

## SLI Diagnostics in Narratives: Exploring the CLARIN-NL VALID Data Archive

Laura Bergmann, Roeland van Hout and Jetske Klatter-Folmer

[j.klatter@let.ru.nl](mailto:j.klatter@let.ru.nl)

### ABSTRACT

In 2014 the Vulnerability in Acquisition: Language Impairments in Dutch (VALID)<sup>1</sup> Data Archive for pathological language data (CLARIN-NL-12-010 grant) was launched. The aim of the VALID Data Archive is to unite various available datasets ranging from metadata, experimental results, and test outcomes to spontaneous speech data, including video recordings, and to develop unambiguous protocols to ascertain the interpretation of research outcomes. In this chapter we report a study that we carried out using the VALID Data Archive. In an earlier project the language development of children with Specific Language Impairment (SLI) had indeed been investigated using a narrative task (retelling of a picture story); the VALID database thus contains transcripts and audio files of the speech of 50 children with SLI and 24 age-matched typically developing (TD) children in the age range between 5;6 to 12;0 years, who all participated in this earlier project. Our study focused on morphosyntactic and lexical accuracy and complexity, in order to determine which language measures are diagnostic indicators of SLI on the basis of this narrative data. Results showed that SLI children performed less well than TD children for morphosyntactic and lexical accuracy and complexity. Interestingly, the results obtained can be compared to results found in three other studies on narratives performed by SLI and TD children. The similarities and differences in the outcomes reveal the urgency to have identical, precise protocols in handling and analysing complex data.

---

<sup>1</sup> <http://validdata.org/clarin-project/datasets/>

---

#### How to cite this book chapter:

Bergmann, L, van Hout, R and Klatter-Folmer, J. 2017. SLI Diagnostics in Narratives: Exploring the CLARIN-NL VALID Data Archive. In: Odijk, J and van Hessen, A. (eds.) *CLARIN in the Low Countries*, Pp. 167–180. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bbi.14>. License: CC-BY 4.0

## 14.1 The VALID Data Archive

The Vulnerability in Acquisition: Language Impairments in Dutch (VALID) Data Archive (CLARIN-NL-12-010 grant) that was launched in 2014 is an open multimedia data archive with data from speakers suffering from language impairments. The aim of the VALID Data Archive is to unite various available datasets ranging from metadata, experimental results, and test outcomes to spontaneous speech data, including video recordings, and to develop unambiguous protocols to ascertain the interpretation of research outcomes. In the CLARIN-NL framework five VALID data resources were curated; an overview of the key information on each of these five data resources is provided in the Appendix. For all datasets concerned, written informed consent has been obtained from the participants or their carers. All materials were anonymised. The audio files were converted into wav (linear PCM) files and the transcriptions into CHAT or ELAN format. Research data that consisted of test, SPSS and Excel files were documented and converted into CSV files. All datasets obtained appropriate CMDI metadata files. A new CMDI metadata profile for this type of data resources was established and care was taken that ISOcat metadata categories were used to optimise interoperability. A full overview of VALID metadata categories can be found in Klatter-Folmer et al. (2014). After curation all data were deposited at the Max Planck Institute for Psycholinguistics in Nijmegen, where persistent identifiers are linked to all resources. The content of the transcriptions in CHAT and plain text format can be searched with the TROVA search engine (cf. Klatter-Folmer et al., 2014; van den Heuvel et al., 2014).

The most important difference with the Child Language Data Exchange System (CHILDES<sup>2</sup>) is that VALID is a specialised structured database for all types of data related to pathological language, ranging from metadata, experimental results and test outcomes to spontaneous speech data, including video recordings. CHILDES, on the other hand, covers the spectrum of first-language acquisition research data, focusing in particular on spontaneous speech data, and with fewer datasets from child clinical groups. Moreover, the VALID Data Archive covers all age groups.

The realisation of the data archive was made possible by a CLARIN-NL grant (12-010) for a pilot project. This pilot enabled us to build up experience in conserving different kinds of pathological language data in a searchable and persistent manner. The conserved datasets reflect current research in language pathology rather well, both in the range of designs and in the variety in pathological problems, such as Specific Language Impairment (SLI), deafness, dyslexia and ADHD (Klatter-Folmer et al., 2014; van den Heuvel et al., 2014). The first author of the present contribution carried out the study presented below (Bergmann, 2015), monitored by Roeland van Hout (VALID data provider) and Jetske Klatter (VALID project leader). A main goal of this study was to test the accessibility of the VALID data archive and to signal problems met in extracting the data.

## 14.2 SLI Diagnostics in Narratives

SLI is a set of speech and language disorders with high co-morbidity with other disorders and impairments. Its definition is based on exclusion criteria and is related to a mix of linguistic, sensory, cognitive, neural-motor, and emotional restrictions. This rather unsatisfactory definition is largely due to the heterogeneous speech and language behaviour of SLI children (Manders, De Bal and Van den Heuvel, 2013), while at the same time no specific causes of SLI have been detected yet (Archibald and Gathercole, 2006). The co-morbidity patterns found do support the idea that SLI is a multi-factorial disorder (Bishop, 2006).

SLI children display a problematic and delayed development in language form, function, and use, where impairments may occur in all language domains, such as phonology, semantics,

---

<sup>2</sup> <http://childes.talkbank.org/>

morphosyntax, and pragmatics (Casalini et al., 2007). Bishop (2006) concluded that SLI children obviously have difficulties in adequately processing information that is being offered in a short time span, as is the case in spoken conversations. For the majority of SLI children, grammar is a difficult area, and weak morphosyntactic skills are correlated with poor lexical-semantic skills (Simon-Cerejido and Gutiérrez-Clellen, 2009; Toppelberg and Shapiro, 2000; Bishop, 2013).

