Speech and Language Sciences Section
School of Education, Communication & Language Sciences

# The Morphology of German-speaking Children with Specific Language Impairment

Dorothea Juliane Hasselaar

This thesis is submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy
November 2014

Declaration of Originality

The material presented in this thesis is the original work of the candidate except as otherwise acknowledged. It has not been submitted previously in part or in whole, for any award at any university, at any other time.

## Acknowledgements

My thanks go first to the children and their parents who participated in this research. It was them who made this study possible. I enjoyed playing and working with the children and am grateful that they co-operated so enthusiastically and full of patience.

Special thanks go also to the speech and language therapists who supported me in my search for participants and who let me use so generously their locations. I would like to thank in particular the Werner-Otto-Institut in Hamburg, as well as Inula Groos, Nina Kuhn-Wierzbinski and Karen Grosstück for their help and for their flexibility in providing me the possibility to work part-time. I am also grateful to the nurseries that I visited to collect data for their collaboration and their hospitality.

I am particularly grateful for the endless support of my principal supervisor Carolyn Letts. Her patience and encouragement helped me to keep going through all ups and downs of the study and writing up process. Her insight and feedback was very stimulating and always motivating.

Many thanks go to Cristina McKean for her supervision and her ability to spot the most relevant issues of my thesis and helping me to keep the focus. Her questions and comments stimulated me to think more theoretically. Thanks also to Thomas Klee for his supervision in the beginning of my study. He was always willing to answer methodological questions and to provide help with statistics. Thank you to David Howard for giving his time and expertise for statistical challenges and for his encouragements.

This research project has been funded with an ESRC grant and a bursary of the Speech and Language Sciences Section of Newcastle University. I am very grateful for this financial support.

Thank you to Annette Fox-Boyer for sharing her enthusiasm in studying and living in Newcastle. I wouldn't have even thought about doing a PhD in

Newcastle without her and am very happy that she infected me with the ‘Newcastle virus’.

Thank you to my friends and family for their patience and encouragements whilst I kept apologising instead of taking part in activities. I am looking forward to the ‘new’ life without thesis.

Particular thanks go to Els and Evert Hasselaar who supported me wherever they could by entertaining their grandson, proofreading and finally even helping out with formatting while I had the chance to get to know their new little grandson.

Thank you to my parents and my family! They made me the person that I am and supported me in many, many ways. At the same time they showed me that academic success is not the essential in life. Sadly, my Mum cannot experience the accomplishment of the thesis anymore but I am glad that the PhD study allowed the flexibility to spend as much time as possible with her before she passed away.

Finally, thanks to Tjalle and Thijs. To Tjalle for accepting a Mum who kept disappearing behind the computer each weekend during the last months, and to Thijs for making me feel less guilty knowing that Tjalle had a great time with his fabulous Dad. Thijs, thank you for your patience, endless support and love! You are by far the best outcome of doing a PhD in Newcastle!

# Abstract

## The Morphology of German-speaking Children with Specific Language Impairment (SLI)

### Purpose

Reliably identifying children with specific language impairment (SLI) remains challenging. Difficulties in morphosyntax have been proposed as potential linguistic markers for SLI across a number of languages. This study sought to investigate the existence of such a clinical marker in German-speaking children with SLI. A potential clinical marker should discriminate successfully between children with SLI and their typically developing peers, but should also stand out from the children's general language difficulties. It is also essential for clinical as well as theoretical purposes, that such a marker represents grammatical rather than phonological impairments (PI), frequently concomitant with SLI.

### Method

A case-control study was completed with 66 children in four groups (1) SLI, (2) PI, (3) age matched typically developing children (ND-A) (4) language matched typically developing children (ND-L). Errors in morphosyntax were analysed from elicitation tasks and spontaneous language samples.

### Results

The performance of children with SLI was poorer than the ND-A group for present tense inflections and case marking. When compared to the ND-L group, the SLI group only differed in spontaneous production of nominative articles and a composite measure of present tense verb inflections; a difference also present compared to the PI group. Production accuracy on this composite is therefore a potential candidate maker for German SLI. Error analysis revealed that the impact of phonological errors is considerable for children with PI, underlining the need to differentiate between grammatical and phonological errors.

**Discussion**

These findings bring new insights into the development of German morphology in SLI, and hence into our understanding of the disorder. The validity and reliability of the use of clinical markers in research and practice is discussed in the light of heterogeneity and qualitative differences found between the participant groups.

# Table of Contents

## List of Tables

# List of Figures

## Nomenclature

| | |
|---|---|
| SLI | specific language impairment |
| PI | (isolated) phonological impairment |
| ND-A | normally / typically developing, age-matched children |
| ND-T | normally / typically developing, language-matched children |
| MLU | mean length of utterances |
| MLU-W | mean length of utterances in words |
| MLU-M | mean length of utterances in morphemes |
| SETK | Standardised language test SETK 3-5: Sprachentwicklungstest für drei- bis fünfjährige Kinder (Grimm, 2001) |
| ELFRA | Standardised parental questionnaire: Elternfragebogen für zweijährige Kinder (Grimm & Doil, 2000) |
| PLAKSS | picture naming task for diagnosis of speech disorders: Psycholinguistische Analyse Kindlicher Sprechstörungen (Fox, 2002) |
| VMI | Developmental Test of Visual-Motor Integration (Beery & Buktenica, 1989) |

.

# 1 Introduction

The PhD thesis focuses on the characterisation of children with specific language impairment (SLI). Research on specific language impairment (SLI) in children is predominantly driven by the motivation to improve our understanding of the phenomenon of language. The focus on this population allows such conclusions about language processing because the children's deficit is thought to be *specific* to language; i.e. children with SLI are considered to experience language difficulties in the absence of another deficit that may cause the language impairment. The study of children with SLI as one group involves a number of decisions over the nature of SLI as illustrated in Chapter 2. These may eventually lead to a better understanding of how language is acquired and how it may be interrupted. However, each of these decisions bears uncertainties and these challenge the presumption of children with SLI as one distinct population. If children with SLI do not turn out to be a distinct population, research on this rather artificial construct would be unlikely to lead to conclusive answers with respect to language acquisition and possible interruptions.

The present study attempts to characterise SLI in German-speaking children in order to test the homogeneity of the population. Starting point of the research project is therefore the assumption that children with SLI are indeed distinct from typically developing children and if so, it should be possible to describe these children with one or few key-characteristics, a so called clinical marker. So far, research on SLI has primarily focussed on English-speaking children with SLI and several candidates were proposed as clinical marker for SLI in this language. However, the identification of a clinical marker that is universal across languages would be of much greater significance because it would be an important step forward in the search for an underlying cause of SLI and it would furthermore justify with hindsight the approach to treat children with SLI as a distinct population.

The presented research tackles two further issues in the decision making process on SLI. With the decision to study children with SLI as one group, researchers imply that there are no valid subgroups of SLI with different underlying aetiologies. One clinical group that often overlaps with SLI are children with an isolated phonological impairment (PI). If this group of children is distinct from children with SLI a clinical marker for SLI should reflect this fact. However, the distinction between phonological and grammatical errors is of high importance in this respect. If children with an isolated phonological impairment, on the other hand, are a subgroup of the SLI population it is questionable whether the population of children with SLI is a construct caused by a single origin.

**Structure of the thesis**

This thesis is organised into seven chapters. Chapter 2 sets the context for the remaining chapters by illustrating how decisions around the construct SLI influence each other. Chapter 2 introduces the concept of SLI including symptoms and possible subclassifications. It further describes approaches to define a clinical marker for SLI and their role on the development of explanatory theories of SLI. Some of these theories are briefly summarised. Finally, Chapter 2 discusses the overlap between SLI and PI and the impact speech errors may have in the identification process of SLI.

Chapter 3 provides the rationale for the present study that is derived from the literature review. The chapter further states the detailed aims of the thesis.

The ‘methods’ chapter, chapter 4, describes and rationalises the general methodological procedures employed, the participants of the study and the analyses carried out.

Chapter 5 describes the verb morphology in German-speaking children with SLI. The chapter illustrates first linguistic characteristics of the verb phrase in German, and reviews then the literature on verb morphology in German SLI. Chapter 5 illustrates further the methodological approach applied in the present study in order to identify a clinical marker for SLI in the children's use of present tense inflections and presents the empirical data from elicitation tasks and spontaneous language samples. Comparisons are made to typically developing control children as well as to children with PI. The potential of the investigated variables as clinical marker is discussed with regards to significance of delay, deviant error patterns and specificity to SLI.

Chapter 6 focuses on case marking in German-speaking children with SLI. It follows a similar structure as chapter 5. The chapter provides first an overview over the case marking system in German and reviews relevant literature on case marking in German-speaking children with SLI. The chapter details then the methodological approach used in the investigation of nominative, accusative and dative marking in German SLI. Results from elicited and spontaneous data are presented and their potential with respect to a clinical marker discussed.

Chapter 7 discusses the empirical data presented in a wider context. The chapter follows three main lines regarding methodological issues arising from the present data collection and analyses, regarding the relation between SLI and PI and finally regarding the validity of the clinical category SLI. This last section ends with recommendations for further approaches to the population of primarily language-impaired children.

Chapter 8 summarises the main outcome of the present study and the recommendations with regards to further research.

# 2 Specific Language Impairment: A Clinical Category

## 2.1 Introduction

Traditionally, specific language impairment (SLI) is seen as primary language disorder. The goal of this thesis is to characterise the morphology of German-speaking children with SLI. This aim involves the investigation in how far such a linguistic characterisation can discriminate children with SLI from another developmental disorder, an isolated phonological impairment (PI). These aims are derived from the assumption that SLI can be seen as a distinct population. The specificity of the children's developmental problems to the language domain has attracted a lot of attention because research on SLI can provide very valuable information on how language develops, which processes interact with language acquisition and how or where it can go wrong. This information could possibly help to prevent language problems and help to develop more efficient treatment concepts. However, Figure 1 illustrates that the assumption of the specificity of SLI is based on a number of decisions impacting one on another. Each of the decision hierarchy's levels raises fundamental questions in the definition of SLI. The present chapter will introduce the concept of SLI following the decision hierarchy: How is SLI defined? Are there valid subgroups of SLI? Are children with SLI distinct from children with an isolated speech output disorder or should children with an isolated PI be classified as SLI, too? Are there distinct symptoms that characterise the population, i.e. a clinical marker for SLI? What are the implications of the findings regarding the aetiology of SLI? What deficits can account for the children's language impairment?
Figure 1 illustrates the path that has often been chosen in SLI research. As highlighted in the hierarchy, the present study offers the opportunity to challenge especially two levels of this path: Firstly, the existence of a distinct

characteristic for SLI, and secondly, the categorisation of children with PI. The answers to these questions, though, impact on the other levels of the decision hierarchy, too. On the one hand side, results could contribute to the validation of the category SLI or, on the other hand, question whether research on a rather artificial construct can lead to conclusive answers with regards to the remaining levels of the hierarchy.

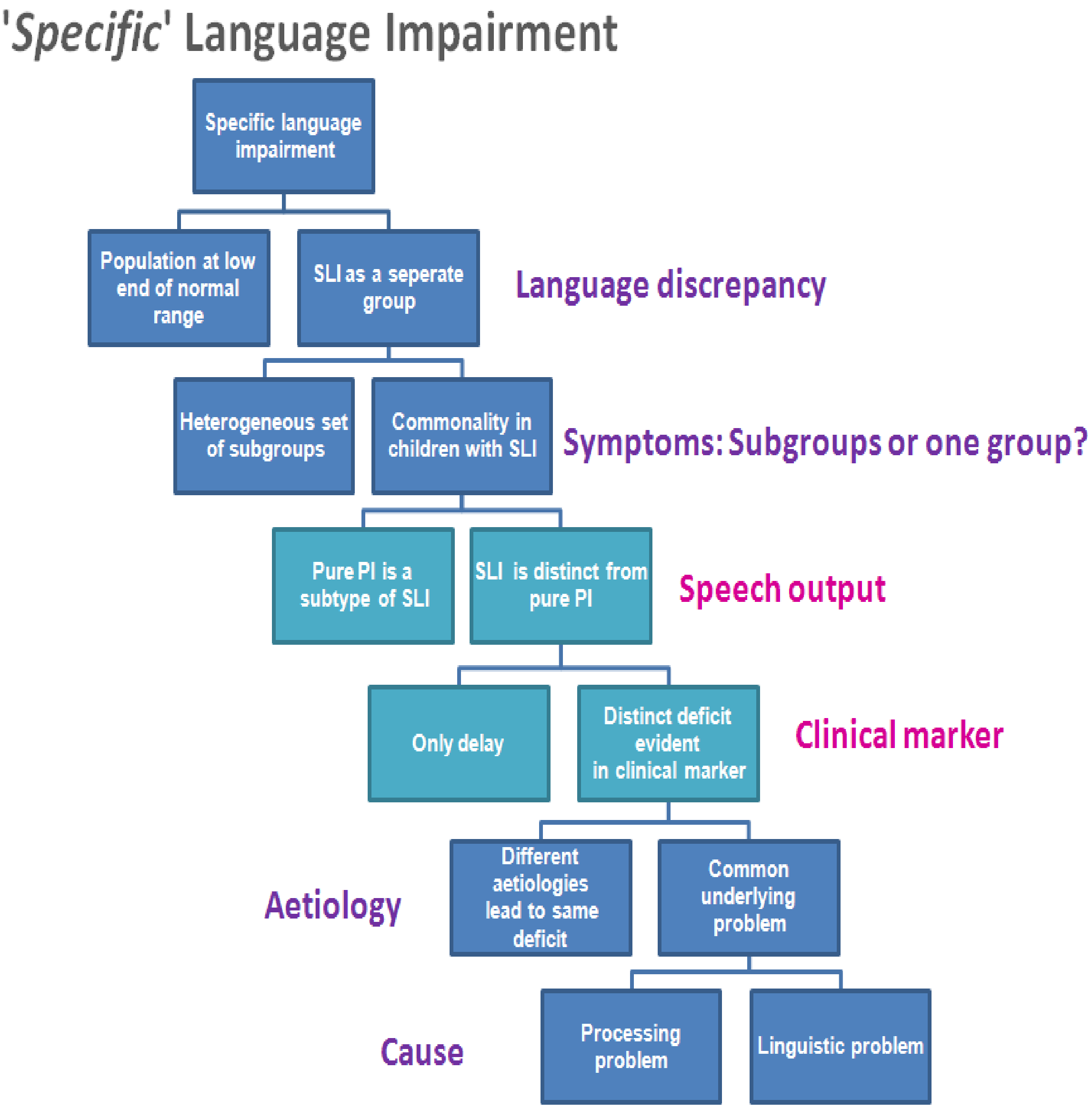



**Figure 1.:** Hierarchy of decisions about the nature of SLI.
Adapted from de Villiers (2003, p.427)

### *Identifying SLI*

This section familiarises the reader with the concept of SLI. Section 2.2 starts with an overview of the defining characteristics of SLI, including the language discrepancy and exclusion criteria. Section 2.3 provides a short overview of symptoms in SLI followed by considerations on subtypes of SLI in section 2.4. and the role of phonological impairment in SLI in section 2.4. and 2.6.. Across different languages, section 2.5 summarises the empirical evidence searching for a clinical marker on SLI. This section sketches the research history on SLI and includes proposals presented in order to explain the findings in a theoretical context.

## 2.2 Definition

Specific language impairment in children has traditionally been defined as a developmental language disorder that occurs in the absence of any identifiable cause. It is characterised by a delayed onset and protracted language development relative to other areas of development and it is normally identifiable from the age of 3 years onwards (Tager-Flusberg & Cooper, 1999). The prevalence of SLI lies around 7% of the general English-speaking US population (J. B. Tomblin et al., 1997) with males more likely to be affected than females (Law, Boyle, Harris, Harkness, & Nye, 2000; Leonard, 1998; Tomblin et al., 1997). Similar prevalences are expected for the German-speaking population although only regional empirical data is available (Langen-Müller, Kauschke, Kiesel-Himmel, Neumann, & Noterdaeme, 2011). Especially, if these language problems remain untreated SLI can have significant consequences throughout the life span. Only approximately 25% of the cases can be expected to resolve spontaneously (Law, Boyle, Harris, Harkness, & Nye, 2000) and the children are at risk of experiencing reading and other academic difficulties (e.g. Conti-Ramsden, Durkin, Simkin, & Knox, 2009; Leonard, 1998; Snowling, Bishop, & Stothard, 2000) as well as experiencing psychological problems later in life (e.g. Beitchman, Brownlie, Inglis, Wild, & et al., 1996; Durkin & Conti-Ramsden, 2010). Although in research literature limited uniformity exists in defining characteristics of SLI most studies have in common the exclusionary nature of their description rather than inclusion criteria. Accordingly, children with SLI seem to develop normally except that they experience significant limitations in their language ability. SLI is thus an impairment that is *specific* to language.

### *2.2.1 Criteria for inclusion: The language discrepancy*

While early research on SLI focussed on clinically identified children, i.e. children in need of language intervention, subsequent inclusion criteria inclusion criteria usually refer to the severity of the language disorder. This requires therefore an evaluation of the discrepancy between the child’s achieved language status and some reference for the expected language status. This reference measure has changed over the last decades. One of the earliest suggestions were the criteria for SLI suggested by Stark and Tallal (1981) relating the children’s language age to their mental age. However, a number of limitations have questioned both the reliability of MA over time as reference measure (e.g. Krassowski & Plante, 1997) and the reliability of calculations based on overall scores such as language age or mental age. These limitations stimulated a different approach.

