–2	0–2

$p = .043$. Post hoc testing showed that the CA group used more complex sentences at T3 than the LA group ($p = .038$). The children with SLI did not differ from either control group. RM ANOVAs for the separate groups revealed that differences between T1–T2 and T2–T3 were not significant, for all three groups. However, T1–T3 differences were significant in all groups, for SLI, $F(1, 29) = 7.34$, $p = .011$, $\eta_p^2 = .202$; for LA, $F(1, 29) = 8.66$, $p = .006$, $\eta_p^2 = .230$; and for CA, $F(1, 29) = 8.49$, $p = .007$, $\eta_p^2 = .227$.

Subordinate Clauses

The total number of subordinate clauses was divided by the number of T-units to compute the percentages of subordinate clauses. For this measure, no interaction between time and group was found, $F(4, 174) = 0.48$, $p = .751$,

Figure 4. Development of markers of grammatical complexity at ages 4 to 8 years in the three groups, specific language impaired (SLI), language age–matched (LA), and chronological age–matched (CA): (a) mean length of T-units (MLTU), (b) percentages of complex sentences, (c) percentages of subordinate clauses, and (d) percentages of relative clauses.



partial η^2 = .011. The main effect of time did not reach significance, $F(2, 174) = 3.00$, $p = .053$, $\eta_p^2 = .033$. The main effect of group also was not significant, $F(2, 87) = 2.44$, $p = .093$, $\eta_p^2 = .053$. Figure 4c illustrates the developmental trajectories in the three groups.

Relative Clauses

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

Summary of Grammatical Complexity

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

Discussion

This study set out to identify developmental patterns in morphosyntactic accuracy and grammatical complexity in Dutch school-age children with SLI. Our main aim was to see whether the disorder impacts different aspects of grammar equally or in different degrees. To this end, we examined error rates in verb-related morphosyntax, error rates in non–verb-related morphosyntax, and the development of complex syntax. A defect of verb-related morphosyntax (i.e., high error rate) was expected; the question was whether this defect also surfaced in non–verb-related morphosyntax and complex syntax. Our first observation is that comparisons of SLI and TD groups using a narrative task yield a different picture than analysis using standardized language tests. Scores on all subtests of the LPT at T1 indicated that the SLI group lagged 2 years behind the CA group and performed just like the LA control group. Thus, on the basis of the LPT scores alone, one would be inclined to label the SLI group as delayed by 2 years. However, our elaborate analysis of the narratives provides a more differentiated picture. A second observation is that the children with SLI did not differ from the CA control group on the production of number of T-units and total number of words. The SLI group also produced longer narratives when compared with the LA group at two time points. The narratives of the children with SLI in our study are, at least quantitatively, comparable to those of the TD groups. This finding is in contrast with studies on narratives of English-speaking children with SLI (Newman & McGregor, 2006; Reilly et al., 2004). Results on morphosyntactic accuracy and complex syntax measures are discussed next and related to the developmental patterns described in the literature.

The analysis of grammatical accuracy yielded large differences between the children with SLI and the TD control groups. On all morphosyntactic accuracy measures, namely, T-units correct, verb-related errors, and non–verb-related errors, the SLI group performed much more poorly than both control groups. The results of the SLI group on MLTU and all standardized LPT subtests would imply a plain 2-year developmental delay, as the children with SLI did not differ from the 2-years-younger LA group on these measures. However, compared with the scores on MLTU and the LPT subtests, the morphosyntactic errors were unexpectedly high. Verb-related errors and non–verb-related errors roughly followed the same developmental trajectories, and error percentages were approximately comparable. Morphosyntactic development in the SLI group would best fit a delay-within-delay pattern (Rice, 2003) or, in Leonard's (1998) terms, an uneven profile.

In contrast, for grammatical complexity, almost no differences were found between the SLI group and the LA or CA groups. The only exception was that the SLI group produced fewer relative clauses than the CA group at T3. The finding that relative clauses are problematic for older children with SLI concurs with those of other studies (e.g., Jensen de López et al., 2014). Different explanations for this phenomenon in the research literature range from a deficit in grammatical knowledge to processing limitations, semantic deficits, or a limited input frequency. Such explanations may even be interrelated. However, a more detailed discussion is beyond the scope of this study. The picture that arises from our analyses is that the development of complex syntax in the SLI group is affected to a much smaller extent than the development of morphosyntax.