Studies focusing on the complexity and accuracy of morphosyntax address a range of features. Smith-Lock (1993) already pointed to differences between SLI and typically developing (TD) children in passive sentence constructions, and Rice performed several investigations into mean length of utterance (MLU), showing that SLI children lag behind in MLU, partly because of the absence of complex morphosyntactic constructions, e.g. subordinate clauses and question clauses (Rice, Redmond and Hoffman, 2006). As for accuracy, research by Vandewalle et al. (2012) showed errors in verb inflection, articles, and word order, when compared to TD children, and Simon-Cerejido and Gutiérrez-Clellen (2009) mentioned deletion of function words. In SLI, the production of complex utterances triggers an increase in morphosyntactic errors, as complex utterances are obviously more demanding (Colozzo et al., 2011).

Considering lexical complexity and accuracy, Bishop (1992) argued that SLI children have difficulties processing linguistic input as a whole, resulting in weak and inefficient connections between words, which in turn leads to longer retrieval time and more errors in word choice (Kambanaros et al., 2014). In a longitudinal study of 500 SLI and TD children, Rice and Hoffman (2015) found that SLI children consistently performed less well than age-matched TD children on lexical tasks.

Gaining more insight into the causes and characteristics of SLI requires a detailed diagnostic procedure. The usual battery of SLI test materials focuses on communication in structured settings, such as inviting participants to select the image that best represents a stimulus word. These experimental settings are unnatural and provide scarce information about linguistic skills in a spontaneous or semi-structured conversation (Peña et al., 2006). This argues for using narrative tests that combine spontaneous quality with a structured content. Retelling a picture story requires quite different competencies to those used in structured settings, such as introducing the characters, explaining the topic and structuring the text. Also, (re)telling a story challenges people to be more explicit and to produce longer linguistic units (Treurniet, 2011; Treurniet and Orgassa, 2011). As in these tasks children tend to show more linguistic variation and produce more utterances, they are an appropriate means for collecting data on morphosyntactic and lexical skills. Several studies confirm that narratives demonstrate the morphosyntactic and lexical problems of SLI children (e.g. Kambanaros et al., 2014; Vandewalle et al., 2012).

The morphosyntactic and/or lexical accuracy and complexity of narratives by SLI children were analysed in three earlier studies in the Netherlands: that of Treurniet (2011); Verhoeven, Steenge and Van Balkom (2011); and Zwitserlood et al. (2015). All three studies mention problems in the morphosyntactic and/or lexical domain for SLI children. Each used a different set of narrative data. The VALID Data Archive contains yet another, new narrative dataset. Departing from the Dutch studies, the following research questions and hypotheses were formulated for our study:

- A. How do SLI children perform with regard to morphosyntactic accuracy and complexity in a narrative in comparison to their typically developing peers?
  - H1: SLI children use morphosyntactically less complex language
  - H2: SLI children are morphosyntactically less accurate
- B. How do SLI children perform with regard to lexical accuracy and complexity in a narrative in comparison to their typically developing peers?
  - H3: SLI children use lexically less complex language
  - H4: SLI children are lexically less accurate

### 14.3 Method

All the tasks performed by participants in a collective research project led by Radboud University Nijmegen and Kentalis on the expression of spatial relations by SLI children in oral language production were stored in the CLARIN-NL VALID Data Archive; this so-called SLI RU-Kentalis database was one of the five sets that were curated. This database contains narratives by SLI and TD children. In this contribution we discuss a new analysis of the *Frog goes to dinner* narrative. Analysing these data from a perspective that differs from the main aim of the original project (which was to study how SLI children expressed spatial relations in this narrative) was a good test to explore whether the VALID Data Archive is easily accessible and usable for new researchers with new questions.

The Netherlands has special schools for SLI children. To be eligible for special education and extra care, children with SLI have to meet certain criteria that have been acknowledged by the Ministry of Education, Cultural Affairs and Science of the Dutch government. A child's communicative and cognitive abilities are assessed in an examination by a speech therapist, a psychologist and if necessary an audiologist. The SLI diagnosis is made when a child has speech or language impairments that cannot be attributed to limited cognitive abilities. Furthermore, it has to be established that the child has problems in two or more of the following language areas: speech production, auditory processing, and grammatical development or lexical-semantic knowledge. Only the children whose scores on standardised language tests for at least two of these aspects of language are 1.5 standard deviation below average are admitted onto a special form of education.

Only the children with grammatical and lexical-semantic problems were included in the original research project. The selection was made based on the children's achievements on standardised language tests given by the SLI schools: children whose language scores were 1.5 standard deviation or more below average on at least one subtest measuring syntactic and semantic development met the criteria to be included in the research project. The children who participated in the study all have Dutch as a first language.

The sample we used included 74 children out of the total of 93 participants in the original research project: 50 SLI children and 24 TD children, 40 boys and 34 girls. Main reasons for excluding children from the sample were that we constrained the analysis to three age groups or that for some participants less than 30 utterances that could be analysed for the narrative were available (see Bishop and McDonald, 2010). The children came from primary school classes 2 to 7 (Dutch school system), and both the SLI and the TD children were divided into three age groups in order to investigate age effects in the development of their language proficiency (Table 14.1).

The picture book that was used in the narrative was *Frog goes to dinner* (Mayer, 1974). The reason for choosing this book instead of the more famous *Frog, where are you?* (Mayer, 1969) was that the *Frog goes to dinner* book contained a large and varied number of spatial elements and relations which suited the research perspective of the main project much better. In this black-and-white illustrated story, a little boy brings his frog to a fancy restaurant. The frog manages to escape from

	N	Gender		Age group								
		Male	Female	5;6 – 7;5			7;6 – 9;5			9;6 – 12;0		
				N	mean	SD	N	mean	SD	N	mean	SD
TD	24	12	12	10	6.60	0.77	8	8.60	0.62	6	10.27	0.62
SLI	50	28	22	10	6.87	0.39	16	8.13	0.68	24	10.62	0.83
Total	74	40	34		20			24			30	

**Table 14.1:** Overview of subjects: gender; age group (N, mean age, and SD).

the boy's coat pocket and causes a number of incidents. The plot is worked out in 30 pictures. Before starting the audio recording, the children were invited to leaf through the book. During the actual retelling the researcher, if necessary, asked open questions to motivate the child to continue the narrative. The narrative was recorded using a Sony MZ-NH 700 minidisk recorder. An external microphone was added to enhance the quality of the recordings.