A review of research on English-speaking children with SLI published in 2003 and 2004 shows that most of the more recent work has been based on standardised scores derived from the normal distribution rather than on age equivalent scores, although some studies did not provide any specific criteria for inclusion (Heilmann, 2004 cited in Miller & Fletcher, 2005; Pena, Spaulding, & Plante, 2006a). The cut-off criteria in the use of standard deviations (SD) ranged from -1 SD to -1.5 SD. In 1996, Tomblin, Records and Zhang developed the diagnostic system EpiSLI and obtained the best results regarding sensitivity and specificity with a cut-off of –1,25 SD in at least two of five composite language measures: (1) vocabulary (receptive and expressive), (2) grammar (receptive and expressive), narrative (receptive and expressive), comprehension (vocabulary, grammar, and narrative) and language production (vocabulary, grammar, and narrative). A similar cut-off point is often recommended in the literature (e.g. Leonard, 1998).

This approach has been questioned by authors who suggest that SLI may not correspond to a distinct clinical category but only to the lower end of the normal curve (e.g. Leonard, 1991; J. B. Tomblin & Zhang, 2006; Weismer, 2007). As Figure 1 illustrates, this challenges not only the inclusion criteria of SLI but also the meaning of SLI research regarding the following levels of the decision hierarchy. Dollaghan (2004b) studied a large number of 3- and 4-year-olds in order to investigate the usefulness of a cut-off point on standardised measures drawing a line between typical and impaired language development. In fact, the author failed to find evidence for a qualitatively distinct group of children with SLI relative to the normally developing children. Nevertheless, these findings may result from a limited scope of measures since Dolloghan focussed mainly on measures of vocabulary. If the qualitative differences between the clinical and the normal population do not fall into this area, results would have been unlikely to identify SLI as distinct group. Tomblin, Zhang, Buckwalter, and O'Brien (2003b) followed a large number of children with primary language impairment (selected according to the EpiSLI criteria) from pre-school years up to the 4th grade. Across this period the authors found that the diagnosis of language impairment remained rather robust. Thus, a dimensional cut-off for standardised language measures seems still to be reasonable at present in order to provoke the search for qualitative differences that accompany the quantitative differences between SLI and normal language development. Such qualitative characteristics could eventually improve the identification of the affected children.

### *2.2.2 Criteria for exclusion: Specific to language*

Driven from the assumption that SLI is an indication for language being largely autonomous of other mental faculties exclusionary criteria are generally acknowledged in the definition of SLI. An array of developmental clinical conditions of genetic, physical, social or emotional origin has been recognised as showing language disorders as characteristic symptoms. Children whose language problems are thought to be the result of such conditions are therefore excluded from the group of children with SLI. The following criteria are usually considered in the definition of SLI in order to distinguish children with primary language problem from children with secondary language problems:

- Nonverbal IQ of 85 or higher
- Passed hearing screening at conventional levels (usually 25dB) and no long-term recurring otitis media with effusion
- No frank neurological deficit in the present or past including cerebral palsy, head trauma or epilepsy
- No peripheral oral motor or sensory deficits or oral facial anomalies
- No symptoms of severe emotional or behavioural problems

The appropriateness of the criterion excluding children with below average non-verbal IQ from the SLI population has often been discussed in the literature, though. As mentioned above, research showed that the non-verbal IQ of children with SLI changed over time and therefore cannot be considered as reliable measures (e.g. Krassowski & Plante, 1997) and that the outcome of the IQ test is likely to depend on the chosen measure (Swisher, Plante, & Lowell, 1994). Consequently, whether a child may meet the SLI criteria may depend on the time of testing or the chosen test. The cut-off of a non-verbal IQ of 85 has further been criticised on the basis that children just failing this cut-off are not distinct from children with SLI regarding their language profile or response to intervention (Fey, Long, & Cleave, 1994; J. B. Tomblin & Zhang, 1999). Botting (2005), on the other hand, provided longitudinal data that did demonstrate an interplay between nonverbal intelligence and language development. Children who showed a decline in their nonverbal IQ showed also significantly lower language performance at an age of 14 than those children with a stable nonverbal IQ. These results confirmed earlier findings from Goorhuis-Brouwer and Knijff (2001) showing that children with age-appropriate nonverbal IQ benefited more from language intervention than children with a cognitive delay. Concluding from these findings, the consideration of nonverbal skills may have some flaws but seems relevant when addressing language development.

## 2.3 Symptoms of SLI

Children who exhibit a significant deficit in language ability independently of the other developmental factors listed above have been identified in many different countries and languages (Leonard, 2000). The individual problems of children with SLI can vary, but some of the universal hallmarks seem to be a late beginning of language development and subsequent lexical problems such as in word storage and word access. Furthermore, the acquisition of word combinations tends to need more time than in normal developing children (Leonard, 2000; Trauner, Wulfeck, Tallal, & Hesselink, 2000).

Besides the delayed lexical development, SLI is often associated with difficulties in morpho-syntax, especially the comprehension and production of grammatical morphemes in both English (e.g. Bedore & Leonard, 1998; Rice, 2003; Rice & Wexler, 1996) as well as languages other than English (e.g. Bortolini, Caselli, Deevy, & Leonard, 2002; Clahsen, Bartke, & Göllner, 1997; Crago & Paradis, 2003; De Jong, 1999; Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999; Leonard, 2000). Affected linguistic structures and error forms, however, vary across languages due to the individual linguistic characteristics of the respective language (Crago & Paradis, 2003; Leonard, 1998, 2000). In any investigated language, children with SLI differ in the degree to which they produce particular morphemes, however, the children usually do not omit the affected morphemes all the time (Leonard, 2000).

Additional pragmatic problems may be apparent secondary to the language impairment. Yet, an isolated pragmatic disability is often seen as part of the autistic spectrum, or as own diagnostic entity rather than as SLI (Bishop, 2000). As regards speech difficulties, phonological problems often co-occur with SLI (e.g. Bortolini & Leonard, 2000; Maillart & Parisse, 2006; Stackhouse & Wells, 1997).

## 2.4 Subgroups of Specific Language Impairment

Despite these commonalities found in children with SLI, the group of children meeting the criteria SLI is a heterogeneous population (e.g.Leonard, 1998; Schwartz, 2009; Tager-Flusberg & Cooper, 1999; Verhoeven & van Balkom, 2004). Children with SLI differ from one another in the number of language areas affected, the severity of language deficit, the area of language forming the greatest difficulty and their receptive abilities in relation to their expressive language skills. The search for stable subgroups within this population has consequently been addressed frequently as the division into valid subgroups would provide valuable potential for future research. Studies on subgroups of SLI can be classified roughly into three different approaches: a clinical, a psychometric and a linguistic approach.

Clinical subtyping procedures are based on clinical judgements of the children's therapists or teachers. One of the best known studies of this type is the classification system of Rapin and Allen (1987). A number of other researchers used a psychometric approach and tried to group the children according to their profiles in standardised tests using statistical procedures (e.g. 1975; Wolfus, Moscovitch, & Kinsbourne, 1980). Wilson and Risucci (1986) as well as Conti-Ramsden, Crutchley and Botting (1997) combined their statistical cluster analysis of psychometric data with clinical data. Most authors covered receptive and / or expressive tasks in all language areas, but a number also included general processing tasks (Conti-Ramsden et al., 1997; Korkman & Häkkinen-Rihu, 1994; van Weerdenburg, Verhoeven, & van Balkom, 2006; Wilson & Risucci, 1986). Linguistic subgrouping procedures follow the attempt to relate linguistic data to a linguistic framework. Fletcher (1992), for example, analysed spontaneous language data from thirty children with SLI and classified the children according to difficulties related to the different modules of the language production model by Levelt (1989).

It would be desirable that the evidence from classification studies on SLI was sufficient to rely on these results for future participant selection. Difficulties in comparing results across studies arise from a considerable variation among the used approaches, though. The obtained subgroups are highly dependent on the nature of data (clinical observations, psychometric tests, spontaneous speech analysis), the specific aspects information focused on in regard to the children's abilities (language production, language comprehension, perceptual / processing / cognitive skills), and the age range of the respective study.

*Expressive-receptive SLI*

Nevertheless, there is one general consensus among the studies reviewed here: all studies identified at least one group of children with SLI experiencing problems in grammar, semantics and phonology which were mirrored in both expressive and receptive language (Aram & Nation, 1975; Conti-Ramsden et al., 1997; Fletcher, 1992; Korkman & Häkkinen-Rihu, 1994; Rapin & Allen, 1987; van Weerdenburg et al., 2006; Wilson & Risucci, 1986; Wolfus et al., 1980). Difficulties in syntax and phonology seem often to be especially pronounced within this group. It appears that children with a language impairment affecting predominantly semantics and / or pragmatics were often described as experiencing a receptive language impairment (Aram & Nation, 1975; Conti-Ramsden et al., 1997; Fletcher, 1992; Korkman & Häkkinen-Rihu, 1994; Rapin & Allen, 1987; van Weerdenburg et al., 2006; Wilson & Risucci, 1986). However, other researchers claim that receptive only difficulties are rare and occur neither in 'late talkers' nor in 'early talkers' (Bates, Dale, & Thal, 1995) nor in SLI (Archibald & Gathercole, 2006b). It is therefore unlikely that language production exceeds language reception in any respective language area. Expressive difficulties in semantics and pragmatics may nevertheless be easier to compensate for than syntactic or phonological difficulties and therefore not always be identified. Also, not all categorisation studies included explicit measures referring to expressive semantic or pragmatic skills (Conti-Ramsden et al., 1997; Fletcher, 1992; Wilson & Risucci, 1986). It can thus be concluded that a receptive language deficit can be considered as receptive-expressive language deficit and can affect any language area. This conclusion is reflected in the description of specific developmental disorders of language by the World Health Organisation (WHO) in the ICD-10 (World Health Organisation, 2010). Although the ICD-10 lists receptive language disorder (F.80.20) its further description illustrates that this subcategory is always supplemental to an expressive language disorder.

*Expressive SLI*

The categories the most difficult to match across studies are those referring to expressive deficits and a phonological deficit. The main problem by comparing the groups categorised as experiencing expressive language difficulties is that no distinct language domain has been referred to. Some problems concern lexical-semantic skills (Conti-Ramsden et al., 1997; Fletcher, 1992; van Weerdenburg et al., 2006; Wilson & Risucci, 1986) whereas other groups are reported to have difficulties in morpho-syntax (Aram & Nation, 1975; Conti-Ramsden et al., 1997), in phonology or articulation (Aram & Nation, 1975; Conti-Ramsden et al., 1997; Rapin & Allen, 1987), or even in both morpho-syntax and phonology (Wilson & Risucci, 1986).

*Speech output disorder*

Some authors suggested classifying children with major problems in phonology as a distinct subtype of SLI (Aram & Nation, 1975; Conti-Ramsden et al., 1997; Rapin & Allen, 1987). In the more recent literature on developmental speech and language disorders, a distinction is commonly made in developmental speech problems between articulatory or phonetic and phonological problems (e.g. Stackhouse & Wells, 1997; World Health Organisation, 2010). The term *articulatory* difficulty is used if children have difficulties with the *production* of a particular sound, independent of its position in the word. The production of a sound depends on the appropriate movement of the articulators. *Phonological difficulties*, on the other hand, refer to the children's ability to *use* the sounds in their correct linguistic context irrespective of the children's ability to produce the sound. Another term that is also used in the literature on classification of SLI reviewed above is *verbal dyspraxia.* Usually, it has been seen as breakdown at the level of motor programming (Dodd, 1995; Stackhouse & Wells, 1997). However, it apparently has sometimes been used as synonym for any kind of speech output disorder, including phonological impairment.

Strictly speaking, only children with phonological problems should meet the criterion of SLI because neither motor planning nor the movement of articulators are based on linguistic skills. More recently, even children with phonological impairment have often been excluded from studies on children with SLI (Schwartz, 2009). It is therefore debatable whether reduced intelligibility of speech resulting from a phonological impairment should be considered as language impairment or as speech disorder. This question will be followed up below in section 2.5.

In summary, the only broad sub-classification which is possibly based on the literature review is the distinction between *expressive-receptive SLI* (or global SLI) and *expressive SLI*. This distinction has indeed recently been used by a number of researchers (e.g. Evans, 1996; Simkin & Conti-Ramsden, 2006) and reflects further the distinction made in the ICD-10 (World Health Organisation, 2010). However, as Leonard argues (Leonard, 2003), subgroups identified for one sample of children with SLI tend to blur or change when applied to a new sample. Conti-Ramsden et al. (1997) followed their sample longitudinally. However, although the same groups were identified again after two years, 45% of the 242 individual participants changed from one to another group within this period. Since this data shows that individual children are likely to change their language profile the suggested distinction between expressive and expressive-receptive SLI seems premature for present research. Eventually, a sub-classification system may be useful if the categorisation takes the children's age into account, too, as proposed by Schöler, Fromm and Kany (1998b). However, to date, further research is necessary in order to identify such a valid and stable classification system. Ideally, this research should include large samples of children clearly diagnosed with SLI, but with a limited age range, and follow the children's changes in longitudinal studies.

## 2.5 Clinical Markers of SLI

Researchers' interest in the aetiology of SLI enhances the importance of an exact identification of this developmental language disorder. A large number of researchers have been searching for key behavioural features that may be useful in the identification process. For any identified condition, the establishment of clinical markers plays an essential role in health sciences for the diagnosis of affected individuals. Classically, research on the identification of clinical markers tries to differentiate between affected and not affected individuals. A clinical marker for SLI, thus, would represent key characteristics in the way that affected children with SLI vary from other children. In order to account for the heterogeneity of the target population, it may be promising to look for key characteristics for different subgroups of SLI respectively. However, the review in Section 2.4 on subgroups of SLI demonstrated that this would be premature up to date. Yet, the discovery of a clinical marker for the population of children with SLI taken as a whole would firstly underline the assumption of specificity in SLI and could secondly contribute significantly to our understanding of this language disorder as well as on screening and intervention methods. The literature review below presents the most important findings in clinical marker research for SLI and illustrates their impact on the development of SLI theories. As they reflect changes in the view of language acquisition, the findings are placed into theoretical frameworks of language development

### *2.5.1 Identification of a clinical marker*

Over the last two and a half decades, several measures have been proposed as potential clinical marker for children with SLI. These measures can broadly be divided into two domains: linguistic markers, such as the production of various morphemes, MLU or lexical measures (e.g. Bedore & Leonard, 1998; Klee, Stokes, Wong, Fletcher, & Gavin, 2004; Rice, 2003), and processing markers, such as the performance on tests of working memory (e.g. Bortolini et al., 2006; Conti-Ramsden, 2003; C. Dollaghan & Campbell, 1998). Regarding their statistical evaluation of the potential clinical markers, researchers followed one of two different paths.

*Delay within delay*

The first approach provides evidence based on comparisons of group mean differences, i.e. a group of children with SLI is compared to a control group in their performance on a given variable. In most studies, at least one control group consists of normally developing children who are age-matched to the children with SLI. The performance of the study group can thus be related to age expectations. Further control groups may be included such as younger, for example language-matched, typically developing children or children with language impairment that is associated with other developmental conditions such as a syndrome or limited cognitive skills. The main motivation for choosing a 3-group design with age- as well as language-matched (or generally younger) typically developing children is to see in which way the affected children vary in their language performance from typically developing children. The evidence that has been obtained in this way allows delay versus deviance interpretations. Differences between the affected and age-matched control children point to a *delay* in the measured language area. However, differences between affected children and language-matched typically developing children pinpoint to a language profile that resembles that of younger normally children in the language area the groups were matched on but is more comprised in other areas (Rice, 2003). In these particular areas, the children with SLI are *deviant* from typically developing children. Rice (2003) proposes a model of *delay-within-delay* in which some elements of the linguistic system may be selectively delayed in relation to other linguistic elements. These elements may be potential candidates for a clinical marker.

*Discriminant analysis*

The second approach employed in the search for clinical markers is based on discriminant analysis which does not only tests the differentiation between the clinical and control group on grounds of a proposed clinical marker but it also calculates as a second step the identification accuracy of the marker. Thus, results of discriminant analyses usually reveal the percentage of affected children who were correctly classified by the marker (sensitivity) and the percentage of children without SLI who were correctly tested negative on the marker (specificity). Although there are no generally applied thresholds, some researchers consider sensitivity as well as specificity levels of 80% as acceptable, but only levels of 90% or higher as good (e.g. Bortolini et al., 2006; Plante & Vance, 1994). Calculated likelihood ratios go even beyond the consideration of the whole population and are independent of the condition's prevalence. Likelihood ratios reflect instead the degree of confidence that an individual person is classified correctly (e.g. C. A. Dollaghan, 2004a). Thus, positive likelihood ratios (LR+) indicate how likely it is that a person who was tested positive on a marker is truly affected, while negative likelihood ratios (LR-) reflect the likelihood that a person not identified as affected by the marker is free of the condition, or in this case is developing language normally. The higher LR+ and the lower LR- the better is the accuracy a given marker classifies both affected and not affected individuals. An LR+ value higher than 10 is considered as good (Klee, Gavin, & Stokes, 2007). Dollaghan states that a LR+ value of 20 or more indicates that an individual is nearly certain to be affected by the tested condition if he / she has a positive test result. A LR- value, on the other hand, is considered to be good if lower than 0.20 but it can basically be taken for granted that the individual is free of the condition if the measure's LR- value is 0.10 and the test result negative. Many researchers did not restrict their focus to one potential marker but examined the usefulness of several different potential markers in parallel. Thus, some studies could even investigate even the accuracy of possible combinations of markers in the identification of children with SLI. The markers below are nevertheless introduced according to their focus on processing and linguistic skills respectively.