When we inspect the different developmental patterns for the separate measures in the three groups, we see that the children with SLI show continuous progress on the morphosyntactic accuracy measures. The developmental trajectories of the SLI and TD groups are more or less parallel, suggesting synchronous growth patterns. Furthermore, on verb-related errors, the SLI group seems to be catching up somewhat at T3 when compared with the CA group. However, such a pattern was not found for the use of dummy auxiliaries. In the LA and CA groups, changes in the use of dummy auxiliaries were seen over time. The spike at T2 in the LA group could be solely attributed to an abundant use of dummy verbs by three of the LA children. In the CA group, the use of dummy auxiliaries gradually decreased. In the SLI group, the "economy strategy"' of inserting a dummy auxiliary instead of using a finite lexical verb (de Jong, 1999; de Jong, Blom, & Orgassa, 2013; Orgassa, 2009) might be considered as a sign of fossilization. Compared with the TD groups, this stagnation pattern may be viewed as asynchronous, with the SLI group showing no sign of catching up with their CA controls.

The observed dissociation between the developmental profiles of morphosyntax and complex grammar in children with SLI is reminiscent of the dissociations between verb-related morphosyntax on the one hand and vocabulary noun morphology on the other, as observed by Rice and coworkers (Rice, 2003; Rice et al., 1998). Such dissociations are in need of an explanation. Rice's account is that verb-related morphosyntax reflects a specific component (module) of the mental grammar, separate from lexical knowledge, as well as from the module controlling nominal morphology. It was hypothesized (Rice, Wexler, & Cleave, 1995) that this grammar module is specifically affected in SLI, as described in the extended optional infinitive model. Our findings are partly compatible with this account (verb-related morphosyntax is seriously impacted), but at the same time, our results suggest that the extended optional infinitive model is too limited, because non–verb-related morphosyntax appears to be markedly poor in Dutch children with SLI as well. However, our results do seem to support the general notion that the mental grammar is composed of several, relatively independent modules or processing systems. If we assume that performance in the domain of complex grammar reflects the capacity to build syntactic structure, our results suggest that this capacity is distinct from the grammar module(s) that subserves morphosyntax (both verb-related and non–verb-related). Apparently, morphosyntax is compromised in SLI (in line with Rice and colleagues' proposal), whereas complex syntax is (virtually) unaffected. The design and results of our study do not allow us to explore the possible causes of this dissociation at a deeper level.

Conclusion

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

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

Clinical Implications

One important clinical implication of this study is that narrative tasks have considerable diagnostic value as compared with standardized tests, because a narrative analysis can offer a more detailed evaluation of the strengths and weaknesses in grammatical complexity and accuracy than can be derived from standardized tests. Grammatical profiles obtained from a narrative analysis can therefore help clinicians to select adequate therapy goals and evaluate the effects of intervention. Another clinical implication is that SLI is a dynamic condition and that grammatical skills of school-age children with SLI continue to develop. Therefore, the provision of language intervention beyond grade 3 still seems beneficial for children with SLI. Finally, our findings that school-age children with SLI do not lag behind in their use of complex syntax suggest that clinicians should not be worried about the possibility of overly burdening their linguistic systems. Instead, it might even be useful to include complex sentences as contexts to help students with SLI enhance their morphosyntactic skills.

Acknowledgments

This study was supported by the Royal Dutch Auris Group. We thank Cito for their kind permission to use the images of the Dutch Language Proficiency Test for All Children storytelling tasks. We also thank the children, parents, and staff of all the schools that participated, and we appreciate Marij van Ewijk, Henaly Leijenhorst, Marjolein van der Horst, and Merel van Witteloostuijn for their help with the transcriptions.

References

Blom, E., Vasić, N., & de Jong, J. (2014). Production and processing of subject–verb agreement in monolingual Dutch children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 57,* 952–965.

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

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

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

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

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

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

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

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

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

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

Newman, R., & McGregor, K. (2006). Teachers and laypersons discern quality differences between narratives produced by children with or without SLI. *Journal of Speech, Language, and Hearing Research, 49,* 1022–1036.

Orgassa, A. (2009). *Specific language impairment in a bilingual context. The acquisition of Dutch by Turkish-Dutch learners* (Doctoral dissertation). University of Amsterdam, Utrecht, the Netherlands. LOT Dissertation Series 220. Retrieved from http://dare.uva.nl/record/1/312821

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

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

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

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

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1412–1431.