The VALID Data Archive contains not only the audio files, but also the transcripts and TextGrids for the PRAAT analysis tool (Boersma and Weenink, 2004). A TextGrid is a transcript of the audio file that can be made visible simultaneously with the audio file in PRAAT. The transcripts are available in CHAT format (MacWhinney, 2000), which makes calculation of MLU5 and Guiraud's index (CLAN) possible by using CLAN tools.

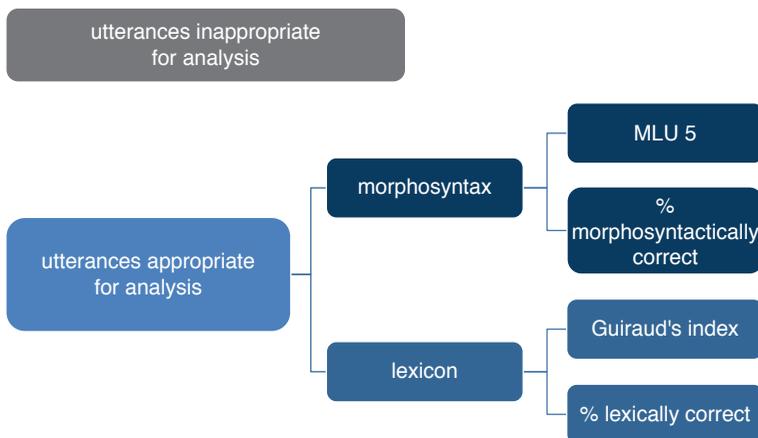
Before starting the analysis, all transcripts were processed to mark and select the utterances appropriate for the analyses to be performed. Examples of utterances labelled as not appropriate for deeper analyses on the utterance level were 'yes' and 'no' answers and straightforward formulaic utterances (e.g. 'ik weet het niet', 'I don't know'). The steps of analysis are depicted in Figure 14.1. In general, in all three age groups, the TD group produced more appropriate or usable utterances than the SLI group. Both the transcripts and the audio files were used to analyse the data.

Morphosyntactic complexity was measured with MLU5, which stands for the mean length of utterances (in words) of the five longest utterances of a child, using CLAN.

Morphosyntactic accuracy, as the percentage of correct utterances, was measured by marking utterances containing errors in: position/inflection of verbs, noun form, word order, omission of function words, and grammatical gender. Utterances that were rectified by the child (self-correction) were labelled as correct.

For lexical complexity Guiraud's index was calculated. First a list was drawn up of all words in the transcripts, excluding proper names, onomatopoeia, and noninterpretable words. Words like 'ja' ('yes') and 'ok' were also left out, because they are not suitable to establish a child's vocabulary size (Schaerlaekens, 2008). Guiraud's index is computed by dividing the number of types by the square root of the number of tokens, resulting in a measure of richness of the productive lexicon.

Lexical accuracy was again expressed by the percentage of correct utterances. Utterances containing incorrect function words (e.g. prepositions, conjunctions) and neologisms were categorised as incorrect. The same procedure was applied if content words were used with a wrong meaning.



**Figure 14.1:** Analysis diagram.

## 14.4 Results

### 14.4.1 Morphosyntactic Complexity

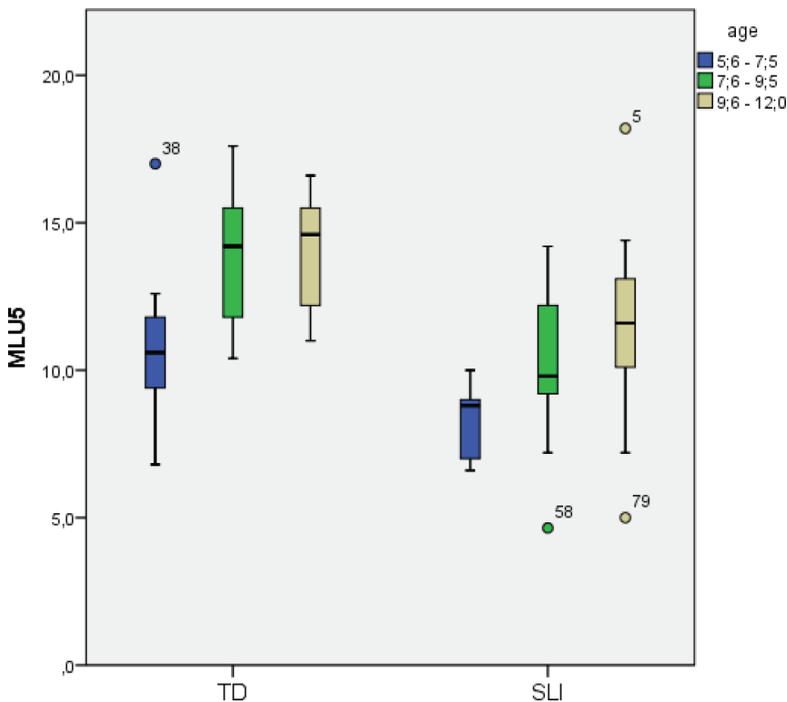
This variable was investigated by selecting the five longest utterances in words that were lexically and morphosyntactically correct: the MLU5. Figure 14.2 gives the relevant box plots for TD and SLI children.