### *2.5.2 SLI within a nativist framework*

For several decades, SLI has been considered within a nativist framework based on Chomsky's suggestion of an innate grammatical knowledge (1959). The specificity of the children's impairment to the domain language is for that reason an important pre-requisite. Considering children with SLI as one clinical category supports thus this modular view of language acquisition being regulated by domain-specific processes. Consequently, the root of a *specific* language impairment should lie within the domain of language. This nativist approach to SLI received support by research indicating a common genetic underpinning of SLI (Bishop, North, & Donlan, 1996; Saccuman, 2006; SLI-Consortium, 2002; Tallal, Hirsch, Realpe-Bonilla, Miller, & Brzustowicz, 2001; B. J. Tomblin, 1996).

#### *2.5.2.1 Linguistic characteristics as clinical marker*

The ambition to identify a linguistic clinical marker reflects such a domain-specific view on language impairment. However, the search for a linguistic marker is also derived from the intention to provide a clinically relevant tool. As the children's main difficulty is the receptive and/or expressive use of language it is sensible to search for a clinical marker within the domain of language.

Linguistic markers were not necessarily derived from experimental probes. Some researchers designed tasks in order to test specific linguistic features in the children's language. An alternative, widely established approach, however, is the reliance on spontaneous language samples as data source. In this approach, obligatory contexts for the linguistic target forms are usually related to the number of their correct productions. Some researchers combined both methods because language samples demonstrated that children do not necessarily produce enough obligatory contexts for all linguistic forms of interest (Dromi, Leonard, & Blass, 2003; Schöler et al., 1998b). Furthermore, MLU values or certain lexical measures such as lexical diversity are not obtainable through the administration of experimental tasks.

Early attempts to isolate specific characteristics in the language of young, English-speaking children with language impairment were based on the analysis of spontaneous language samples. Two studies presented promising sensitivity and specificity rates for the discrimination between children with SLI and typically developing children. Results indicated that the number of one-word utterances and phrase complexity alongside verb errors seemed to be distinctive for young children with SLI aged between 2;0 and 4;2 (Gavin, Klee, & Membrino, 1993), whereas slightly older children could be characterised through a limited lexical verb diversity in conjunction with verb errors (Fletcher & Peters, 1984). Dunn, Flas, Sliwinski and Aram (1996) analysed language samples of a much larger group of children with SLI (n=201), aged between 2;10 and 6;11. A set of variables with the subvariables age, MLU and percentage of structural errors could identify the clinically diagnosed children with SLI with and accuracy of 96.5%. Only the poor value of specificity (49%) indicated that too many typically developing children were identified as having SLI. Similar concern arose from a subsequent paper of Klee, Gavin and Stokes (2007) presenting LR values for the compositional variable ‘unmarked verbs + verb types’ of Fletcher and Peters, as indication of accuracy independent of the individual sample. The rather broad linguistic variables are likely to be impaired in children with SLI. However, it cannot be inferred that every child showing limitations in these areas, should be diagnosed as language-impaired. Researchers moved consequently on to searching for more defined clinical markers.

Such a perspective was adopted by Rice and colleagues who attempted to narrow down the linguistic structures that are particularly difficult for children with SLI. Instead of employing discriminant analysis, Rice and colleagues searched for linguistic characteristics in which children with SLI performed not only below their age-matched peers but also below younger, typically developing children. Their search focused on a ‘delay-within-delay’ (e.g. Rice, 2003, p. 65) which could serve as clinical marker. Based on comparisons between the three groups of children (mean age SLI: 4;10) Rice and colleagues proposed a composite tense measure as clinical marker consisting of the morphemes representing 3rd person singular *–s*, regular past inflection *–ed*, as well as the copula *be* and the auxiliary *do* (Rice, 2003; Rice & Wexler, 1996; Rice, Wexler, & Cleave, 1995). With the mastery level set at 80% correct production, Rice and Wexler (1996) found 97% of the 5-year-old children with SLI failed this criterion whereas 98% of their age-matched peers had mastered tense marking. It is noteworthy that the authors came to the conclusion of a 'composite tense' as clinical marker irrespective of the source of data. Both spontaneous language samples and experimental tasks resulted in similar findings. This proposition did not only present a milestone in the search for a clinical marker but led further to the development of the extended optional infinitive account (EOI) (Rice et al., 1995; Rice, Wexler, & Hershberger, 1998) which is outlined shortly below in section 2.5.2.2.1.

A number of researchers carried the idea of a clinical marker in the children's morphology further. A review by Klee et al. (2007) of five clinical marker studies indicated that a verb composite proposed by Bedore and Leonard (1998) is a highly promising clinical marker for SLI in English-speaking children aged 3;7-5;9 years. Klee et al. calculated LR values which were very good for this composite (LR+ 33.0; LR- 0.18) derived from the production of 3rd person singular inflection, regular past tense and copula and auxiliary forms of the verb *be*. Thus, this composite is almost identical to the verb composite suggested by Rice and colleagues and has recently been confirmed in its high discrimination accuracy (Gladfelter & Leonard, 2013). The second most promising marker in the review by Klee and colleagues is the combination of two further markers presented by Bedore and Leonard: a noun composite and MLU (LR+ 35.0; LR-: 0.13). Since LRs as well as sensitivity and specificity values are slightly less promising for each of these two variables individually, the verb composite seems nevertheless the best candidate for a clinical marker. This is supported by longitudinal data from children with SLI aged 6-15 years indicating a persistent deficiency in grammatical judgment of *be* and *do* auxiliaries and copula (Rice, Hoffman, & Wexler, 2009). A deficit in verb finiteness seems consequently a good clinical marker for English SLI with potential for both young and adolescent children.

Several subsequent studies on a clinical marker for SLI in English tried to narrow a clinical marker even further down and focused on morphemes individually rather than using composite scores. Conti-Ramsden (2003) tested 5-year-olds on a past tense task and on a noun plural task besides two processing tasks. Only 16% of the children with SLI could successfully be identified on the basis of the noun plural task. The past tense task, on the other hand, revealed satisfactory sensitivity (81%) and specificity (81%) values. However, one third of the children with SLI had been excluded from this task as they were not able to perform the exercise. Both LR values for Conti-Ramsden's past tense task provided by Klee et al. (2007) were unsatisfactory.

Past tense remained yet the most frequently employed linguistic task in clinical marker studies on English SLI. In comparison to other linguistic tasks, such as a 3rd person singular task, past tense marking seemed the most accurate linguistic task in discriminating between language-impaired and age-matched typically developing children (Conti-Ramsden, 2003; Conti-Ramsden, Botting, & Faragher, 2001). Further research indicated that past tense marking has further potential to discriminate between children with SLI and other clinical groups such as children with autistic spectrum disorder (Botting & Conti-Ramsden, 2003), children with attention deficit disorder (Redmond, 2005) and older children with hearing impairment (C. F. Norbury, Bishop, & Briscoe, 2001). If a clinical marker shall comprise only one morpheme, tense marking remains consequently the inflection with the best discrimination accuracy for English-speaking children. However, certain task designs seem to be challenging for language-impaired or very young children and LR values indicate that the choice of a single morpheme as marker may not be as useful as a verb composite instead.

The results for English-speaking children stimulated the search for a clinical marker in the grammar of children speaking languages other than English. Research demonstrates clearly that symptoms vary across languages (for review see Leonard, 1998, 2000) and a linguistic marker is consequently likely to vary, too, the more specific it is.

Three studies reported by Bortolini et al. (2002) compared three different marker candidates from experimental tasks for Italian-speaking children: definite singular articles, 3rd person plural inflection and direct-object clitics. The discrimination accuracy of these grammatical morphemes was compared across three different samples of children and in varying combinations. Object clitics and 3rd person plural inflection revealed the best sensitivity and specificity results. In a follow-up study, Bortolini et al. (2006) replicated the successful discrimination using direct-object clitics. However, 3rd person plural yielded a lower sensitivity value than in the previous study. These results highlight the importance of testing a marker across several samples and the significance of LR values in order to judge a marker's implication. LR values for direct-object clitics were unfortunately not provided by Bortolini et al. (2006) but sensitivity (91%) and specificity (100%) were good and even higher than for the processing task (nonword repetition) included in this study. Thus, besides a composite measure, the production of direct-object clitics can currently be considered as the most promising candidate for a clinical marker for Italian SLI.

The search for a linguistic clinical marker for SLI has further been reported for Spanish- and for Cantonese-speaking children. Spanish-speaking children with SLI did not demonstrated difficulties in verb morphology; neither in present nor in past tense (Bedore & Leonard, 2001, 2005). Nevertheless, they performed below age-matched as well as language-matched typically developing children in the production of direct-object clitics and noun plurals in both elicitation tasks (Bedore & Leonard, 2001) and spontaneous language (Bedore & Leonard, 2005). As in Italian, the production of direct-object clitics (possibly only plural clitics) might therefore be an appropriate clinical marker for SLI in Spanish. Most recently, however, Grinstead and colleagues (2013) questioned the reliability of previous results regarding tense marking in Spanish SLI. The authors claimed that a great number of errors had not been analysed as such but that these would reflect non-finite verb stems instead of a correct verb ending. Changing their measure and analysis method accordingly, Grinstead and colleagues found good discrimination accuracy for tense marking in Spanish. Klee et al. (2004) examined the spontaneous language of 15

Cantonese-speaking children with and without SLI. The 4-year-old children with SLI could successfully be differentiated from both age-matched and language-matched controls. Although the language-matched children did not differ in their MLU and lexical diversity from the children with SLI, the differentiation accuracy for the three groups was 98% if MLU and lexical diversity were related to the children's age. Similar results were obtained if the language samples were controlled for number of words across the groups. The LR values (LR+ 10.33; LR- 0.03) reported in Klee et al. (2007) were good, too. A replication study, though, failed to confirm these promising findings (Wong, Klee, Stokes, Fletcher, & Leonard, 2010). Thus, although MLU and lexical diversity resulted in very good discrimination accuracy within one sample this was not the case for another sample. This underlines again the importance of testing the potential of a clinical marker across several samples.

The trend back to more general language measures such as MLU and lexical diversity is mirrored in the research on English-speaking children, too. One major advantage of this approach is obviously that such a linguistic marker may be more language independent than a marker featuring a language-specific form or composite. Owen and Leonard (2002) investigated the contribution of lexical diversity as such a marker for English SLI. The findings, however, revealed that the children with SLI, aged 3;7 to 7;3, showed a comparable lexical diversity to age-matched, typically developing children if the language samples were matched for number of words. Klee et al. (Klee et al., 2007) replicated their study on Cantonese with 47 English-speaking children, although the English sample (age 2;0 – 4;2) was considerably younger. Classification accuracy of the composite of MLU, lexical diversity and age was at 83% lower than the first results for Cantonese-speaking children. In order to determine whether this marker discriminated less reliable due to crosslinguistic differences or due to the younger age of the English-speaking sample, Klee et al. repeated the analysis with the 2-year-olds excluded. The classification accuracy derived from this sample was higher at 88% while the classification accuracy for the 2-year-olds decreased. The results presented by Klee et al. are therefore certainly promising but three facts raise caution. First of all, the findings summarised above indicate that the lexical diversity of children with SLI can well be at an age-appropriate level. Secondly, Klee et al. did not match the language samples for word number across groups. This implies a certain advantage for the typically developing children who had in their samples larger MLUs, a higher total number of words, and possibly, as a consequence, a greater lexical diversity. Thirdly, the number of participants was in the English study considerably reduced due to the exclusion of 2-year-olds. A replication of the study with a larger sample is therefore necessary before a marker reflecting MLU, lexical diversity and age can be considered as reliable.

In summary, the review of the literature on linguistic markers illustrates that general language measures, composite measures as well as specific linguistic features have been proposed as candidates. The more specific the proposed marker the more restricted is its applicability crosslinguistically. Past tense marking may be one of the most promising linguistic markers for English but Italian- or Spanish-speaking children can better be identified on the basis of their use of direct object clitics. Composite scores involving the production of several morphemes result usually in better LR values and are therefore more likely to discriminate accurately between SLI and typical development independent of the population sample. However, these composite scores are usually more difficult to compute or involve a time consuming data collection such as the transcription of spontaneous language samples. Conflicting findings were reported regarding the discrimination accuracy of even more general measures such as MLU and lexical diversity. However, these are the easiest to compare across different languages. Many studies show methodological limitations as they often involve relatively small samples sometimes in combination with a large age range. Several researchers attempted to overcome these limitations through the application of the findings to another population sample. The often varying outcome from these different samples highlights the importance of replication studies before any clinical marker can be considered as reliable.

### *2.5.2.2 Linguistic theories on SLI*

The linguistic characteristics of children with SLI identified stimulated the development of theories on SLI, as indicated in the decision hierarchy of SLI in Figure 1. Three influential linguistic accounts of SLI are briefly introduced below as examples of theories derived from a domain-specific view on SLI. Thus, all theories are based on the assumption of an innate domain-specific language device (Noam Chomsky, 1965; van der Lely, 1997).

2.5.2.2.1. *Extended Optional Infinitive stage*

As mentioned above, the discovery of a marked deficit in tense marking in English SLI led Rice and colleagues to the development of the extended optional infinitive hypothesis EOI (Rice et al., 1995; Rice et al., 1998). This hypothesis claims that children with SLI take longer than typically, language-matched children to acquire the features tense and agreement, i.e that finite forms are obligatory in main clauses. The children remain therefore for a protracted period in a stage of optional tense and agreement marking while typically developing children are considered to have proceeded beyond this stage by the age of 4-5 years.

Confirming the EOI's predictions, studies by Rice and colleagues (Rice & Wexler, 1996; Rice et al., 1998) demonstrated that it is morphemes signalling tense and agreement which are affected in English-speaking children, such as past tense *–ed* and present tense 3$^{rd}$ person singular *–s*, yet not non-tense related morphemes, such as plural marking *–s,* progressive *–ing* or prepositions *'in'* and *'on'* . This is especially remarkable because 3$^{rd}$ person singular inflection and plural marking are homophonous morphemes which are affixed in a similar phonological position and show a similar surface structure: word final –*s*. On tense related morphemes, though, the children did not only lag behind their chronological age peers but also behind their controls matched for language according to MLU. Paradis and Crago (2001) described similar patterns for tense-related morphemes compared with non-tense related morphemes in French-speaking children with SLI. The children with SLI in their study produced the 3$^{rd}$ person singular form *'a'* of the auxiliary *'avoir'* with a significantly lower accuracy rate than the homophonous preposition *'à'*, i.e. a non-tense morpheme. Difficulties in tense marking have been reported for further languages such as German (Rice, Noll, & Grimm, 1997), Dutch (De Jong, 1999) and Swedish (Hansson, 1997), too. Children acquiring a null-subject language, on the other hand, were predicted not to show such deficits and demonstrate indeed much less difficulties with tense marking than English-speaking children with SLI (Bedore & Leonard, 2005; Bortolini et al., 2006).

Although there is a vast body of evidence confirming the predictions of the EOI, other data is less easy to accommodate with this account. Certain error types, for example, are difficult to explain on the basis of the EOI. While the EOI predicts the use of verb infinitives it cannot account for bare stems or incorrect inflections as observed in Dutch (De Jong, 1999) and German (Clahsen et al., 1997; Roberts & Leonard, 1997). Furthermore, while several studies confirmed the prediction that children with SLI acquiring a null-subject language do not show optional infinitives the EOI cannot account sufficiently for the difficulties these children do show. Another challenge to the EOI are studies examining the influence of verb frequency and phonological aspects. According to the EOI, it would not be expected that these factors influence the children's production of tense and agreement markers. However, findings in English SLI show the contrary (Marchman, Wulfeck, & Weismer, 1999; Norbury, Bishop, & Briscoe, 2001; Oetting & Horohov, 1997). In their production of past tense, the children show sensitivity to the phonological complexity of the verb stem as well as to the verb frequency. The EOI can consequently be considered as a theory on SLI that can successfully explain error patterns of several grammatical morphemes across different languages alongside a high degree of accuracy of other morphemes. However, besides the fact that the EOI can account 'only' for difficulties regarding tense and agreement marking further limitations are apparent in a number of exceptions from the model's predictions found in empirical data.

### *2.5.2.2.2 Missing Agreement Hypothesis*

One of the earliest accounts viewing the grammatical difficulties in SLI as a deficit in the linguistic system is the Missing Agreement hypothesis by Clahsen (1989) which originated from findings in German-speaking children with SLI. The grammatical feature agreement requires that one functional category controls another category. For example, finite verb forms must agree with person and number of the subject. However, according to the Missing Agreement Account, children with SLI are restricted by a selective inability to match grammatical features of different functional categories within one sentence while the grammatical system is otherwise preserved. Affected likewise are those inflections reflecting agreement in gender, number and case between determiners and nouns as well as between adjectives and nouns. If correct forms can be found in the children's language these are rote-learned. Examples of inflections that do not involve agreement are plural marking in nouns and past tense marking which are consequently thought to be less challenging for children with SLI. Indeed, noun plural marking does not usually pose particular problems to English-speaking children with SLI (e.g. Rice & Oetting, 1993) whereas the use of auxiliaries, subject-verb agreement in main verbs (e.g. Rice & Wexler, 1996; Rice et al., 1998) and case marking in noun phrases does (e.g. Leonard, 1995; Loeb & Leonard, 1991; Wexler et al., 1998).

However, although tense marking does not involve structural relations between different functional categories and should therefore be unimpaired, past tense marking in English-speaking children with SLI has been reported to fall behind that of MLU controls (e.g. Rice, 2003; Rice & Wexler, 1996). Findings of Italian-speaking children with SLI are even more difficult to explain because they show only very few problems with verb-subject agreement, except of in 3$^{rd}$ person plural marking (e.g. Bortolini et al., 1997; Bortolini, Leonard, & Caselli, 1998). The Missing Agreement Account, though, cannot provide any reason why Italian-children should have access to agreement features in verbs whereas English- and German speaking children do not. In the meantime, even Clahsen and his colleagues (Eisenbeiss, Bartke, & Clahsen, 2005) have called the Missing Agreement Account into question because Eissenbeiss et al. (2005)

failed to find evidence for deficits in case-marking in German SLI. Hence, the Missing Agreement Account provides a good basis for some symptoms in SLI but it cannot be applied to the full range of grammatical symptoms.