Shetter, W. Z., & Ham, E. (2007). *Dutch. An essential grammar* (9th ed.). New York, NY: Routledge.

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

van Bon, W. H. J. (1986). *Coloured progressive matrices: Manual of Dutch norms*. London, UK: Lewis.

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

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

Verhoeven, L., & Vermeer, A. (2006). *Verantwoording Taaltoets Alle Kinderen* (TAK) [Justification of the Language Proficiency Test for All Children]. Arnhem, the Netherlands: Cito.

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

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

Appendix

Verb-Related Errors: Errors in Subject–Verb Agreement, Tense, Overregularizations, Verb-Second, Root Infinitives, Subject and Object Omissions (Mean Percentages, Standard Deviations, Raw Numbers/Number of Clauses and Ranges) on T1, T2, and T3 for the Groups

	SLI group			LA group			CA group		
	M (SD)			*M (SD)*			*M (SD)*		
Measure	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
Agreement errors									
M % (*SD*)	13.7 (11.6)	8.1 (8.0)	4.6 (4.4)	7.6 (8.6)	4.7 (7.8)	1.7 (2.9)	2.9 (5.7)	1.2 (2.0)	0.8 (1.6)
Raw/clauses	91/679	67/791	38/906	41/571	28/628	11/643	16/637	8/782	7/806
Range	0–13	0–8	0–4	0–6	0–5	0–2	0–5	0–1	0–1
Tense errors									
M % (*SD*)	5.0 (6.5)	6.5 (6.9)	5.0 (7.5)	1.9 (4.7)	2.3 (6.5)	1.3 (2.9)	1.1 (2.4)	0.8 (1.7)	1.0 (2.1)
Raw/clauses	35/697	49/79	44/906	12/571	13/628	8/643	7/637	6/78	8/806
Range	0–6	0–6	0–11	0–5	0–5	0–2	0–2	0–1	0–2
Overregularizations									
M % (*SD*)	1.8 (4.7)	1.7 (2.8)	2.6 (3.4)	1.4 (3.7)	0.5 (1.5)	0.8 (1.9)	1.9 (4.3)	0.1 (0.6)	1.1 (2.9)
Raw/clauses	13/679	15/791	25/90	7/571	3/628	5/643	12/637	1/782	8/806
Range	0–4	0–3	0–5	0–2	0–1	0–1	0–3	0–1	0–3
Verb-second errors									
M % (*SD*)	3.5 (6.1)	1.9 (4.7)	2.5 (3.0)	2.6 (10.2)	1.0 (1.9)	0.0 (0.0)	0.4 (1.7)	0.1 (0.6)	0.4 (1.1)
Raw/clauses	22/679	19/791	24/906	14/571	7/628	0/643	3/637	1/782	3/806
Range	0–5	0–11	0–4	0–10	0–1	0–0	0–2	0–1	0–1
Root infinitives									
M % (*SD*)	3.0 (6.1)	1.1 (2.7)	0.7 (1.6)	1.7 (4.6)	0.2 (1.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0
Raw/clauses	20/679	8/791	5/90	7/571	1/628	0/643	0/637	0/782	0/806
Range	0–6	0–3	0–1	0–2	0–1	0–0	0–0	0–0	0–0
Subject omissions									
M % (*SD*)	8.3 (8.2)	4.9 (4.3)	4.5 (4.5)	3.7 (8.7)	1.0 (2.0)	0.4 (1.5)	0.8 (2.0)	0.2 (0.9)	0.3 (1.3)
Raw/subject–verb clauses	54/660	39/777	38/872	12/560	6/622	2/630	4/62	1/777	2/773
Range	0–6	0–3	0–3	0–2	0–1	0–1	0–1	0–1	0–1
Object omissions									
M % (*SD*)	13.8 (13.0)	7.7 (9.9)	8.8 (9.9)	4.5 (8.0)	4.8 (7.9)	1.2 (4.3)	1.3 (4.3)	0.2 (1.1)	0.3 (1.5)
Raw/subject–verb–object clauses	27/282	22/344	28/405	9/258	10/292	3/314	3/300	1/356	1/399
Range	0–3	0–2	0–3	0–2	0–2	0–2	0–1	0–1	0–1