The (TD vs. SLI) Group ( $F(1, 68) = 21,73, p = 0.000, \eta_p^2 = 0,24$ ) and Age ( $F(2, 68) = 9,42, p = 0.000, \eta_p^2 = 0,22$ ) factors turned out to be significant in ANOVA. The interaction effect between these two variables was not significant ( $F(2, 68) = 0,37, p = 0.692, \eta_p^2 = 0,01$ ). Posthoc tests for age group (Tukey HSD) showed differences between the youngest and the middle group, but not between the middle and oldest group. The TD children reached higher MLU5 values than the SLI group. MLU5 increased with age between the youngest and the two older groups.

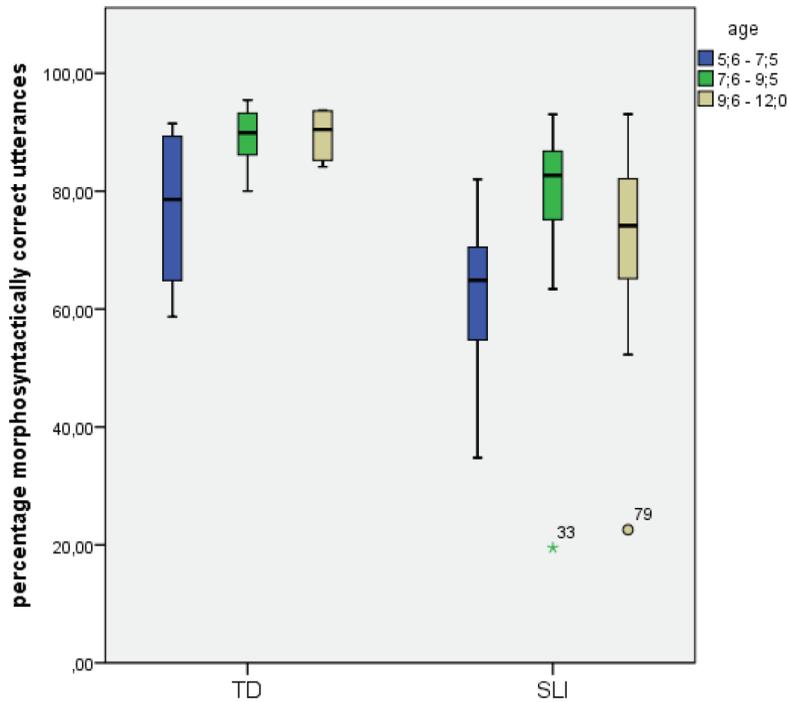
### 14.4.2 Morphosyntactic Accuracy

In Figure 14.3 accuracy in morphosyntax is shown by the box plots of the percentage of correct utterances.

There was a significant effect for Group ( $F(1, 68) = 15,79, p = 0.000, \eta_p^2 = 0,19$ ) and Age ( $F(2, 68) = 5,91, p = 0.004, \eta_p^2 = 0,15$ ), but there was again no interaction effect ( $F(2, 68) = 0,22, p = 0.805, \eta_p^2 = 0,01$ ). Tukey HSD tests pointed again to differences between the youngest and the two older age groups.



**Figure 14.2:** Box plots of the mean length in words of the five morphosyntactically longest utterances (MLU5), by child (TD or SLI) and age group.



**Figure 14.3:** Box plots showing the percentage of morphosyntactically correct utterances by child (TD or SLI) and age group.

#### 14.4.3 Lexical Complexity

The box plots of the scores on Guiraud's index for lexical richness are shown in Figure 14.4.

Group ( $F(1, 68) = 4,39, p = 0.040, \eta_p^2 = 0,06$ ) and Age ( $F(2, 68) = 4,27, p = 0.018, \eta_p^2 = 0,11$ ) turned out to be significant, with no significant interaction effect ( $F(2, 68) = 1,68, p = 0.194, \eta_p^2 = 0,05$ ). Tukey HSD did not show differences between the age groups, indicating that the age differences are not large.

#### 14.4.4 Lexical Accuracy

The box plots of accuracy in vocabulary are given in Figure 14.5.

ANOVA showed that Group caused a significant effect ( $F(1, 68) = 0,16, p = 0.017, \eta_p^2 = 0,08$ ), which did not apply to the Age factor ( $F(2, 68) = 1,97, p = 0.854, \eta_p^2 = 0,01$ ). An interaction effect between age group and SLI was established ( $F(2, 68) = 3,80, p = 0.027, \eta_p^2 = 0,10$ ). Tukey HSD did not point to differences between age groups.

### 14.5 Discussion

In this section, we first relate the results to our hypotheses. Table 14.2 gives the partial eta squares ( $\eta^2$ ) for the effects that turned out to be significant.

**Morphosyntactic skills** H1: SLI children use morphosyntactically less complex language.  
H2: SLI children are morphosyntactically less accurate.

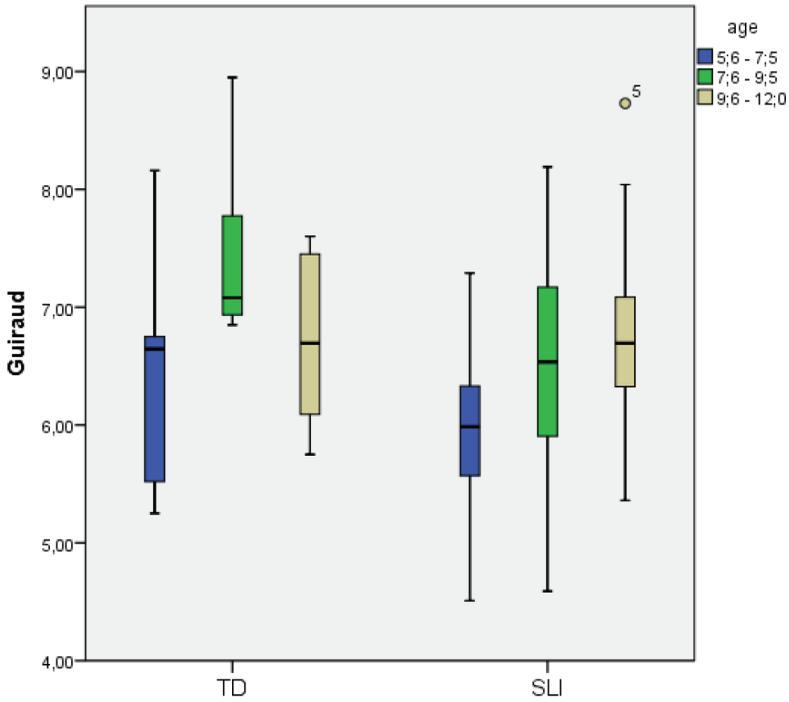


Figure 14.4: Box plots for Guiraud's index, split out by child (TD or SLI) and age group.