##### *2.5.2.2.3 Deficit in Computational Grammatical Complexity Hypothesis*

Another theory that assumes a deficit in the syntactic system of children with SLI is the deficit in Computational Grammatical Complexity (CGC) hypothesis (Marshall, 2006; van der Lely, 2005) which was developed from the Representational Deficit for Dependent Relationships hypothesis (RDDR) by van der Lely (1994, 1996). However, it needs to be pointed out that the authors refer only to one specific subgroup of the population with SLI which van der Lely identified as 'Grammatical-SLI children' (G-SLI) (e.g. van der Lely, Rosen, & McClelland, 1998; van der Lely & Stollwerck, 1996). This group is characterised by a persistent deficit at the age of nine years and beyond in syntax, morphology and often phonology (although they are intelligible for known words), whereas the lexicon as well as derivational morphology are relatively spared. All other aspects, such as pragmatics, articulation but also auditory processing are reported to be broadly unaffected.

The CGC hypothesis claims that children with G-SLI have an underlying deficit in representing linguistic, structurally-complex forms in the three components of the computational grammatical system: syntax, morphology and phonology. For example in syntax, children with G-SLI are described having particular difficulties in the comprehension as well as production of elements marking syntactic dependency (van der Lely, 2003). A dependent structural relationship is necessary if sentence constituents, such as grammatical features and lexical categories, need to be linked. This, for example, is the case in subject-verb agreement, case marking, tense marking and the matching of semantic roles onto syntactic functions in complex sentences such as passive constructions and questions. According to the CGC, especially long-distance relationships, being characteristic of more complex syntax, are difficult for children with G-SLI and result often in ambiguous interpretations. Particular difficulties with the production of Wh-questions were reported by van der Lely and Battell (2003);

and indeed object questions (longer distance relationship) were more difficult than subject questions (shorter distance relationship). This difference between subject and object questions is described as being greater for the children with G-SLI than for the language-matched control children.

As the CGC hypothesis further regards phonological complexity as a form of grammatical complexity, it stands out considerably from other linguistic accounts. In this way, the CGC accounts for phonotactic effects in children with G-SLI but not in typically developing children (Marshall & van der Lely, 2006, 2007). Marshall and van der Lely (2006, 2007) found structural constraints had a higher impact on the use of grammatical morphemes in children with SLI than in younger, typically developing children; similar results were found in an investigation of English and of Italian SLI (Bortolini & Leonard, 2000). However, in this study also noun plural marking was affected by structural constraints although this is not a long-distance relation and should therefore be spared according to the CGC's predictions on syntax. Generalisations from the subgroup G-SLI to the whole population of children with SLI are not possible, though. The CGC hypothesis is consequently only relevant for a highly selective subgroup of SLI and even van der Lely (2005) states that this is a very rare form of SLI.

##### *2.5.2.2.4 Conclusion on linguistic accounts of SLI*

This overview of the most established and most frequently discussed linguistic accounts of the clinical symptoms found in children with SLI does not claim to be exhaustive in scope. However, it outlines how different the predictions can be depending on the account of SLI, and even depending on the linguistic framework that is underlying the account. One major advantage of these very detailed and precise predictions has been in the resulting, similarly precise research questions and study designs, which contributed enormously to our knowledge about SLI. Nevertheless, this overview also demonstrates that we still have not arrived yet at the end of the road. Most linguistic theories of SLI struggle to explain all symptoms even within one language, but especially to confirm predictions cross-linguistically.

### *2.5.3 SLI within a domain-general framework*

A domain-specific approach to SLI has been challenged by researchers who propose that the language deficit is a consequence of limitations in non-linguistic, cognitive processes. The shift from a domain-specific towards a domain-general approach reflects changes in models of language acquisition. While early nativists assume that language acquisition is based on an innate language acquisition device (Noam Chomsky, 1965) that results in domain-specific, modular language skills, the influence of environmental factors as well as domain-general skills were more and more recognised within other linguistic frameworks. The role of language input has even been incorporated within a nativist framework as triggering language development (Noam Chomsky, 1995; Guasti, 2002). Cognitive linguistic frameworks, on the other hand, consider both as highly relevant the nature of language input and cognitive processes. According to cognitive approaches, language acquisition is not based on an innate language faculty but that language acquisition is a cognitive process emerging from the interaction between an innate set of domain-general abilities and the child’s environment (Bates et al., 1999; Tomasello, 2003). Language development is thus assumed to be a learning process. From the point of view of an emergent and a neuroconstructivist framework, the result of this process may be language-specific, such as lexical or grammatical knowledge, but the process of language acquisition is characterised by the gradual emergence of particular cognitive processes facilitating language learning. Depending on the level of language development and the language input, these processes may change over time (see Thomas, 2005 for overview). SLI is therefore considered as an interruption of language learning rather than a deficit in language knowledge. The interaction between environmental factors, cognitive processing and language development has been supported by findings from an intervention study by Pikho and colleagues (2007). Pikho et al. demonstrated that a systematic intervention program could improve language skills in children with SLI as well as cerebral functions which are associated with language processing.

#### *2.5.3.1 Processing tasks as clinical marker*

Findings from the last two decades demonstrated that children with SLI do not only fall behind their age-matched and / or their language-matched, typically developing peers on language tasks but also on tasks testing their working memory. Aspects of these investigations were simultaneous processing and storage on the one hand and phonological working memory on the other hand. The findings from research focussing on these latter aspects led to proposals of processing markers for SLI. Nonword repetition tasks as well as tasks involving sentence repetition and digit recall are considered as measurement tools of phonological working memory and were tested as potential clinical markers.

**Nonword repetition**

One of the studies that set the ball rolling into research in nonword repetition in SLI was the work by Gathercole and Baddeley (1990). The authors compared six children with SLI on a task of nonword repetition with two groups of typically developing children, age-matched children and younger children. The children with SLI scored significantly below both control groups on three- and four-syllable words and they showed a performance that was comparable to that of children aged on average four years younger than the children with SLI. Bishop, North and Donlan (1996) used the same task in a study involving children with SLI (all twins), their unimpaired twin siblings as well as resolved cases of SLI. Results showed that not only the children with SLI performed significantly below typically developing children on Gathercole and Baddeley’s nonword repetition tasks but also the children with resolved SLI. Moreover, the performance on nonword repetition was highly inheritable. Monozygotic twins of SLI children were more likely to perform poorly on nonword repetition than dizygotic twin siblings.

Following these highly promising studies, Dollaghan and Campbell (1998) were the first researchers who applied these findings to the search for a clinical marker. The authors designed their own nonword repetition task in order to avoid high impacts of lexical knowledge or speech disorders. In accordance with the previous findings, Dollaghan and Campbell’s results were promising.

Using different cut-off points for passing and failing the task, an overall accuracy of 98% was obtained in the discrimination between children enrolled in speech and language intervention and typically developing children. Data of a pilot study with 20 children in the SLI group showed furthermore that there was only little within-group variation and literally no overlap in accuracy levels between typically developing children and language impaired children. These characteristics make nonword repetition a good candidate for a clinical marker.

A large number of studies followed this research in order to investigate the nature of nonword repetition or to validate its accuracy as clinical marker in the identification process of children with SLI. All studies on English-speaking children found significant differences between children with SLI and typically developing age-matched or language-matched children (e.g. Botting & Conti-Ramsden, 2003; Briscoe, Bishop, & Norbury, 2001; Conti-Ramsden, 2003; Conti-Ramsden et al., 2001; Conti-Ramsden & Hesketh, 2003; Ellis Weismer et al., 2000; Gray, 2003a; Marton & Schwartz, 2003). Furthermore, reliable differences were also reported independent of the cultural or dialectal background of English-speaking children (Ellis Weismer et al., 2000; Oetting & Cleveland, 2006; Rodekohr & Haynes, 2001) as well as across languages such as Italian (Bortolini et al., 2006) and Spanish (Calderon, 2004; Girbau & Schwartz, 2007). Only Cantonese-speaking children with SLI (mean age 4;11) performed at a similar level as their age-matched peers (Stokes, Wong, Fletcher, & Leonard, 2006).

Results of longitudinal studies or studies involving resolved cases of SLI supported the findings by Bishop et al. (1996) showing that the children's limitations in nonword repetition remain relatively stable (Conti-Ramsden & Durkin, 2007; Thal, Miller, Carlson, & Vega, 2005). However, not all of the studies on English-speaking children replicated the high accuracy and likelihood levels of Dollaghan and Campbell (1998). Ellis Weismer and colleagues (2000), for example, applied Dollaghan and Campbell's methodology to a very large sample of school-age children but the LR+ of 6.5 was unsatisfactory. This value declined further to 2.6 if the gold standard was not the fact whether a child was enrolled in speech and language intervention (as employed by Dollaghan and Campbell) but the actual absence or presence of language impairment as diagnosed by a clinician. Mixed results were presented by Conti-Ramsden and colleagues with specificity values above 80% but lower sensitivity values (Conti-Ramsden, 2003; Conti-Ramsden et al., 2001; Conti-Ramsden & Hesketh, 2003). Recently, Archibald and Joanisse (2009) tested a large sample of school-age children on a very short screening employing nonword repetition. The convincing results summarised above could not be replicated with a sensitivity level of 46% and specificity of 66% in their study.

Graf Estes, Evans and Else-Quest (2007) provide a thorough review of the studies on nonword repetition in English-speaking children. Although some researchers suggested an influence of the children's age on the identification accuracy of nonword repetition (e.g. Bortolini et al., 2006), Graf Estes et al. found the results from the studies in their meta-analysis being independent from age. Sensitivity and specificity (or effect size as in the meta-analysis computed) vary instead according to the measurement tool used. The various nonword repetition measures employed in the studies differ in their number of items, wordlikeness, articulatory complexity, nonword length and scoring method. Most studies used one of the following two procedures: the above described Nonword Repetition Test (NRT) by Dollaghan and Campbell (1998) or the Children's Test of Nonword Repetition (CNRep) by Gathercole and Baddeley (1996). The CNRep comprises more than twice as many items than the NRT, contains English words and affixes within its nonwords, includes articulatory simple as well as complex items and is scored on a word by word approach rather than phoneme by phoneme. However, even in this test the child receives credit for consistent phonological errors in the sense that the child does not score lower due to a consistent speech disorder. Graf Estes et al. found that sensitivity and specificity levels were usually higher in studies that employed the CNRep in comparison to the NRT. This conclusion has been confirmed by Archibald and Gathercole (2006a) who compared these two test within the same group of children. The lack of a between-group difference in Cantonese-speaking children (Stokes et al., 2006) further highlights the influence of the test design. Cantonese is a phonologically relative simple language and the nonwords designed for this language may consequently challenge the children's phonological working memory less than nonwords designed for phonologically more complex languages. Another suggestion of the authors is that nonword repetition tasks designed for other languages may draw more on knowledge of existing lexical entries than the task designed for Cantonese. This would underline the influence of wordlikeness on test results which is evident from studies showing an effect of phonological probability on nonword repetition accuracy (Gathercole, 2006; McKean, Letts, & Howard, 2013a; Munson, Edwards, & Beckman, 2005). However, more research is necessary within and across languages in order to identify those test characteristics that result in the most reliable discrimination between specifically language impaired and unimpaired children.

*Specificity of nonword repetition*

Dolloghan and Campbell (1998) designed their NRT carefully, in order to avoid that children being identified due to a speech disorder rather than their language problems. Nevertheless, the golden standard used by the authors sheds some doubts in how far language rather than phonological problems may have caused the problems on the NRT and resulted in the excellent identification accuracy. Instead of using a standardised language test, children were selected on grounds of their enrolment in language therapy. In this respect, it is important to note that children with speech output impairment are more likely te be referred for speech and language intervention than children with language disorders only (Zhang & Tomblin, 2000). Furthermore, from a selection of different speech and language measures, output phonology was found to be the best predictor for accuracy in nonword repetition in Swedish-speaking children (Sahlén, Reuterskiold-Wagner, Nettelbladt, & Radeborg, 1999). Consequently, as particularly evident in Ellis Weismer et al's study (2000), the gold standard employed in the different studies is another factor that highly influences the outcome of the research. Differential diagnoses need therefore to be considered carefully when setting the inclusion criteria.

Further doubts on the specificity of nonword repetition as clinical marker arise from studies involving not only language impaired and language unimpaired children but also children with other developmental disorders. Nonword repetition has been reported to be limited in children with reading difficulties (e.g. Conti-Ramsden & Durkin, 2007; Roodenrys & Stokes, 2001), in hearing impaired children (Hansson, Forsberg, Löfqvist, Mäki-Torkko, & Sahlén, 2004), as well as in children with Down Syndrome (e.g. Jarrold, Baddeley, & Hewes, 2000; Keller-Bell, 2001). Furthermore, nonword repetition failed to discriminate successfully between children with SLI and children with autism or pragmatic disorder (Botting & Conti-Ramsden, 2003) and also between SLI and language impairment accompanied by limitations in nonverbal IQ (Ellis Weismer et al., 2000). It is therefore doubtful in how far limitations in nonword repetition are really characteristic for children with SLI with regards to other developmental problems.

Taken together, nonword repetition is clearly deficient in most children with SLI compared to their typically developing peers across different cultural as well as linguistic backgrounds. However, the appropriateness cannot be overgeneralised to other languages without investigation. Results that vary in their identification accuracy across studies indicate further that nonword repetition may be useful in the identification of children with SLI but it is not sufficient on its own as clinical marker. The influence of test characteristics needs to be investigated in greater detail and it is important to choose carefully the gold standard against which the clinical marker is tested. Moreover, the evidence up-to date suggests that nonword repetition as clinical marker may over-identify children with developmental limitations other than SLI. Nonword repetition should consequently not only be used in conjunction with another marker but also in conjunction with the common exclusion criteria.

**Sentence repetition and digit recall**

In addition to nonword repetition, two other measures of verbal working memory have been tested as potential clinical marker for SLI: sentence repetition and digit recall. These two alternative processing markers received relatively little attention, though, in comparison to nonword repetition.

In the case of digit recall, this little attention results from findings that show promising specificity values of above 90% but sensitivity values, on the other hand, of just above 50% (Conti-Ramsden, 2003; Conti-Ramsden & Hesketh, 2003). Although children with SLI have been demonstrated to perform below typically developing age-matched peers, digit recall is consequently no adequate clinical marker for SLI and has not been followed up further in research.

In the case of sentence repetition, the limited attention may result from the difficulties to draw conclusion about a causal relationship between a poor performance on sentence repetition and SLI. Sentence repetition taps a number of additional processes in comparison to nonword repetition and it is therefore more difficult to identify the process that is the most relevant in causing both limitations in sentence repetition and in language development in general. Lombardino and Potter (Lombardi & Potter, 1992; Potter & Lombardi, 1998) carried out a number of experiments demonstrating that sentence repetition assesses not only the verbatim short-term storage of the sentence but draws also on lexical entries of the long-term memory. Priming effects for the syntactic structure indicate that processes are involved that are alsorelevant for normal sentence production. As sentence repetition relies therefore on both language and memory abilities and involves more linguistic processes than nonword repetition it is less surprising to find children with SLI performing poorly on this measure. Nevertheless, some researchers selected sentence repetition in order to test its potential as clinical marker. The task has been referred to as sentence repetition, sentence imitation or recalling sentences but it always refers to the immediate reproduction of auditory presented sentences. One early study that investigated sentence repetition in individuals with SLI compared to sentence repetition in individuals without SLI was a telephone task carried out by

Tomblin, Freese and Records (1992). The authors demonstrated that adults with a positive history of SLI could best be differentiated from adults without history of SLI on the basis of their performance in sentence repetition. Thus, the results seemed promising that sentence repetition may be a clinical marker for SLI that is stable over time and independent of whether the individuals still experience the language problems or not. Several studies confirmed that children with SLI have marked difficulties in sentence repetition relative to typically developing children (Archibald & Joanisse, 2009; Briscoe et al., 2001; Conti-Ramsden et al., 2001; C. F. Norbury et al., 2001; Redmond, 2005). Conti-Ramsden et al. (2001) could further replicate the findings that this marker gives very satisfying accuracy levels independent of the current language status. From a selection of four different tasks tested as potential marker (nonword repetition, sentence repetition, 3rd person singular task and past tense marking), sentence repetition showed the highest accuracy level with 88% (sensitivity 90% / specificity 85%) and was furthermore highly successful in the identification of resolved cases of SLI.

*Specificity of sentence repetition*

The discrimination of SLI from other developmental disorders by means of sentence repetition is slightly more difficult than from typically developing children. Redmond (2005), for example, reported that children with attention deficit disorder, too, struggle to repeat orally presented sentences as correctly as typically developing children. Moreover, Botting and Conti-Ramsden (2001) showed that sentence repetition was not only the most effective marker for SLI in comparison to nonword repetition and a past tense task, but this was also the best marker in identifying autistic children or children with pragmatic impairment from typically developing children. Nevertheless, even in the discrimination between children with SLI and autistic children or children with pragmatic disorder, sentence repetition proved to be a satisfying tool with accuracy levels of 90% respectively. Moreover, Norbury et al. (2001) demonstrated that children with SLI were distinctly limited in sentence repetition whereas hearing impaired children could use their semantic skills better to fill in their gaps and performed within normal levels.

*Sentence repetition across languages*

Sentence repetition as identifier of SLI has so far been tested predominantly in English-speaking children. One exception is the work by Stokes et al. (2006) on Cantonese-speaking children. In contrast to nonword repetition, sentence repetition could successfully be used to differentiate between typically developing children and those with SLI. The obtained value for specificity was with 97% very satisfactory, only the value sensitivity was marginal at 77%. However, also language impaired children with a language background other than English or Cantonese have also been reported to lag behind their typically developing peers on sentence repetition. Schöler, Fromm and Kany (1998b), for example, demonstrated that German-speaking children with SLI make significantly more errors than age-matched control children while repeating sentences. The authors did not examine the identification accuracy of the task, though. One fact that needs to raise caution is that sentence repetition forms a subtest of several standardised language tests. Evidence showing high accuracy levels is little surprising and less persuasive if a similar subtest has been part of the gold standard applied in the study such as in the results presented by Archibald and Joanisse (2009). Such an overlap of measures can distort results.

To sum up, from all processing tasks investigated as potential clinical marker for SLI sentence repetition appears to be the most promising. Good accuracy levels have been achieved besides evidence indicating stable performances over time and independent from current language status. Some findings suggest that sentence repetition can not only be applied to the discrimination between typically developing children and children with SLI but also for differentiation from other disorders. Further research is necessary in order to validate these findings although it should carefully be avoided to use sentence repetition tasks in the participant selection process, too.

### *2.5.3.2 Domain-general approaches to SLI*

The exclusion criteria for SLI are aimed to ensure a pure and specific impairment in the area of language only. Yet, as the suggestion of processing markers for SLI shows, many children with SLI have subtle limitations in non-linguistic skills additionally to their language difficulties (see Johnston, 1999; Leonard, 1998, for review). These difficulties in non-linguistic skills are often not considered as exclusionary criteria for SLI or, for example in the case of a lower non-verbal IQ, are not so severe that they would influence the inclusion to the SLI population. Besides difficulties in the children's working memory, a number of further limitations were frequently reported, such as a slowed reaction time on a number of different tasks regarding visual-spacial processing (Johnston & Weismer, 1983; Schul, Stiles, Wulfeck, & Townsend, 2004) and auditory processing (Tallal & Piercy, 1973a, 1973b, 1974, 1975). Furthermore, linguistic skills were shown to be influenced by processing demands evident for example from studies manipulating argument structure complexity (Grela, 2000; Grela & Leonard, 2000; Pizzioli & Schelstraete, 2008) or showing priming effects for particular syntactic frames (Leonard et al., 2002; Leonard, Miller et al., 2000).

Taking these findings and changes in theoretical frameworks on language development together, accounts on SLI emerged that attempt to relate the children's non-linguistic deficits with their language impairment. The core question these domain-general accounts on SLI have in common is which of these difficulties are concomitant but unrelated to the linguistic difficulties and which of these non-linguistic difficulties play a major role in the phenomenon SLI. Some of the most important theories are shortly summarised below. This review illustrates that early domain-general theories proposed very specific processing limitations at the root of SLI whereas more recent theories assume a multi-factorial approach to SLI and acknowledge that it is probably the combination of different factors that lead to the language difficulties observed.

#### *2.5.3.2.1 Deficits in Temporal Processing*

The rapid auditory processing hypothesis was derived from a series of studies by Tallal and Piercy (Tallal & Piercy, 1973a, 1973b, 1974, 1975) and resulted in the proposition that SLI is caused by impaired auditory processing of brief input and of rapidly successive transitions. This theory has been very influential. Although most research has been carried out with English-speaking participants, the rapid auditory processing hypothesis has prompted a substantial amount of research projects focussing on various aspects of temporal processing in children with SLI and their controls. The evidence to date, though, has been inconsistent and controversial (D. V. Bishop et al., 1999; Bishop, Carlyon, Deeks, & Bishop, 1999; Corriveau, Pasquini, & Goswami, 2007). Although it seemed a decade ago that a deficit in processing of brief or rapidly changing stimuli is ‘an important piece of the SLI puzzle’ (Leonard, 1998, p. 145) the ‘causal’ link could not be consistently validated. The findings of more than three decades of research are characterised by a high variability across studies. This variability reflects the wide range of factors that influence the outcome. For example, several task factors have been identified that influence the children’s performance: the nature of the auditory stimuli (e.g. synthesised or natural speech contrasts) (e.g. Coady et al., 2005), frequency range and magnitude of contrast (e.g. Rinker et al. 2007) the phonetic context the stimuli occurs in (Leonard et al., 1992) or task complexity (e.g. Coady et al., 2007; Hanson & Montgomery, 2002). Furthermore, differences across studies in the composition of participant groups according to age or concomitant deficits are probable to have contributed to the conflicting results. High variability within the studies demonstrates that apparently only a subset of children with SLI experience the hypothesised difficulties (e.g. Ahmmed, Parker, Adams, & Newton, 2006; McArthur, Ellis, Atkinson, & Coltheart, 2008; Uwer, Albrecht, & von Suchodoletz, 2002). Yet, it remains unclear whether the problems in temporal auditory processing tasks result from a misperception of the presented contrast due to temporal constraints (as according to the rapid auditory processing hypothesis), from problems in the rapid formation and / or storage of specified mental representations of the input including the temporal order, and /

or from problems in the accurate retrieval of the stored representations necessary for the response (Gathercole & Baddeley, 1990).

#### *2.5.3.2.2 Phonological Storage Hypothesis*

The finding that children with SLI experience marked difficulties in nonword repetition resulted in the proposition of the Phonological Storage Hypothesis. Nonword repetition has commonly been seen as a measure of phonological short-term memory capacity. Gathercole and Baddeley (1989, 1990) were the first who hypothesised a causal relationship between SLI and a specific deficit in one particular component of working memory, the phonological loop. According to the model of working memory by Baddeley and Hitch (Baddeley, 2000; Baddeley & Hitch, 1974), the phonological loop consists of two subcomponents, each responsible for a different functions: Firstly, a temporary storage buffer which can hold acoustic and verbal information for a few seconds, and secondly, a subvocal or articulatory rehearsal system that refreshes and maintains the information. The phonological loop is assumed to play an important role in language acquisition, language comprehension and expression. For example, in sentence comprehension as well as in learning new vocabulary, the phonological buffer is thought to hold temporarily the incoming words while they are processed further in the language system.

A deficit in phonological storage is consequently thought to impact on the children's word learning skills because the acquisition of a stable phonological representation for a novel word would be more difficult. Poorer word learning abilities have indeed been reported in different contexts, too, involving word recognition or expressive tasks under different learning conditions (Alt, Plante, & Creusere, 2004; Ellis Weismer, 1996; Gray, 2003b, 2006; Hansson et al., 2004; Oetting, Rice, & Swank, 1995). Results showed for example that children with SLI need a higher number of exposures to acquire novel words (Gray, 2004, 2005; Rice, Oetting, Marquis, & Bode, 1994). Deficits in the use of grammatical morphemes, on the other hand, are explained by an impact of a deficient phonological storage on the comprehension of grammar. In particular, long and syntactically complex sentences which cannot be easily interpreted online are most likely to be affected according to Gathercole and Baddeley

(1993). However, also the limitations in lexical learning are thought to have a secondary affect on the children's grammar as vocabulary learning was shown to predict the acquisition of new bound morphemes (Nakamura, Plante, & Swisher, 1990). Stimulated by the clinical marker research on SLI, suggesting nonword repetition as a suitable marker for SLI, the investigation of a link between working memory and language impairment has received much attention in research over the last three decades[1]. The fact alone that children with SLI have problems in the temporary storage and the reproduction of phonological input as well as problems in different language areas does not yet imply a causal relationship between both.

Up to date, the nature of the relationship between difficulties in nonword repetition and the language deficit of the children could not be unambiguously identified. While Gathercole and Baddeley suggested that it is a deficit in phonological short-term memory, other researchers suspected the root of the children's problems in nonword repetition lay with abilities such as speech perception, phonological encoding, phonological abstraction, articulation and lexical access (see for review Coady & Evans, 2008; J. Edwards & Lahey, 1998; Munson, Kurtz, & Windsor, 2005; Snowling, Chiat, & Hulme, 1991). A number of studies demonstrated that the significant group difference between children with SLI and their age-matched peers is no longer evident if the children's vocabulary size is covaried out in statistical analysis (McKean et al., 2013a; Munson, Kurtz et al., 2005; van der Lely & Howard, 1993). This may imply, as van der Lely and Howard argued, that the measured memory deficit resulted from the linguistic nature of the stimuli in nonword repetition. Thus, the children may present the difficulties in nonword repetition because their lexical knowledge is deficient. Indeed, the linguistic composition of the nonwords has a further influence on the children's performance, as the children's accuracy in nonword repetition decreases with increasing word length and low phonological

[1] The terminology used by different authors differs slightly and has changed over time. *Phonological short-term memory* and *phonological working memory* can be considered as synonyms. V*erbal working memory* refers to both components of the verbal working memory as proposed by Baddeley and Hitch (2000), the phonological storage buffer and the articulary rehearsal system.

probability[2] (Gathercole, 2006; McKean et al., 2013a; Munson, Kurtz et al., 2005). McKean, Letts and Howard (2013a), though, demonstrated different developmental trajectories regarding nonword repetition skills between typically developing children and children with SLI, especially evident with increasing age. The authors suggest that the children with SLI may develop a different strategy of lexical processing which is efficient in early language development but leads eventually to a plateau in language processing skills.

Overall, to date, there seems common agreement that an isolated deficit in phonological storage in unlikely to cause SLI. Even Gathercole (2006) presented a revised version of the original hypothesis acknowledging a wider range of processes involved in nonword repetition. However, correlation studies focussing on the relation between nonword repetition and specific language abilities did not always result in consistent findings. This may indicate that those cognitive deficits causing the ‘double hit’ do not necessarily have to be the same for the whole population of SLI or that they may change over time.

#### *2.5.3.2.3 Surface Account*

Another, very influential theory which is based on the assumption that limitations in the children’s general processing capacity impede their language development is ‘the Surface Account’ proposed by Leonard and his colleagues (Leonard, 1989, 1998; Leonard et al., 1997; Leonard et al., 1992). This theoretical approach to SLI has been referred to as the ‘surface’ account or hypothesis because of its emphasis on the surface characteristics of the particularly vulnerable linguistic forms. Especially, the deficits in temporal processing suggested by Tallal and Piercy (Tallal & Piercy, 1973a, 1973b, 1974, 1975) formed an important basis for the development of Leonard’s hypothesis. Leonard proposed that children with SLI have particular problems in perceiving grammatical morphemes and inferring their grammatical function due to limitations in general processing capacity. The linguistic domain of grammar

[2] Phonological probability is a measure referring to the frequency of phonemes and phoneme sequences in a language. A high phonological probability implies that the word consists of phonemes and phoneme sequences that are frequent in the respective language whereas a low phonological probability implies infrequent phonemes and phoneme sequences.

itself is considered to be intact as well as the children are thought to be generally able to perceive the phonetic form of the morphemes. However, because of the processing limitations, the children fail to discover the morpheme's function if the morpheme's challenging surface characteristics place additional burden on processing.

A large number of research studies were carried out across languages to verify the hypothesis' predictions. The surface account can best account for the linguistic profiles of children learning English or a Roman language such as Italian, Spanish or French. In these languages only minor irregularities occur which are difficult to bring in line with the surface account, such as the magnitude of difficulties or specific error patterns. The morphemes that are affected, though, meet the characterisation of Leonard and colleagues (Leonard, 1989, 1998; Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992; Leonard, Eyer, Bedore, & Grela, 1997; Leonard, McGregor, & Allen, 1992). This characterisation applies further to many morphemes that were reported to be susceptible in Germanic languages. However, not always are all those morphemes indeed affected which should be impaired according to their surface pattern or morphemes of low salience are even substituted for morphemes of higher salience as in Dutch or German (Clahsen et al., 1997; Orgassa & Weerman, 2008). The data of Hebrew-speaking children indicated furthermore that additional morphemes may be affected without being challenging in their surface appearance (Leonard, Dromi, Adam, & Zadunaisky-Ehrlich, 2000). Findings suggested that the processing load may in fact play a greater role than the phonetic salience. At first glance, this does not seem contradictory to the surface account. It is, according to Leonard (1998), the combination of being more challenging to perceive *and* bearing a complex grammatical and / or semantic function that hamper the acquisition of particular morphemes. Nevertheless, if the surface characteristics do not constitute an obstacle, sufficient resources should be available to perform the necessary processing operation. This, however, does not seem to be the case for Hebrew-speaking children. Also Cantonese-speaking children with SLI were reported to have difficulties with aspect markers although their surface characteristics do not place a particular burden on the children's processing resources (Fletcher, Leonard, Stokes, & Wong, 2009). It is consequently noteworthy that not all morphemes problematic in SLI fulfil the criteria of low salience. Another symptom frequently described for children with SLI learning German, Dutch or Swedish are word order problems (Clahsen, 1989; De Jong, 2004; Håkansson, 2001; Hansson, Nettelbladt, & Leonard, 2000). These cannot be accounted for by the surface hypothesis. The surface account is accordingly not sufficient in order to explain all observed symptoms in children with SLI of varying language background. Results from cross-linguistic comparisons contradicted further with the surface account as morphemes of equivalent salience and grammatical complexity are affected to differing degrees across languages. These findings resulted in the development of the morphological richness account, which is generally compatible with the surface account.

#### *2.5.3.2.4 Morphological Richness Hypothesis*

The morphological richness hypothesis (Dromi et al., 1999; Leonard, 1998, 2000; Leonard, Sabbadini, Leonard, & Volterra, 1987) was derived from the observation that morphemes corresponding to a specific grammatical function are used in one language to a higher degree than in another language, although these morphemes share a low salience and are in both languages less accurately used by children with SLI than by their language-matched controls. For example, a review by Leonard and Deevy (2006) illustrates crosslinguistic differences: English-, German- and Dutch-speaking children with SLI of same age and with comparable severity of the disorder use 3$^{rd}$ person singular present tense inflections to different degrees, group means varying from 34% to 61%. The English-speaking children seemed to be weakest for 3$^{rd}$ person singular inflection, but there is even a considerable difference between German- (50%) and Dutch-speaking (61%) children, although the inflection has exactly the identical surface pattern /-t/. Findings like this indicate that the surface account is not sufficient in explaining crosslinguistic differences. Instead, the proficiency with which children use a certain morpheme seems further to be influenced by the typology of the respective language.

Leonard and colleagues were inspired by the competition model from Bates and MacWhinney (1989). This model suggests that children can use different grammatical cues in processing language (e.g. word order, verb morphology…). Which of these competing cues provides the most relevant information is considered to differ across languages. In English, for example, word order is the most prevalent grammatical cue and English-speaking children are therefore considered to devote their resources largely to this cue. Verb morphology, on the other hand, is very sparse in English. Due to an economical use of their resources, English-speaking children pay consequently less attention to the word endings. The morphological richness account predicts that the resources of children with SLI that remain available for less salient cues is insufficient because of the children's processing limitations. Since fewer resources are available English-speaking children require a larger number of encounters with a morpheme until this is acquired. Children learning a language with a rich morphology, though, devote their processing resources first of all to this grammatical aspect. As a consequence, these children lag less behind their typically developing peers in morphology than for example English-speaking children with SLI. The morphological richness of a language determines hence how proficient or deficient the children are in this area. This theoretical framework can consequently close several gaps of the surface account and if both theories are taken together the morphological richness account can be considered as valuable complement.

#### *2.5.3.2.5 Phonological Theory*

Chiat (2001) proposed a theoretical framework of SLI that is based on a slightly broader view than that of the phonological storage hypothesis (Gathercole & Baddeley, 1989, 1990) or the surface account (Leonard, 1989, 1998). Also, according to Chiat's *Phonological Theory*, limitations in phonological processing abilities are at the root of developmental language impairments. The core of the phonological theory forms a mapping theory of language acquisition from which the disruptions caused by a phonological processing deficit are inferred.

According to Chiat (2001), mapping processes are an essential prerequisite for language acquisition and refer to the step of relating form and meaning to each other. Mapping processes can be divided into phonological, semantic and syntactic mapping or 'bootstrapping'. First of all, the child needs to identify word boundaries and to learn lexical items. The cues the child is using for this phonological segmentation are rhythmic and segmental characteristics, for example the occurrence of stressed or phrase- and clause-final lengthening. Secondly, semantic cues are taken from the context in which the utterance occurs and initiate semantic bootstrapping as such as the child can infer the meaning of the utterance from the context. According to Chiat, the less observable a word's meaning the more the child is dependent on the word form, its phonological cue. This in turn is even more relevant when it comes to the mapping of grammatical morphemes or function words to their meaning. It is only the phonological cue, and the discovery of such in the first place, that prompts the child to look for the particular context in which this cue occurs. For example, for the acquisition of person, number and tense marking in English present tense, the child must first notice the phonological variation of the verb with and without third person singular morpheme /-s/ (e.g. *walk / walks; play / plays*, etc.). Only then, the child will subsequently search for the context in which verbs carry the suffix /-s/ and discover subject-verb agreement in present tense. This leads eventually to syntactic bootstrapping when the child discovers syntactic relations between elements.

Chiat refers to phonology as the 'key player' in lexical and syntactic development (p.124). The child's ability to use prosody and phonological details as cues is central in the whole developmental process beginning with the segmentation of utterances, followed by the storage of words, the identification of their semantics and finally the discovery of their syntactic relations. In children with SLI, though, limitations in phonological processing abilities hamper the children's lexical and syntactic development due to their impact on mapping processes. Chiat does not define precisely the locus of interruption on phonological processing. The affected children are considered to have reduced access to phonological details within rhythmic structures which inhibits their perception, storage and/ or retrieval of speech units. On the other hand, any other level of phonological processing may possibly be affected additionally and impact on subsequent levels of phonological processes and development.