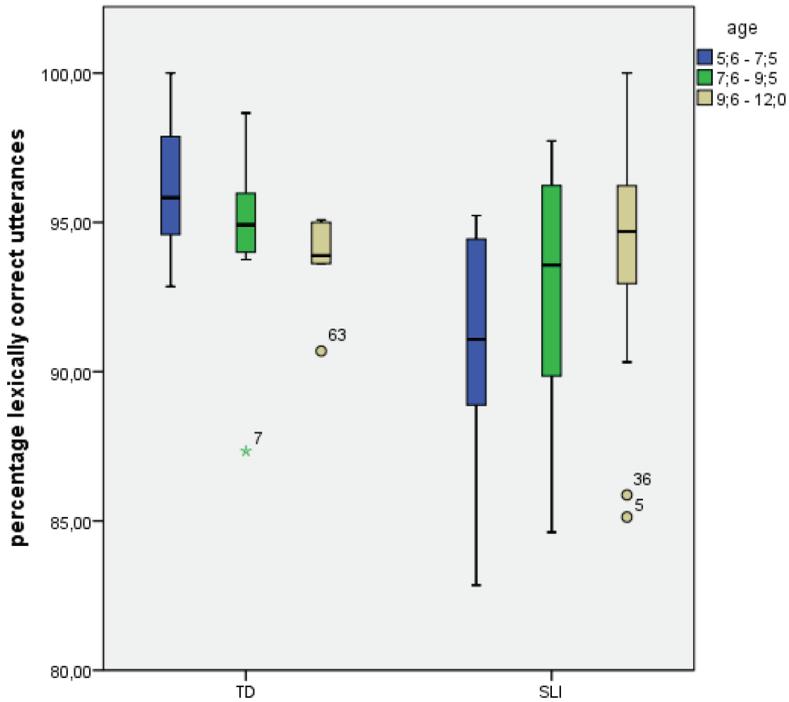


Figure 14.5: Box plots showing the percentage of lexically correct utterances, by child (TD or SLI) and age group.

Language modality	Variable	SLI	Age	SLI × Age
Morphosyntax	MLU5	0.24	0.22	—*
	% correct	0.19	0.15	—
Lexicon	Guiraud	0.06	0.11	—
	% correct	0.08	—	0.10

\* No significant effect found

**Table 14.2:** Overview of significant effects of the ANOVA analyses ( $\eta^2$ ).

The results of our study support both hypotheses on morphological skills. In all age groups SLI children produced morphosyntactically less complex and less accurate utterances than age-matched TD children. There is an age effect as well, indicating development over time in both groups. The morphosyntactic skills of SLI children were lower than those of their TD peers. These findings correspond to findings in other studies: Heilmann, Miller and Nockerts (2010) consider MLU a valid diagnostic variable for SLI, and so do Dunn, Flax and Sliwinski (1996). Smith-Lock (1993) has advised to carefully interpret MLU because raw data do not warrant any direct conclusions about the complexity of language, but in combination with other linguistic variables, e.g. the percentage of morphosyntactically correct utterances, MLU is useful in identifying SLI (Moyle et al., 2011; Simon-Cerejido and Gutiérrez-Clellen, 2009). Colozzo et al. (2011) also concluded that SLI children have problems in telling a grammatically accurate story; they therefore investigated the content quality of the story, and observed that children with weak morphosyntactic skills delivered less consistent narratives. In our study an assessment of the content quality was not carried out, but in future research (using the CLARIN-NL VALID Data Archive) content quality could be used to better assess the language production processes in SLI children.

**Lexical skills** H3: SLI children use lexically less complex language

H4: SLI children are lexically less accurate

The two lexical hypotheses were also corroborated by the data. In comparison to their TD peers, children with SLI produced less lexically complex and less accurate utterances. The age effect was straightforward for complexity: complexity increased with age. There was no main effect of age for accuracy, but an interaction between group and age, indicating that the age difference only affected the SLI group. The TD children already had high lexical accuracy scores in the youngest group, and the mean score does not show a pattern of change over time. The SLI lexical scores are lower than the TD lexical scores, but importantly the SLI lexical accuracy scores are much higher (all above 80%) than the SLI morphosyntactic accuracy scores (with scores as low as 20%). Lexical problems apparently are less strong or surface less strongly than morphosyntactic problems in SLI children, a conclusion supported by the lower partial eta squared values in the lexical outcomes. The morphosyntactic and lexical accuracy scores are marked by outliers, i.e. children who have scores far higher or lower than their age and child group. The variability in scores seems more typical of SLI children, who can have extremely severe impairments in specific linguistic domains. Remarkably, there are high correlations between the lexical and the morphosyntactic accuracy score, i.e. 0.82 for the TD children and 0.92 for the SLI children. This result deviates from the findings in Kambanaros et al. (2014), who investigated lexical proficiency with similar variables: SLI children performed less accurately on the lexical and morphosyntactic level, but contrary to our study no relation between the two skills was observed.

How do our results compare to results found earlier in analyses of narratives of monolingual Dutch TD and SLI children? Treurniet (2011; see also Treurniet and Orgassa, 2011 compared 7-year-old SLI children to 5-year-old TD children. The TD children, although two years younger, turned out to have the same scores as the SLI children for lexical and morphosyntactic accuracy.