The focus of the phonological theory lies on the nature of the linguistic deficits that arise from the hypothesised limitations in phonological processing. In this sense, several predictions can be derived from the phonological theory regarding the impacts on the children's language development. First of all, it is likely that the children exhibit problems in phonological processing that are independent of semantic and syntactic factors. For example, the children may show phonological speech errors or experience difficulties in the discrimination, the analysis or the repetition of auditory input. Secondly, the disruptions in vocabulary, morphology or syntax reflect in how far the child had to rely on phonology in accessing the semantic features of the respective lexical, morphological or syntactical forms. The more dependent a semantic representation is on phonology the more difficult it is for the child with phonological processing limitations to acquire this representation. The principle of 'observability', for example, implies a certain hierarchy which elements are easier (e.g. verbs easily to observe (e.g. *run, eat, throw*,…)) or more difficult to acquire (e.g. verbs impossible to observe (e.g. *dream, guess, assume*,…)). However, the extent of disruption of forms on the same hierarchical level of semantic difficulty depends on the individual phonological surface pattern. Strong and salient phonological structures are less likely to be affected than phonologically weak forms. Hence, those aspects of language that are least crucial to the message (e.g. grammatical morphemes) conveyed are those most at risk to be affected by phonological processing limitations.

Although the wide scope of Chiat's phonological theory is certainly one of its strength it is also this factor that complicates the investigation of the theory's validity. Chiat provides a certain hierarchy in her outlined framework but it is difficult to derive predictions on grounds of the interdependencies between the individual factors or parameters. For example, studies demonstrated that children's vocabulary is more restricted on verbs than in nouns (see Leonard, 1998; Watkins, Rice, & Moltz, 1993). Single case studies of language impaired children demonstrated further that the children are not consistent in their use of the same verb. The very same verb may be affected in one but not in another context (Chiat, 2000; Rice & Bode, 1993). Black and Chiat (2003) propose several factors that may account for these disproportional difficulties in the acquisition of verbs. Some of these factors are language specific, others are relevant across languages. In English, the trochaic stress pattern is the most common one. However, more verbs than nouns deviate from that and have iambic stress pattern (e.g. *accept, repeat*). Moreover, English verbs often have fewer syllables than nouns and are consequently shorter. Another factor that shortens their duration in sentences is the position of verbs. Whereas nouns occur often in sentence initial or final position and are therefore easier to perceive, verbs occur frequently in medial position. Semantically, verbs are harder to map to their meaning than nouns as they represent relational concepts such as activities, changes of state, or causal relations. Nouns, in contrast, refer to objects, persons or other ‘things’ which are more easy to perceive. Another consequence of the relational reference of verbs is that they usually link a number of arguments with each other and that they are therefore more complex than nouns. According to Black and Chiat, it is not just one of these factors that causes greater difficulties in the acquisition of verbs than of nouns but the combination of all of these factors.

This multifactorial view on language development makes it difficult to present counter-evidence for the theory, as the range of possible symptoms seems extensive. To date, few empirical data are available that clearly contradicts the phonological theory. However, this is due to the wide ranging scope of the hypothesis which makes it difficult to pinpoint and subsequently contrast certain predictions. Difficulty to scrutinising the hypothesis does not imply its invalidity, though. The phonological theory can be seen as a promising framework that can successfully account for a wide range of linguistic deficits in children with SLI, and thus for the population’s heterogeneity, as well as for many cross-linguistic differences in SLI.

#### *2.5.3.2.6 Learning Deficit in SLI*

Recent models of language acquisition attach great importance to both genetic and environmental influences. Considering SLI, genetic factors are thought to influence the neurodevelopment and should therefore result in measurable irregularities in the human brain. Ullman and Pierpont (2005) provided a thorough review of neuroimaging and neurophysiological studies and concluded that children with SLI are indeed likely to present brain abnormalities that underlie procedural memory functions (especially in frontal / basal ganglia circuits). Ullman and Pierpont describe the language impairment of most children with SLI therefore as *procedural language disorder*. In contrast to the working memory, the procedural memory system is relevant for long-term knowledge and is thought to be involved in the implicit acquisition, storage and use of such knowledge (Gabrieli, 1998; Willingham, 1998). As the children experience difficulties with procedural learning it is especially hard for them to learn syntactic and phonological rules. It is, however, not the rule extraction itself that is considered to be difficult for the children, but rather the implicit learning of underlying structure from the input. This implicit learning is prompted by statistical features in the input such as a high frequency of sequences and a high variability in the input. Brain structures involved in declarative learning, on the other hand, were reported to be preserved. Rote learning is thus suggested to be relatively intact in children with SLI and becomes evident in an advantage for forms that are not predictable from rules, such as irregular forms. Declarative learning is further thought to be highly involved in the acquisition of vocabulary and of semantic knowledge. An advantage for irregular forms as opposed to regular morphemes had been described already in the early SLI literature (Gopnik & Crago, 1991) but also has been questioned since (Bishop, 1994; Gopnik, 1994; Serratrice, Joseph, & Conti-Ramsden, 2003). The clinical marker research, though, showing a marked deficit in the children's morphology supports the procedural deficit hypothesis.

Hsu and Bishop (2011) proposed another form of learning deficit in children with SLI. They suggested that not only procedural learning but statistical learning in general is deficient. Statistical learning is especially important within a framework that considers the acquisition of grammar not as the acquisition of rules but as the accumulation of probabilistic knowledge. The children acquire therefore knowledge that is based on statistical relationships, e.g. frequencies, of co-occurrences of particular structures. These relationships are discovered as the children search for statistical regularities in the input when listening to language. As a result, Hsu and Bishop predict that children with SLI show a high dependency on their language input. It is therefore unlikely that the existence of an underlying rule is crucial but that the frequency of the input is essential for children with SLI. Serratrice, Joseph and Conti-Ramsden (2003) demonstrated indeed that a dissociation between regular and irregular forms is no longer evident in the language of children with SLI if the input is controlled for frequency. This even larger input dependency in the statistical learning deficit hypothesis than in the procedural deficit hypothesis implies that also other language areas than grammar and phonology may be affected, such as vocabulary. Nevertheless, Hsu and Bishop question the origin of this deficit in statistical learning. While it is possible that this learning deficit lies at the root of SLI it is further possible that statistical learning is impacted by other deficits, such as in phonological processing or in working memory. Further research is therefore necessary in order to investigate these different pathways to language impairment.

#### *2.5.3.2.7 Conclusion on domain-general accounts of SLI*

As the overview over linguistic theories on SLI, the overview over domain-general accounts of SLI is limited to some of the most important theories. While assumptions on language acquisition may differ across the domain-general accounts of SLI, all theories have in common that it is the language input that cannot be proficiently perceived, stored, processed or computed and that these limitations are at the root of the secondary language impairment. One challenge to these theories is the accumulation of non-linguistic limitations in SLI which makes it hard to pinpoint one of these limitations as the origin of the language impairment. Furthermore, as Gillam and Hoffman (2004) state, cognitive deficits could be the cause of SLI as well as the recipient of their consequences. In view of an emergent framework, such interplay of cognitive processes could result in a very dynamic phenomenon of language impairment. However, research focussing on the genetics of SLI and neurobiological characteristics of SLI back-up the approach to consider the population of children with SLI nevertheless as one category. Nonword repetition, for example, has proven to be limited in most children with SLI across several languages and points into the direction of a common genetic origin. Most importantly, though, domain-general accounts offer an explanation for a wide variety of linguistic symptoms and seem to account better for cross-linguistic differences than linguistic accounts on SLI.

## 2.6 Speech Output in SLI

Section 2.5 illustrated that the literature suggests different types of clinical markers for SLI: processing markers such as nonword repetition and linguistic markers. Linguistic markers described in the literature refer often to the production of particular morphemes in the respective language. Thinking about factors that can cause such difficulties, a speech output disorder could obviously impact on both nonword repetition as well as the production of grammatical morphemes. It is therefore important to ask in how far speech output disorders that could account for the difficulties with potential markers indeed co-occur in children with SLI. However, maybe even more important is the question whether children who have an isolated speech output disorder without any further semantic or grammatical difficulties would be categorised as experiencing SLI or not. Both questions will be followed up below. First of all, though, some terminology will be specified.

*Speech disorder*

Speech (output) disorder is a general term that refers to different types of speech problems as described in Section 2.4. Three terms were introduced that currently are often used for differential diagnosis:

*Phonetic / articulatory difficulty:* difficulty with the pronunciation, thus the *production* of a particular sound, independent of its position in the word. The production of a sound depends on the appropriate movement of the articulators. According to Fox (Fox, 2003), only those speech disorders can unambiguously be categorised as phonetic difficulty if they do not involve the loss of a phonemic contrast. In German, only two articulatory error patterns qualify consequently as phonetic errors: a lisp, thus the realisation of /s, z/ as [θ, ð], or a lateralisation of /s, z/ or /ʃ/ as[ɬ]. For none of these substitutional sounds are a phoneme of German this does not result in the loss of a phonemic contrast. Consequently, an articulatory disorder, as defined by Fox, would not impact on either of the suggested clinical markers for SLI.

*Phonological difficulty*: breakdown in the *use* of sounds in their correct linguistic context irrespective of the children's ability to produce the sound. The child may be able to produce the sound in isolation or in particular word positions but omit or substitute the same sound in other linguistic contexts. It is commonly assumed that the child has an underlying problem in the formation of phonological representations of words and the mapping of these representations into speech forms (Ingram, 2008). As a consequence, phonemic contrasts are lost for affected sounds. However, the children's speech is not disorganised but results mostly in consistent phonological error patterns, such as stopping of certain fricatives, the reduction of consonant clusters etc.. The application of incorrect phonological rules can have potential impact on both nonword repetition and the production of certain morphemes. Phonological impairment is thus of particular interest for this study.

*Verbal dyspraxia / developmental apraxia of speech (DAS):* breakdown at the level of motor programming (Dodd, 1995; Stackhouse & Wells, 1997) resulting in an inconsistent and often severely impaired speech output. Effecting only approximately 3-5% of children with a speech output disorder (Shriberg, 1994), DAS compromises a relatively small proportion of speech disorders. Nevertheless, DAS impacts clearly on both potential markers and should be carefully diagnosed as differential diagnosis to both a phonological impairment and SLI.

In the past, these terms seem to have often been used interchangeably in the literature on SLI. For example, Stark and Tallal (1981) excluded children from the SLI population if they have an articulation disorder of greater severity than their language deficit (more than 6 months difference). Considering the definition of articulatory difficulty by Fox (2004), such a severe disorder could not possibly occur and would correspond far more likely to a phonological impairment or DAS. Pennington and Bishop (2009), too, describe a change of view in the literature. Originally, speech disorder was considered a functional articulation disorder. However, careful analysis of error patterns and the identification of deficits in phonological tasks led researchers to the conclusion that the majority of children with a speech output disorder have a form of phonological impairment. The terminology of speech output disorders in the literature on SLI quoted in this thesis is therefore treated with caution and partly regarded as ambiguous.

### *2.6.1 Comorbidity of speech disorder and SLI*

A child with a phonological impairment or DAS is likely to fail on nonword repetition but may also have difficulties with the production of grammatical morphemes. For example, if the child stops the fricative [s] to [t], the 2$^{nd}$ verb inflection –*st* in German present tense is no longer distinguishable from the 3$^{rd}$ person singular inflection. The co-existence of speech disorder and SLI is therefore a crucial aspect when drawing conclusions from the potential clinical markers for SLI. In order to avoid going in circles, however, it is in this respect important to use the term SLI to refer to deficits in language areas other than phonology. The studies cited in the following fulfil this criterion.

Only few studies investigated the comorbidity of speech disorder and SLI, i.e. the co-existence of both disorders in the same child. Some of these studies have been summarised in Table 1. One of the largest studies is the work by Shriberg, Tomblin and McSweeny (1999) using data from a large-scale study on the 'Epidemiology of Specific Language Impairment' by Tomblin et al. (1997; 1996). The diagnostic system 'EpiSLI' which was derived from this study has been reported in Section 2.2.1. Applied to the general population of 6-year-olds in the United States, results indicated a comorbidity of SLI and speech delay of only 0.51%. Speech delay was defined as functional articulation disorder or developmental phonological disorder of unknown origin. Comorbidity estimates vary according to the population they relate to. For example, the younger the children are the less likely it is that one or the other disorder has resolved. It is assumed that already 75% of preschool children with speech disorder have normalised their speech by an age of 6 years (Shriberg, 1994). The sample-wide comorbidity of SLI and speech output disorder is therefore higher in younger children. In 5-year-olds living in Canada, for example, 4.6% children are affected by co-occurring speech and language deficits (Beitchman, Nair, Clegg, & Patel, 1986), although no exclusion criteria guaranteed the specificity of the language deficit in this study.

Sample-wide comorbidity estimates from epidemiological studies mirror the prevalence of cases in the population sample with both disorders co-occurring. Another way to calculate comorbidity rates is with reference to an index disorder. For example, if considering in Shriberg et al.'s (1999) research the population of 6-year-old children *with SLI* as index disorder, 5-8% of these children showed a co-occurring speech disorder. Using speech disorder on the other hand as index disorder, 11-15% of these children were additionally affected by SLI. With regards to the influence on clinical marker research for SLI, those studies with SLI as index disorder seem more appropriate than those with a speech disorder.

Nevertheless, comorbidity rates vary even considerably even across studies referring to a population derived from the same index disorder. According to a literature review provided by Shriberg et al. (1999), estimates range from 9-75% in children ascertained by language disorder as index disorder. This wide range, mainly due to methodological differences, makes it difficult to draw conclusions. More recent studies have been selected and summarised briefly in Table 1.

Most of the studies presented in Table 1 were conducted by research groups around Shriberg and follow similar criteria in the definition of both speech and language disorders, i.e. cut-off points, which is a crucial precondition for comparison. Only two studies provide estimates for the comorbidity of speech and language disorders with SLI as index disorder (Shriberg & Austin, 1998, study 4; Shriberg et al., 1999), both using data from the EpiSLI study (J. B. Tomblin et al., 1997; J. B. Tomblin et al., 1996). Comorbidity rates in children with SLI vary from 5-9%. When analysing results for receptive and expressive language impairment[3] separately, both groups show a rather similar risk of having a co-occurring speech disorder (Shriberg & Austin, 1998, study 4). The

[3] see Section 2.4.2 for comments on the use of the term *receptive language impairment*

fact that studies with speech disorder as index disorders tend to report higher comorbidity rates (3-53%) than those with language impairment as index disorder, may be explained by the participants' age because the speech disorder may already be resolved. The data of the EpiSLI study is therefore constrained by the fact that the children had already an age of 6 years, probably implying a high number of resolved cases of speech disorders.

A comparison between the comorbidity rates reported by Shriberg and colleagues and Broomfield and Dodd (2004) highlights another important factor influencing the outcome. While Shriberg and colleagues selected an epidemiological sample on the basis of the EpiSLI criteria, Broomfield and Dodd categorised within 15 months all new referrals to pediatric speech and language therapy services in a British Primary Care Trust. In this respect, it is important to note that this sample did not only refer to general language impairment, and also that children with speech output impairment are more likely to receive speech and language intervention than children with language disorders only (Zhang & Tomblin, 2000). Studies which used clinically ascertained samples of children with (S)LI may therefore be biased towards children with co-existing language *and* speech impairments. It is therefore likely that true comorbidity rates may be lower than those reported by Broomfield and Dodd.

Shriberg and Austin (study 1-3) analysed their data further according to degree of severity. The authors found that an increased severity of speech disorder seemed to be related to an increased probability of receptive as well as expressive language problems, especially regarding grammar.

To sum up, many children show co-occurring speech and language difficulties, however, a substantial number of children have one disorder but not the other. The sample-wide prevalence of the co-existence of both disorders is markedly lower than the prevalence of both disorders considered independently of each other. Evidently, developmental phonological disorders and language impairment do not always co-occur, even in young children. It is therefore unlikely that a speech disorder can generally account for a clinical marker in SLI.

**Table 1.** Comorbidity Estimates for Speech and Language Disorders

| Study | n | Mean age | Sample wide | *Index disorder* | | Comments |
|---|---|---|---|---|---|---|
| | | | | Speech index | Language index | |
| Shriberg et al. (2005) | 72 | 4;9 | | 53% | | Cognition not tested. |
| Broomfield & Dodd (2004) | 730 | 0-16 | | 38,7 % receptive LI<br>25,6 % receptive-expressive LI | Receptive LI: 65%<br>Receptive-expressive LI.: 61% | No exclusion criteria applied, clinical sample. |
| Shriberg, Tomblin & McSweeny (1999) | 1328<br>51<br>108 | 6 | 0.51% | 11.1-15.2% | 4.8-7.6% | 1.3%<br>30-39% if not controlled for cognition<br>SLI criteria considered in study. |
| Shriberg & Austin (1998, Study1) | 58 | 4;2 | | 7% / 3%<br>Low receptive vocabulary / vocabulary + grammar<br>41%<br>Low expressive grammar | | Criteria for SLI considered.<br>Clinical speech disorder only. |
| Shriberg & Austin (1998, Study2) | 42 | 4;0 | | 9 % / 16%<br>Low receptive vocabulary / grammar<br>11%<br>Low expressive grammar | | Criteria for SLI considered.<br>Clinical speech disorder only. |
| Shriberg & Austin (1998, Study3) | 40 | 5;3 | | 26 % receptive LI<br>39% expressive LI | | Criteria for SLI considered.<br>Clinical speech disorder only. |
| Shriberg & Austin (1998, Study4) | 79 | 6;0 | | | 9 % | Criteria for SLI considered.<br>Clinical speech disorder only. |
| Beitchman et al. (1986) | 1655 | 5 | 4.56% | | | No exclusion criteria. |

### *2.6.2 Phonological characteristics in SLI*

Still, many, especially young children with SLI have additional problems in their speech output and it is worth to have a closer look at the phonology of these children. Several studies investigated expressive phonological profiles of children with SLI. Many of their authors used age controls as a reference group or compared their findings with data reported in the literature. Fewer studies employed comparisons between children with SLI and younger normally developing children matched for language competence. The general picture is that many of the phonological characteristics in children with SLI are also typical for younger, normally developing children.

Phonological characteristics were described in different languages. Bortolini and Leonard (2000), for example, provided data for English-speaking and Italian-speaking children. Both reported studies showed that there was a marked difference between children with SLI and younger, normally developing children matched for MLU and consonant inventory size. Overall, in both languages the children with SLI produced significantly fewer words accurately than their language controls and they showed a greater inconsistency in speech production indicating unstable phonological representation. Similar significant delays in phonology were reported for French- (Maillart & Parisse, 2006) and Spanish-speaking children with SLI (Aguilar-Mediavilla, Sanz-Torrent, & Serra-Raventos, 2002). However, rather than using measures of accuracy, Bortolini and Leonard focussed on structural characteristics of the children's phonology because the authors were interested in the impact of phonological limitations on the children's morphology. The analysed 'structural constraints' were consonant cluster reduction and weak syllable deletion, and for English also word-final consonant deletion. Both English- and Italian-speaking children showed a higher degree of consonant cluster reduction and weak syllable deletion. Due to language-specific differences the structural constraints were analysed differently in English and Italian concerning word positions. However, a significant difference between children with SLI and their language-matched controls was found for all studied variables. As Fee (1995) notes, word-final consonant deletion and consonant-cluster reduction in final word-position can even continue into adulthood. The studies carried out by Bortolini and Leonard provide strong evidence that children with SLI have serious limitations in their phonology that go beyond their limitations in utterance length and consonant inventory. Nevertheless, although many grammatical difficulties such as omission of tense marking resemble structural phonological difficulties, not all of them covaried with the seemingly related structural constraint (e.g. omission of infinitival *to* and initial weak syllable deletion). Phonological difficulties occurred further in contexts less relevant for grammatical morphemes, such as word initial consonant accuracy. According to the authors, it is therefore likely that deficits in phonology and grammatical morphology can co-occur without the former necessarily affecting the later.

### *2.6.3 The role of phonology in the identification of SLI*

Comparing results from comorbidity studies with results regarding the phonology of children with SLI, both seem to contradict with each other. Findings that children with SLI as a whole group fall behind even younger, normally developing children in their speech output would suggest higher comorbidity rates than those reported. One possibility to account for these contradicting findings would be major differences in inclusion and exclusion criteria. Obviously, in the case of SLI as index disorder in comorbidity studies, speech output deficits were no sufficient inclusion criteria to the language impaired group. However, children may have been included simply on grounds of their speech output in the studies on phonology in SLI. For these SLI studies, researchers applied very different inclusion criteria focussing on different language areas, partly also on phonology (e.g. Fee, 1995). Other researchers recruited their participants from clinical samples (Bortolini, Caselli, & Leonard, 1997). Both may result in a bias towards children with a speech output disorder co-occurring with the children’s language impairment.

Looking at the in the literature described markedly phonologically impaired subgroup of children with SLI, it is necessary to come back to the question which children should be included to the population of SLI. With regards to the possible impact on the performance on clinical marker tasks, it is of special interest how a child with an isolated phonological impairment should be categorised. In contrast to articulation, phonology can be considered as language domain. Thus, does a child having a deficit in this language area have a language impairment by definition, even if it is an isolated phonological impairment, or should an isolated phonological impairment rather be a differential diagnosis from SLI?

In order to answer this question, it is necessary to consider studies published on the identification of SLI. Guidelines for the identification of SLI which were formulated at a workshop on the phenotype of SLI recommend the consideration of all language areas, including phonology (Tager-Flusberg & Cooper, 1999). The biggest study to mention, on the other hand, is the EpiSLI system for the diagnosis of SLI in kindergarten children (J. B. Tomblin et al., 1996). In this large-scale study, the performance on language subtests assessing vocabulary, grammar and narration was crucial for the diagnosis of SLI, respectively in both comprehension and production. Diagnostic criteria yielded consistent results with clinical diagnosis and results of previous research. Accordingly, children with a pure phonological impairment were not considered as SLI. Another study examining the stability of two methods of defining SLI used measures for semantics, syntax and vocabulary but not for phonology either (Cole, Schwartz, Notari, Dale, & Mills, 1995). Even a third study, investigating the clinical and research congruence in the identification of children with SLI, used measures targeting mainly semantics and syntax but hardly phonology (Aram, Morris, & Hall, 1993).

A systematic review of studies on prevalence and natural history of primary speech and language delay found that children with “language delay in the absence of speech delay” were the group most commonly studied in the research reviewed (Law et al., 2000). A closer inspection of some of the reviewed studies, though, suggests that a speech delay did not necessarily have to be absent for the inclusion in the respective studies but rather that the authors of these studies simply did not consider speech delay in their participant selection (e.g. J. B. Tomblin et al., 1997). Yet, this review as well as the cited studies above show that an isolated phonological disorder is commonly not considered as SLI in research but as a separate diagnostic entity. This is also reflected if considering the question from the perspective of research on speech disorders: Authors who study children with speech rather than language problems do usually not refer to their participants as children with SLI (e.g. Broomfield & Dodd, 2004; Dodd, 2005; Lewis et al., 2006; Shriberg & Kwiatkowski, 1994; Shriberg et al., 2005; Stackhouse & Wells, 1997). It is therefore little surprising that both the DSM-IV manual (American Psychiatric Association, 2000) as well as the ICD-10 (World Health Organisation, 2010) adopted this distinction between SLI and disorders of the speech sound system.

### *2.6.4 Conclusion*

In summary, Stark and Tallal's (1981) criterion for SLI to exclude children with speech problems of greater severity than their language deficit seems to be too strict and is hardly considered in neither research nor clinical decisions. Although children with SLI are likely to have speech problems additional to their language problems researchers do not tend to identify children as having SLI if their only problem is the correct production of word forms. In accordance to these findings, children with an isolated phonological impairment will not be treated as SLI in the present study. Instead, an isolated phonological disorder will be regarded as a differential diagnosis to SLI. This implies furthermore, that a distinction should be made between errors that are likely to arise from SLI and errors that are likely to arise from a phonological impairment.

## 2.7 Conclusion

The present chapter illustrated several uncertainties in the classification and investigation of children with SLI. However, although the results presented above are not entirely congruent across studies, the review shows clearly that there are several characteristics that make children distinct from typically developing children and that these findings contributed considerably to our understanding of both typical and impaired language development. The reported differences in findings are likely to result in part from methodological differences across research projects. Samples vary for example in size, age, age range or even language background and different measures were employed as potential markers as well as gold standard. The differing outcomes of clinical marker studies highlight the fact that a marker is first of all only valid for the particular population sample employed in the respective study. Only replication studies can show whether the marker may be universal for one language and crosslinguistic research is essential in order to test the marker's applicability to other languages. Up to date, the number of languages investigated is too low in order to draw conclusions for the most promising candidate for a universal marker across languages. This may be a processing task, such as sentence repetition, a composite measure involving morphology or a general language measure tapping utterance length and lexical ability. Differing outcomes across studies that result from different gold standards, on the other hand demonstrate once more how important the diagnostic criteria are. However, there is an implied circularity here. Different gold standards weaken clearly the empirical evidence obtained from different studies, as it is more difficult to draw conclusions across these studies. The fact, though, that researchers used different gold standards in the first place highlights yet again the need for widely excepted, common diagnostic criteria. The search for a clinical marker tries to fill exactly this gap and faces here at the same time one of its greatest challenges.

The motivation for finding a clinical marker for SLI, however, goes beyond this gap in the diagnosis process. A valid clinical marker would not only be key symptom of the disorder but would further be important in the definition of a phenotype of SLI. This in turn can assist considerably in the search for a genetic base of SLI and in the identification of underlying causes. However, only a clinical marker that is universal across languages can comprise this significance.

# 3 Research Questions and Rationale

## 3.1 Hierarchical Model of Decisions about SLI

Looking back at the hierarchical model of decisions about the nature of SLI in Figure 1 in Chapter 2, many of the raised issues are difficult to answer clearly. Nevertheless, by studying children with SLI as one distinct group, researchers follow this path – often without questioning or reflecting on the decisions or assumptions made on other levels of the hierarchy that justify this approach. In more than three decades, a lot of resources have been drawn on in order to find the ultimate answers for English-speaking children. An alternative approach to verify the illustrated path in this hierarchy is applying it to another language. If results from other languages support the hypothesis that children with SLI can or should be considered as a distinct group it is worth exploring the questions further regarding English-speaking children. If data from other languages provide definite answers that lead to a dead end in the decision hierarchy it may be worth reconsidering the assumption of a distinct group with SLI. Alternatively, this assumption may be valid but language specific. Any conclusion for children with SLI from a particular language background could consequently only be applied to this particular population and cannot be generalised and considered as being universal.

Several levels of the hierarchy would be suitable as starting point for such a crosslinguistic approach. For example, the researcher could investigate children with language impairment in order to identify subgroups or to see whether this proves to be as difficult as in English, leaving us again as a whole with a heterogeneous group of primarily language impaired children. However, what advantages would this crosslinguistic information add to the theory of SLI? The evidence for English does not positively identify stable subgroups across different population samples; on the other hand, it does not reject the existence of subgroups either. Research on children learning a language other than English is therefore likely to add similarly ambiguous evidence. If cluster analysis can successfully categorise different groups, further research projects would be necessary in order to study the classification's specificity for the respective language, or its applicability to English and further languages in order to evaluate eventually its general validity for children with SLI of different language backgrounds. It is consequently more useful to start at another level of the hierarchy for crosslinguistic comparison; preferably at a level for which the data of English- speaking children is more consistent.

One level that seems suitable as such is the search for a clinical marker for SLI. The evidence for English-speaking children has been fairly consistent in showing a significant delay in linguistic markers regarding the production of certain morphemes or in processing markers such as nonword repetition or sentence repetition. The discovery of an effective clinical marker has two potential implications: First of all, it obviously could be used in the identification process of children with SLI and improve this process. Secondly, a clinical marker that is universal across languages would bring us one step further in the identification of a cause of SLI independent from the language to be learned. This understanding of the underpinnings of SLI can eventually offer a path for most effective intervention. The longitudinal impact of SLI on different aspects of life highlights this need for effective therapy methods. This goal seems highly relevant, too, given the findings from neurobiological research demonstrating that language skills as well as cerebral functions can be altered successfully by intervention (Pihko et al., 2007). The search for a clinical marker across languages is therefore an important starting point in order to improve our understanding of language impairment further.

Nevertheless, looking at SLI from an emergent, neuroconstructivist point of view, it seems questionable whether it is realistic to search for *one* common underpinning in SLI. So far, research could not identify one single factor that can account for the heterogeneous symptoms of children with SLI. Furthermore, a multifactorial and dynamic model of language acquisition as assumed in emergentist and neuroconstructivist frameworks suggests that influential factors interact with each other and may even change over time. Disentangling a causal relationship is therefore hindered. It appears as a result most reasonable to focus first on linguistic symptoms in the search for a clinical marker. The identification of a processing marker may bear more potential crosslinguistically but it has the disadvantage of being rather general. The identification of *one* particular linguistic skill or set of skills, on the other hand, would have three advantages. First of all, a linguistic marker is more useful in diagnostic procedures as it reflects those limitations which cause the children fall into the group of SLI in the first place. Clinicians as well as educational staff and parents could gain relevant knowledge to identify language impaired children from their language output. From this perspective, it would be even counterproductive if the clinical marker is evident in both acute as well as resolved cases of SLI as has been reported from studies on processing markers. Secondly, the results for SLI in English indicated that there is such a cardinal linguistic characteristic. The finding that very specific linguistic structures fail to emerge in other languages, too, may provide more concrete clues to the underpinnings than limitations in a processing task. In this way, it is clear that the marker would refer to the effect rather than to a measure of an underlying cause of SLI. The direction of a causal relationship would consequently be unambiguous and would possibly facilitate the search for the origin of the language limitations. Thirdly, linguistic markers appear more specific to SLI than processing markers. The applicability of the marker would therefore bear more potential in clinical contexts than an unspecific marker. In addition to these arguments based on the assumption that a clinical marker may exist, the opposite outcome may justify the search for linguistic rather than cognitive commonalities, too. First of all, if results demonstrate that a linguistic marker for SLI is unlikely to exist in another language than English the challenge to the construct SLI would be larger than finding inconsistencies regarding a processing marker. The latter would rather raise questions concerning the nature of the underlying cause than questioning the validity of combining children with different linguistic profiles within one single category. It seems consequently the most appropriate to start the search for a clinical marker within the linguistic domain.

## 3.2 SLI and Phonological Impairment

In order to identify an appropriate clinical marker for a certain population, consistent guidelines should be followed concerning the question of which individuals are part of this respective population. As section 2.5.2 in Chapter 2 illustrates, inconsistency is found across studies, though, with regards to the inclusion of children with speech impairment. Asking the question whether children with an isolated speech impairment should be considered as SLI is one step. If we conclude that these children should not be treated as part of the SLI populations, we should further be careful with the impact of children's speech difficulties on their language performance.

Section 2.6.4 concludes with the assumption that children with an isolated phonological impairment should be treated as a distinct clinical group rather than a subgroup of children with SLI. Data that call this conclusion into question comes from studies on subgroups of SLI, though. The only consistent outcome of these classification studies is the identification of one group with both receptive and expressive problems in addition to a group with expressive deficits only. This latter group with output problems only could possibly be or include a group of children with isolated phonological impairment. It is therefore important not to ignore this group of children. If a clinical marker can be identified for SLI it is crucial to ask the question whether children with isolated phonological impairment would be identified by it too, or whether these children would fall outside of the population with SLI. Eventually, an important goal of phenotyping a certain disorder is to differentiate it from other disorders. A clear distinction between SLI and PI is furthermore important regarding the underlying cause of SLI. As section 2.5.3 demonstrated, processing limitations have frequently been reported in children with SLI. Children with phonological impairment have by definition limitations in their phonological processing. It may be possible that the same cause results in either speech output problems or language problems that may possibly be accompanied by speech problems. However, in order to disentangle conclusions from these two populations about underlying deficits, children with SLI should preferably be investigated with children with isolated phonological impairment excluded. Results may otherwise be biased towards phonological processing deficits.

The second concern that needs to be raised is the impact of the children's speech errors. Even if children with an isolated phonological impairment are treated as a group distinct from the group of children with SLI, language impaired children can nevertheless experience concomitant phonological impairment. Consequently, speech errors of children with SLI should be treated carefully in error analyses. Morphological errors, for example, could be a result of phonological processes in the speech of children. If a child generally deletes word-final consonants the child is likely to omit inflections represented as word-final consonants, too. In this case we would return to the previous question: Are phonological errors *part* of the SLI profile or might they occur *additionally* to SLI? If children with isolated phonological impairment are considered as a group distinct from the SLI population, co-occurring speech errors should be treated in a different way than language errors. This distinction in turn must include speech errors which may affect other language areas and may seem superficially to be morphological or semantic errors. Only then we can draw conclusions on the *language* errors of the children with SLI.

## 3.3 German as Language for Investigation

The study presented here focuses on German-speaking children. The most obvious reason for choosing German is that this is the author's native language. Nevertheless, German seems to be an appropriate choice even with regards to the literature. As the morphological richness account (e.g. Leonard, 1998) illustrates, the typology of the language being acquired influences highly the appearance of language impairment. In comparison to English, German has a richer morphology and follows different rules for word order. Distinct characteristics are, for example, the heavy reliance on grammatical case or the fact that nouns can have different genders. These features make German a valuable candidate for research and imply that the focus on German-speaking children with SLI can add relevant information. Only the study of a variety of languages with different typologies can contribute to the identification of typological characteristics that are most relevant for the phenomenon SLI. To date, no clinical marker research for SLI has focussed on German.

The data that is available on German-speaking children with SLI is limited in comparison to SLI in English. First of all, far less studies have been carried out on German SLI. Secondly, even more striking, the number of participants is also comparatively low especially since several authors used data from existing corpora rather than collecting new data. Thirdly, few researchers compared children with SLI not only with age-matched children but also with children matched on language. This latter comparison is of particular importance for the search of a clinical marker and also for a number of accounts of SLI. Nevertheless, German-speaking children have been reported to experience larger difficulties in grammar than in other language areas (Grimm, 1993; Lindner & Johnston, 1992). Findings like this indicate that grammar may be the language area most promising for the search for a clinical marker for SLI. A potential grammatical marker would probably have a better practical applicability than a more general linguistic marker such as MLU or lexical diversity. The purpose of the current study is therefore to replicate reported findings on German SLI and secondly to provide the possibility of a comparison with language-matched children. This in turn intends to broaden the view of SLI from a crosslinguistic perspective.

## 3.4 Research Questions

In summary, the present study addresses the following research questions:

1. Does a grammatical clinical marker for SLI exist in morphology for German-speaking children?
2. If a marker exists, can this potential clinical marker discriminate only between children with SLI and normally developing children or also children with SLI and children with isolated phonological impairment (PI)?

# 4 General Methods

## 4.1 Introduction

In order to investigate the potential existence of a clinical marker for German SLI, a study was designed which investigated the expressive morphology of German-speaking children. The aim of the study was to identify linguistic key characteristics of German-speaking children with SLI and to study their potential in distinguishing between children with SLI, normally developing children and those with an isolated phonological impairment (PI). According to this aim, the language of children was studied across four different groups: children with SLI, children with PI, typically developing children of similar age and typically developing children of similar language level. The project follows thus a quasi-experimental research design. The participant groups as well as the procedure employed are presented below.