As in our study the scores on morphosyntactic accuracy (53%, a score lower than our scores in the same age groups) were much lower than lexical accuracy ones (88%, a score comparable to our scores). Their average Guiraud scores (5.6) seem a bit lower than ours, but they had a different frog story.

A second study including monolingual TD and SLI children is the one performed by Verhoeven et al. (2011), who measured MLU on all usable utterances and grammatical accuracy, with criteria similar to ours. Their MLU returned a significant SLI effect ( $\eta^2 = .06$ ), a significant age effect (two age groups: 7- and 9-year-olds;  $\eta^2 = .19$ ) and no interaction effect. Although they computed the MLU on all usable utterances their effect sizes were lower than ours – this perhaps demonstrates that it is preferable to restrict the MLU to the subset of longest utterances. Their grammatical accuracy scores give an SLI effect ( $\eta^2 = .38$ ), an age effect ( $\eta^2 = .06$ ), and no interaction. The scores (TD with an average of 80% grammatical accuracy and SLI with an average of 46% grammatical accuracy) are lower than we found, and the gap between SLI and TD children seems wider.

A third study on narratives is the longitudinal study by Zwitserlood et al. (2015), which included monolingual SLI and TD children as well, ranging in age from 6.5 to 8.5 years. The authors found an SLI effect ( $\eta^2 = .27$ ), an age effect ( $\eta^2 = .16$ ), and no interaction effect between age and SLI for MLU. They found particular strong effects for grammatical accuracy: the SLI effect was  $\eta^2 = .72$ , the age effect  $\eta^2 = .34$ , and there was no interaction effect; the TD scores, with an average of 85%, are similar to ours, while the SLI scores, with an average of 48%, are much lower.

The positive outcome of all studies, including ours, is that the effects found are comparable. The effects all show that our measures, in particular morphosyntactic/grammatical accuracy, proved to be useful in diagnosing SLI children. On the other hand there is an obvious overlap in scores between the TD and SLI groups, restricting the diagnostic value of our measures. At the same time the results of the four studies show substantial variation in effect sizes. Stronger effect sizes are crucial when it comes to powerful diagnostics. The differences in outcomes between the four studies may have multiple sources, e.g. the homogeneity of the groups of participants, the design type (cross-sectional vs. longitudinal), and the coding schemes applied to the complicated, rich narrative data. The data curated in the VALID Data Archive, accessible to all interested researchers, demonstrates the necessity to combine and accumulate data not only in their raw format but also through coding schemes and coded data, to increase the analytic power and to calibrate our research tools. That seems particularly relevant when the data are intricate, as in the case of narratives.

Finally, it is important to observe that precise details on data analysis are often lacking in published articles, including the ones we discussed. One often needs the detailed, original protocols to understand how specific decisions are being taken in defining utterances (or perhaps T-units) and subordination (directly relevant to the MLU). The same applies to the definition of words, morphemes and accuracy measures. These decisions have direct consequences for the outcomes and may obscure the (dis)similarities between different studies.

## 14.6 Conclusion

As frequently stated in the research literature on language development, narratives are an appropriate and attractive option for gathering rich data on the linguistic, cognitive, and social competencies of SLI children (Befi-Lopes, Bento and Perissinoto, 2008). Gathering, transcribing and coding narratives is laborious and time-consuming, however. Given their richness it seems self-evident to store such data sources in accessible, standardised formats. The CLARIN goals made the VALID Data Archive possible and we see this as a first step in establishing the availability of data sources to improve and widen the research perspectives on language and speech pathology (Rietveld et al., 2005).

We analysed the dataset on SLI children available in the VALID Data Archive and found several small infelicities that could be remedied – the Data Archive is now more easily usable. Using and trying out the Data Archive is an important step not only in improving the database, but also in adding to the data new information that came out of the new analyses. It also shows how important it is to have other research data available. CLARIN also made available the data used in Treurniet (2011) and the Functional Elements in Specific Language Impairment (FESLI)<sup>3</sup> data (Treurniet and Orgassa, 2011); it seems self-evident to link the FESLI corpus (12 bilingual children without SLI, 25 monolingual children with SLI, 20 bilingual children with SLI) more directly to the VALID Data Archive. The Verhoeven et al. (2011) and Zwitserlood et al. (2015) data are unfortunately not accessible. The main goal of the VALID enterprise is to include other databases in the archive. An archive is pivotal in evaluating the robustness of experimental outcomes in terms of reproducibility and replicability. VALID is a proper medium to guarantee the quality and comparability of datasets.

The results of the present investigation motivate more in-depth research on morphosyntactic and lexical variables in order to improve the diagnostics and treatment of SLI children. The effects for morphosyntax were strongest, but in absolute terms were still weak. Morphosyntactic variables are generally considered important indicators for identifying SLI (Dunn, Flax and Sliwinski, 1996). The availability of larger amounts of standardised and enriched data sources might sharpen our analyses by enabling us to focus more on correlational patterns (what is the relation between morphosyntactic and lexical problems, a correlation we found in our analysis but which is not reported in the other Dutch studies) and on actual speech patterns (by applying for instance machine learning techniques on string data). CLARIN and its data formats offer the proper perspective on establishing the rich data sources we need and, hopefully, will motivate researchers to make their data available on the internet.

## References

- Archibald, L., & Gathercole, S. (2006). Prevalence of SLI in language resource units. *Journal of Research in Special Educational Needs*, 6, 3–10. doi: 10.1111/J.1471-3802.2006.00054.x
- Befi-Lopes, D., Bento, A.C.P., & Perissinoto, J. (2008). Narration of stories by children with specific language impairment. *Pró-Fono Revista de Atualização Científica*, 2, 93–8. doi: 10.1590/S1516-18462012005000105
- Bergmann, L. (2015). *Taalvaardigheden in narratieven van kinderen met SLI*. Nijmegen: Radboud University (master thesis).
- Bishop, D.V.M. (1992). The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry*, 33, 3–66. doi: 10.1111/j.1469-7610.1992.tb00858.x
- Bishop, D.V.M. (2006). What causes Specific Language Impairment in children? *Current Directions in Psychological Science*, 15(5), 217–221. doi: 10.1111/j.1467-8721.2006.00439.x
- Bishop, D.V.M. (2013). Neuroscientific studies of intervention for language impairment in children: Interpretive and methodological problems. Emanuel Miller Memorial Lecture 2012. *Journal of Child Psychology and Psychiatry*, 54(3), 247–259. doi: 10.1111/jcpp.12034
- Bishop, D.V.M., & McDonald, D. (2010). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language and Communication Disorders*, 44, 600–615. doi: 10.1080/13682820802259662
- Boersma, P., & Weenink, D. (2004). *PRAAT*. Amsterdam: Universiteit van Amsterdam.
- Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, S., Pecini, C., Roncoli, S., & Burani, C. (2007). Non-word repetition in children with specific language impairment: A deficit in phonological working memory or in long-term verbal knowledge? *Cortex*, 43, 769–76.

<sup>3</sup> <https://dev.clarin.nl/node/1948>

- Colozzo, P., Gillam, R.B., Wood, M., Schnell, R.D., & Johnston, J.R. (2011). Content and form in the narratives of children with Specific Language Impairment. *Journal of Speech Language and Hearing Research, 54*, 1609–1627. doi:10.1044/1092-4388(2011/10-0247)
- Dunn, M., Flax, J., & Sliwinski, M. (1996). The use of spontaneous language measures as criteria for identifying children with Specific Language Impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech Language and Hearing Research, 39*, 643–654. doi: 10.1044/jshr.3903.643
- Heilmann, J.J., Miller, J.F., & Nockerts, A. (2010). Using language sample databases. *Language, speech, and hearing services, 41*, 84–95. doi: 10.1044/0161-1461(2009/08-0075)
- Kambanaros, M., Grohmann, K.K., Theodorou, E., & Michaelides, M. (2014). Can vocabulary size predict narrative abilities in children with SLI? *Language Disorders, 6*, 60–81. doi: 10.1080/14790718.2012.705846
- Klatter-Folmer, J., van Hout, R., van den Heuvel, H, Fikkert, P., Baker, A., Jong, J. de, Wijnen, F., Sanders, E. & Trilsbeek, P. (2014). Vulnerability in Acquisition: Language Impairments in Dutch Creating a VALID Data Archive. *Language Resources and Evaluation Conference Proceedings (LREC)*, Reykjavik, 26–31 May 2014, 357–364.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. 3rd edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Manders, E., Bal, C. De, & Heuvel, E. Van den (2013). *Taalontwikkelingsstoornissen: fenomenen, onderzoek en behandeling [Language development disorders: phenomena, research, and treatment]*. Antwerpen, Belgium: Grant.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.
- Mayer, M. (1974). *Frog goes to dinner*. New York: Dial Press.
- Moyle, M.J., Karasinski, C., Weismer, S.E., & Gorman, B.K. (2011). Grammatical morphology in school-age children with and without language impairment: A discriminant function analysis. *Journal of Language, Speech and Hearing Service, 42*, 550–560. doi: 10.1044/0161-1461(2011/10-0029)
- Peña, E.D., Gillam, R.B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research, 49*, 1037–1057. doi: 10.1044/1092-4388(2006/074)
- Rice, M.L., Redmond, S.M., & Hoffman, L. (2006). Mean Length of Utterance in children with Specific Language Impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. *Journal of Speech, Language and Hearing Research, 49*, 793–808. doi: 1092-4388/06/4904-0793
- Rice, M.L., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without Specific Language Impairment (SLI): A longitudinal study from 2½ to 21 years of age. *Journal of Speech Language Hearing Research, 58*, 345–359. doi:10.1044/2015\_JSLHR-L-14-0150
- Rietveld, T., Stolte, I., van den Heuvel, T., Klatter, J., van Balkom, H., Dupont, J., & Ruiter, M. (2005). *Taal- en spraaktechnologie en communicatieve beperkingen*. 's-Gravenhage: Nederlandse Taalunie.
- Schaerlaekens, A. (2008). *De taalontwikkeling van het kind [The language development of the child]*. Groningen, The Netherlands: Noordhoff Uitgevers.
- Simon-Cerejido, G., & Gutiérrez-Clellen, V. (2009). Spontaneous language markers of Spanish language impairment. *Applied Psycholinguistics, 28*, 317–339. doi: 10.1017.S0142716407070166
- Smith-Lock, K.M. (1993). Morphological analysis and the acquisition of morphology and syntax in specifically-language-impaired children. *Status Report on Speech Research, 114*, 113–138.
- Toppelberg, C.O., & Shapiro, T. (2000). Language disorders: A 10-year research update review. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 143–152. doi: 10.1097/00004583-200002000-00011.

- Treurniet, M. (2011). *Tweetaligheid en SLI: morfosyntactische en lexicale vaardigheid in narratieven* [Bilingualism and SLI: Morphosyntactic and lexical skills in narratives]. Nijmegen, The Netherlands: Radboud University.
- Treurniet, M. & Orgassa, A. (2011). 'Kikker, waar ben jij?' Grammaticale en lexicale vaardigheden in SLI en tweetaligheid ['Frog, where are you?' Grammatical and lexical skills in SLI and biligualism]. *Logopedie en Foniatrie*, 12, 392–398.
- Van den Heuvel, H., Sanders, E., Klatter-Folmer, J., van Hout, R., Fikkert, P., Baker, A., Jong, J. de, Wijnen, F. & Trilsbeek, P. (2014). Data curation for a VALID Archive of Dutch Language Impairment Data. *Dutch Journal of Applied Linguistics* 3(2), 127–135, doi: 10.1075/dujal.3.2.02heu
- Vandewalle, E., Boets, B., Boons, T., Ghesquière, P., & Zink, I. (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. doi: 10.1016/j.ridd.2012.05.004
- Verhoeven, L., Steenge, J., & Van Balkom, H. (2011). Verb morphology as clinical marker of specific language impairment: Evidence from first and second language learners. *Research in Developmental Disabilities* 32, 1186–1193. doi: 10.1016/j.ridd.2011.01.001
- Zwitsersloot, R., van Weerdenburg, M., Verhoeven, L., Wijnen, F. (2015). Development of morphosyntactic accuracy and grammatical complexity in Dutch school-age children with SLI. *Journal of Speech, Language, and Hearing Research* 58, 891–905. doi: 10.1044/2015\_JSLHR-L-14-0015

## Appendix

### (1) *The SLI RU-Kentalis Database*

Informants: 63 SLI + 24 controls; Characteristics: 56 boys and 31 girls ; 5 – 12 years old; Specific Language Impairment (SLI); Aim of data collection: investigation of the expression of spatial relations by children with SLI and typically developing children in their spoken language production; Materials available: Tests: Raven, WISC (Block Pattern, Mazes), Peabody: SPSS data files; Photo/Film Task – audio recorded: transcript in Praat; data processed and coded in SPSS; Route Description Task – audio/video recorded: transcript in Praat; data processed and coded in SPSS; TAK Narrative – audio recorded: transcript in Praat; data processed and coded in SPSS; Frog Narrative – audio recorded: transcript in Praat; data processed and coded in SPSS; Size: 1st measure 13.9 GB, 2nd measure +/-13.9 GB (Photo/Film); 18.9 GB (Route Description); 1st measure 1.75GB, 2nd measure 1.46 GB (TAK); 3.10 GB (Frog); 1 MB (test and background data).

### (2) *The UU SLI-Dyslexia Project Database*

Informants: two longitudinal cohorts: (a) babies, from 19 months to approximately 37 months; N ≈ 110. (b) toddlers; 3;2 (years; months) at the onset; about 5;0 at the last test session; N ≈ 140; Characteristics: baby cohort: ~70 children at familial risk (FR) of dyslexia; ~40 controls; toddler cohort: ~70 FR children, ~40 controls, ~30 children (tentatively) diagnosed with Specific Language Impairment (SLI). Children from both cohorts have returned to the lab at age 8 for follow-up tests: toddler cohort: n = 107; baby cohort: n = 65; Aim of data collection: to explore early language development in children at FR of dyslexia; to compare developmental language profiles in FR children and children with SLI. Materials available: preferential listening experiment; measurement of listening times in several trials; categorical perception experiment; tests of productive phonology (elicited naming); various procedures (book reading; card matching); digital recordings of speech, (partly) transcribed in IPA and coded for phonological errors; word–picture matching experiment;

eye gaze to corresponding pictures (one out of two per trial) was recorded; lexical decision experiment: words (presented in combination with pictures) were correctly or incorrectly pronounced (phonemic errors); many speech elicitation experiments (various designs; digital audio recording; partial transcriptions); auditory grammaticality judgement task; all coded responses in Excel / SPSS formats; WISC digit span task; Snijders-Oomen nonverbal intelligence test; N-CDI's: standardised communicative development inventory, completed by participants' parents; Size: raw estimate of 60 GB.

### *(3) The Bilingual Deaf Children RU-Kentalis Database*

Informants: 11 deaf children, longitudinal; Characteristics: 5 boys and 6 girls ; 3 – 6 years old; prelingual deafness (hearing loss of minimally 80dB Fletcher Index on the best ear), no mental restrictions; Aim of data collection: investigation of the bilingual language and communication development of young deaf children in Sign Language of the Netherlands (SLN) and Dutch (D);

Materials available: Tests: Nijmeegse Observatieschaal voor Kleuters (NOK; SLN & D), Reynell Test voor Taalbegrip (SLN & D), Dutch version Assessing British Sign Language Development (SLN): data processed and coded in SPSS; Spider Story (SLN & D): data processed and coded in SPSS; Semi-structured conversations with deaf and hearing adults – video recorded (SLN & D): a selection of five minutes communication per recording has been selected and transcribed in a CHAT-like format (104 recordings); Size: 4 GB complete video recordings; 1 GB selected parts video recordings; 0.1 GB selected parts transcripts; 0.5 GB test and background data.

### *(4) The ADHD and SLI Corpus UvA Database*

Informants: 26 Dutch children with ADHD, 19 Dutch children with SLI, 22 Dutch children controls; Characteristics: ages between 7 and 8 years; 80% male, 20% female; intelligence within normal ranges; Aim of data collection: to compare the language and executive functioning profiles of children with ADHD to that of children with SLI and TD children; Materials available: Tests: Sentence repetition task; Non-Word repetition task; Frog story narratives, processed in SPSS on morphological, syntactic and pragmatic measures; Children's Communicative Check-list II; CANTAB EF tasks for executive functioning; Size: 4 GB (67 recordings).

### *(5) The Deaf Adults RU Database*

Informants: 46 deaf Dutch adults, 38 hearing Turkish adults, 24 hearing Moroccan adults, 10 Dutch controls; Characteristics: males: 22 deaf + 31 Turkish/Moroccan + 5 controls; females: 24 deaf + 31 Turkish/Moroccan + 5 controls; Aim of data collection: investigation of the acquisition of Dutch by deaf Dutch adults (late L1/early L2) and comparison to hearing Turkish and Moroccan-Arabic L2-learners of Dutch (late L2) on morphosyntactic aspects; Materials available: Test: standardised C-test Instaptoets Anderstalige Volwassenen (IAV); coded and processed in SPSS; Writing task The Frog Story: recorded and stored in ScriptLog (Holmquist), data coded and processed in Excel and SPSS; Size: 2 GB.