## 4.2 Language Area

German-speaking children with SLI were reported to experience larger difficulties in grammar than in other language areas (Grimm, 1993; Lindner & Johnston, 1992). Several grammatical characteristics have been identified in German-speaking children: case marking errors (Clahsen, 1989; Kany & Schöler, 1998), difficulties in verb finiteness in present tense (Clahsen, 1989; Clahsen, Bartke & Göllner, 1997), the omission of copula and auxiliaries (Clahsen, 1989), errors in the production of the past participle (Clahsen & Rothweiler, 1992), verb placement errors (Clahsen, 1989, 1991; Grimm, 1993; Grimm & Weinert, 1990; Hamann et al., 1998; Lindner, 2002; Roberts & Leonard, 1997; Kany & Schöler, 1998) and elicited plural marking, particularly in non-words (Clahsen, Rothweiler & Woest, 1992; Grimm, 1993; Grimm &

Weinert, 1990; Kany & Schöler, 1998). For the current study, a set of elicitation tasks was designed which involves morphology probes for two verb inflections in present tense (2nd and 3rd person singular) and definite articles in nominative, accusative and dative. Plural marking was not included in these elicitation tasks because this would have been circular to one of the subtests of the standardised language test SETK 3-5 (Grimm, 2001). Present perfect marking was excluded from the study, too, as no marked difficulties were to expect according to the literature (s. chapter 5 for details). An additional spontaneous speech sample was taken in order to analyse subsequently for the same morphemes as in the morphology screening and for further morphemes not covered by the elicitation tasks.

## 4.3 Participants

Sixty-six monolingual German-speaking children participated in the present study on German SLI and were distributed across four groups. Seventeen children aged 3;0-5;1 years met the criterion SLI (SLI group), 16 children had an isolated phonological speech output impairment and were age and gender matched to the children with SLI (PI group), 17 children were typically language developing and age and gender matched with the SLI group (ND-A group), and a further 16 children aged 2;6-3;11 years were typically language developing and matched with the SLI group on gender and language comprehension (ND-L). This design allowed to draw different conclusions: firstly, whether a potential difficulty is characteristic for SLI only or also for children with PI, and, secondly, whether such a difficulty may represent a *delay* (children with SLI are weaker than age-matched typically developing children) or a *delay-within-delay* (children with SLI lag behind both age-matched as well as language-matched typically developing children)[4].

The following section describes the procedure of the study and gives details on each of the participant groups. An overview across the four participant groups is

[4] See Section 2.5.1 for details on *delay* versus *delay within delay* as suggested by Rice (2003, p. 65)

provided in Table 7 at the end of the participant section and includes basic language measures derived from a 20-minute language sample taken from each child.

### *4.3.1 Recruiting process*

All children are monolingual German-speakers living in Hamburg, North Germany, or Marburg, in the county of Hessen. The identification process of the clinical groups (SLI, PI) covered several steps. All 33 children of these two groups were initially referred to speech and language therapists in Hamburg by paediatricians or paediatric audiologists as presenting difficulties in their speech or language development. Subsequently all children were diagnosed by qualified speech and language therapists through clinical assessment. For this, the respective clinician relied on information from a number of different sources such as speech and language tests and / or the analysis of a spontaneous language sample. Due to the limited number of available norm referenced tests at the time, the tests were not necessarily standardised (yet). For the interpretation of such data clinicians relied therefore on data on normal and disordered language development in German reporting the age of acquisition as published by e.g. Clahsen (1986) or Fox (2002). The diagnosis SLI was applied if the children had marked difficulties in grammar and / or vocabulary. All children were usually tested with a word-naming test in order to identify potential phonological processes and with a language test targeting a number of language areas such as the *SETK 3-5* (Grimm, 2001) or the *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen* (Christina Kauschke & Siegmüller, 2002) or language tests targeting a particular language area such as grammar (e.g. *TROG-D* (Fox, 2006) or vocabulary. Having passed this procedure, the children were recruited for the present study either from private clinics for speech and language therapy or from the Werner-Otto-Institute[5]. The children were preselected on the basis of the independent clinician's diagnosis SLI or PI. Further testing was then carried out in order to verify the children's

[5] Institute for child development, dedicated to the early diagnosis and treatment of developmentally challenged of disabled children and teenagers.

recruitment to participate in the study and their alignment to one of the clinical groups: SLI or PI.

In order to avoid an influence of intervention, children included in the present study had a maximum of five therapy sessions before they participated. The inclusion of children who are enrolled in an intervention programme may impact on the study's outcome depending on the language areas being targeted and the generalisation effect on other language areas. It is impossible to control for these potential effects and children with more than five therapy sessions (including sessions necessary for assessment) were consequently excluded from the study. The cut-off of five sessions was chosen in order to allow the recruitment from the caseload of speech and language therapists. The clinical diagnosis was hereby validated through the diagnosis of an independent clinician. Furthermore, the parents of children with SLI and PI were more approachable if the children were already enrolled in intervention. The exclusion of children with a longer history of speech and language intervention, though, had as consequence that the inclusion criterion for children with SLI had to be taken with less restriction. All children of the SLI group performed at minimum 1.1 standard deviations below their age average in at least one of three subtests of the SETK. This cut-off point corresponds to a maximum percentile of 13.57. Following the literature review in chapter 2, a cut-off point of -1.25 SD would have been preferable rather than -1.1 SD (see also Tomblin, Records & Zhang, 1996). However, as the children were required to be still very young in order to fulfil the criterion of not or hardly having undergone intervention, it was more difficult to find these children. It would have been easier to identify enough participants if they could have been already involved in therapy, but the individual intervention was considered to have a possible impact on the research outcome. Therefore, the cut-off point was instead lowered to 1.1 SD below the children's age average.

All normally developing children (ND) attended nurseries and were pre-selected by their kindergarten nurses as developing language normally. The children's language-status was confirmed on the basis of a standardised language-test carried out by the investigator.

### *4.3.2 Measures for participant selection*

#### *Language measures*

#### *SETK 3-5*

The current status of the children's language development was tested by using subtests of a test battery for German-speaking children aged 3 to 5: Sprachentwicklungstest für drei- bis fünfjährige Kinder (referred to hereafter as the *SETK 3-5*) (Grimm, 2001). Besides the *SETK 3-5* only three other standardised test batteries for German language development existed at the time of data collection. However, these tests were based on normative data collected several decades ago, did not provide the appropriate age range or were very long in administration (Angermaier, 1974; Grimm & Schöler, 1991; Häuser, Kasielke, & Scheidereiter, 1994). Accordingly, the *SETK 3-5* seemed to be the best choice.

The *SETK 3-5* consists of six subtests. Depending on the child's age, four subtests were selected and administered with each child. The subtests were scored individually and are of different complexity or target different language areas in the two age groups 3;00 – 3;11 and 4;00 – 5;11. For all administered subtests, t-scores were available referring to the normative data of the respective age group divided into 6 to 12 month intervals. The mean of these standard scores is 50 with a standard deviation of 10. Table 2 provides details of the subtests administered in the study. As it illustrates, two subtests were not administered to the entire sample of participants because of age restrictions. Children aged 4;0 years or above were asked to repeat sentences of increasing complexity for the subtest *Satzgedächtnis* (English: *Sentence Repetition* (*SR*)), whereas children aged 3;0 to 3;11 years were required to describe situation pictures for the subtest *Encodierung semantischer Relationen* (English: *Encoding of Semantic Relations* (*ESR*) instead.

Two subtest were not administered for participant selection: *Phonologisches Arbeitsgedächtnis für Nichtwörter* (English: P*honological Working Memory for Nonwords* (*PMN*)) and *Gedächtnisspanne für Wortfolgen* (English: *Memory Span for Word Series* (*MW*)). The *PMN* was excluded because individual speech patterns could influence the outcome especially for children with a phonological impairment. The *MW* was not administered because it is not based on normative data and consequently not standardised.

Concluding from this, three subtests were considered per age-group in order to determine the children's language status as age-equivalent or impaired. For the 3-year-olds, this was *SC*, *ESR* and *MR*; for 4- and 5-year-olds *SC*, *SR* and *MR* were considered.

### *ELFRA 2*

Some of the children of the ND-L group were too young to judge their language status from the norm data of the *SETK 3-5*. In order to ensure that these children were developing language normally, the standardised parental questionnaire *ELFRA-2* (Grimm & Doil, 2000) was filled in by the parents. The questionnaire provides three scores, respectively for expressive vocabulary, syntax and morphology. The vocabulary score reflects the size of the productive vocabulary. Children with an expressive vocabulary of 80 words or less are seen at risk. The maximum score for syntax is 47 and 16 for morphology. Scores must not fall below a critical value of 7 and 2 respectively.

### *Speech production*

### *PLAKSS*

For both clinical groups, additional data was available from the *PLAKSS* test (Psycholinguistische Analyse Kindlicher Sprechstörungen) (Fox, 2002), a picture naming task developed for the diagnosis of speech disorders in German-speaking children. This test was administered during the children's first assessment through a speech and language therapist and made available to the investigator. If this initial administration, dated back more than six weeks, the *PLAKSS* was repeated at the beginning of the data collection. The *PLAKSS* is based on the classification model of Dodd (1995) and enables the identification of phonological processes in the children's speech including their classification into developmental and deviant processes. The *PLAKSS* was not administered to the typically developing children. However, those phonological processes that could impact on the production of analysed morphemes were coded in the transcript of a spontaneous language sample for all four groups. Although the typically developing children were within age expectations regarding their speech production, some developmental phonological processes may still have been present. In accordance with the criteria of the PLAKSS, a phonological process was considered as active if it occurred at least twice. However, if the relevant phonological context occurred infrequently (three instances or less), one phonological error was sufficient in order to consider the respective phonological process as active.

### *Visual-motor skills*

#### *VMI*

In addition to the language and speech assessments, it was decided to administer the Developmental Test of Visual-Motor Integration (VMI) (Beery & Buktenica, 1989).

The Beery-Buktenica Developmental Test of Visual-Motor Integration (3rd Revision) requires the children to copy developmentally sequenced geometric shapes and assesses visual-motor integration. The test can be administered non-verbally and research during the standardisation has shown that the test can be used across different cultures. Furthermore, weak to moderate correlations with other measures suggest that the test can be used as indicator for performance IQ (Beery & Buktenica, 1989) and academic performance (Taylor Kulp, 1999).

Standard scores are given below for all participants from age 4 years, with a mean of 100 and a standard deviation of 15. For younger children, however, only age equivalents are provided (s. tables within the participant section below). The lowest raw score obtained on the VMI by the younger children of the current study, though, would have been equivalent to a standard score of 86 in 4-year-old children. The criterion of a performance within the age expectations is therefore considered as fulfilled in all participants both above and below the age of four years.

**Table 2.** Measures for Participant Selection

| | | |
|---|---|---|
| *SETK 3-5 - Expressive:* | | |
| Morphologische Regelbildung MR | *English:* | Application of morphological rules |
| | *Age:* | 3;0 – 3;11: real words<br>4;0 – 5;11: real words + nonwords |
| | *Task:* | plural production of given nouns |
| Encodierung semantischer Relationen ESR | *English:* | Encoding of semantic relations |
| | *Age:* | 3;0 – 3;11 |
| | *Task:* | description of situation pictures |
| Satzgedächtnis (SG) | *English:* | Sentence repetition (SR) |
| | *Age:* | 4;0 – 5;11 |
| | *Task:* | repetition of 15 sentences with increasing complexity (6-10 words): 9 sentences grammatically correct but semantically incorrect |
| *SETK 3-5 - Receptive:* | | |
| Verstehen von Sätzen (VS) VS / SC | *English:* | Sentence comprehension (SC) |
| | *Age:* | 3;0 – 5;11 |
| | *Task:* | comprehension of sentences (picture pointing task) and following instructions |
| *ELFRA 2* | | |
| Standardised parental questionnaire | *Age:* | 2;0 |
| | *Content:* | Scores for expressive vocabulary, syntax and morphology |
| | *Filled in:* | only by parents of 2-year-old participants of ND-L group (instead of *SETK 3-5*) |
| *PLAKSS* | | |
| Picture naming task | *Aim:* | diagnosis of speech disorders / analysis of phonological processes |
| | *Administered:* | with SLI and PI group |
| *VMI* | | |
| Developmental Test of Visual-Motor Integration | *Task:* | copying geometric shapes |
| | *Administered:* | all participants; standard scores > 4 years |

***Measures for matching***

The literature review presented in chapter 2 illustrated that the research on clinical markers in SLI involved group comparisons. The starting-point for exploration of any area of dysfunctioning is the comparison with what could be expected at the children's chronological age. In other words, children with SLI are compared in the area of interest with a group of children of similar age. Furthermore, no difference is expected to the comparison group either, with respect to other areas of development, such as cognitive, physical or emotional development. If no significant difference for a language measure can be found between the experimental group and the age-matched controls, then the investigated variable is not a candidate for a clinical marker. If a difference is found on that variable, though, it is worth having chosen a 3-group design which enables further comparisons with another group of children, e.g. children who are on a similar language level as the children with SLI or another clinical group. Comparisons with younger, typically developing children allow delay versus deviance interpretations as described in greater detail in Section 2.5.1. The rationale behind the inclusion of another clinical group, such as the PI group, is that it gives indications about whether the candidate marker is likely to be specific to children with SLI. As both questions are relevant in the search for a clinical marker, four instead of three groups were included in the present study.

A common language matching tool in research on SLI is mean length of utterance (MLU). MLU is commonly measured either in words (MLU-W) or in morphemes (MLU-M). The main purpose of the present study was to investigate the grammar of German-speaking children for the potential existence of a clinical marker for SLI. According to de Jong (1999, p.42), a suitable matching measure should fulfil two criteria: "it should be relevant with respect to the dependent variable (…) and its *value* should be independent from the value of the dependent variable." Applied to the context of the present study, this implies that the matching tool should be relevant to the morphological development but on the other hand, it should not be dependent on morphology. This latter criterion, however, would not be fulfilled with MLU as matching measure because MLU is an approach which evaluates morphology in a quantitative manner. The more morphemes a child produces, the higher is the child's MLU.

MLU-W would be slightly less redundant than MLU-M in relation to the children's morphology but there is nevertheless a strong link between MLU-W and morphology. The production of a larger quantity of words involves a higher proportion of function words and creates more potential contexts for the affixation of morphemes. MLU was consequently considered as too dependent on the dependent variables of the present study and therefore not used as matching measure. The younger, normally developing children were instead matched with the children with SLI on the basis of the children's language comprehension as measured with the subtest SC of the SETK 3-5. Measuring language comprehension does not comprise the measurement of the production of morphemes but the comprehension of morphemes is relevant for both the production of morphemes and language comprehension in general.

Mervis and Robinson (2003) raise a further issue concerning group matching. Any matching process is based on the assumption that both matched groups do not differ on the respective matching variable, i.e. the acceptance of the null-hypothesis. Although this variable has been controlled for on a certain measure tool (e.g. age scale, language test, etc.) and statistical analysis did not indicate a significant difference on the matching variable, any assumption of this kind bears the risk of a type II error. The type II error would entail that the acceptance of the null-hypothesis is incorrect and that the groups differ on the control variable. There is no way to entirely rule out this type II error but the higher the alpha value (*p* value) in the statistical comparison of both groups the less likely it is that the assumption is not true. A *p* value of just above 0.05 (i.e. rejecting a significant group difference) may not be sufficient. Mervis and Robinson illustrate that a *p* value of 0.5 or above is preferable in order to consider two groups as matched. The *p* values are provided in the sections describing the individual control groups.

### *4.3.3 Participant groups*

#### *4.3.3.1 SLI group*

The SLI group consisted of 12 boys and 5 girls with SLI and ranged in age from 3;0 to 5;1 with a mean age of 3;11. All children of the SLI group performed at a minimum of 1.1 standard deviations below the mean for their age in at least one of the three SETK 3-5 subtests considered for grouping (SC, ESR and MR for 3-year-olds or SC, SR and MR respectively for 4- and 5-year-olds).

The participants also met the following criteria:

- Children demonstrated a score of 85 or above on the Developmental Test of Visual-Motor Integration (VMI) (Beery & Buktenica, 1989).
- Children had passed their routine hearing screening with the paediatrician and did not have a history of a long-term recurring otitis media as reported by the child’s parents.
- Children had no history of neurological dysfunction as reported by the child’s parents, nurse or paediatrician.
- Children did not present emotional or behavioural problems as reported by the child’s parents, nurse or paediatrician.
- Children did not have peripheral oral motor or sensory deficits or any oral facial anomalies as diagnosed by the speech and language therapist.

Table 3 gives the children’s individual scores for each SETK subtest as well as their score of the VMI. MLU and NDW values are presented in Table 7.

In order to avoid an intervention effect, the children were on the waiting list for intervention or had undergone a maximum of 5 session of their treatment programme when examined for this study. Nine of the 17 children were additionally classified as having a phonological delay (Dodd, 1995; Fox, 2004) as they showed developmental phonological processes (3 of the children only minimal). Another three children showed deviant phonological processes and for this reason were classified as experiencing a phonological disorder additional to their language problem. A further two children did not show speech errors. The remaining three children of the SLI group had a lexicon too limited for phonology assessment.

**Table 3. SLI-Group: SETK- and VMI-scores**

| *SLI-Group* | | | *SETK 3-5 / t-score* | | | | *VMI* |
|---|---|---|---|---|---|---|---|
| **Code** | **Gender** | **Age** | **SC** | **ESR** | **SR** | **MR** | |
| SLI1 | m | 3;0 | 46 | **38** | - | 48 | /3/ |
| SLI2 | f | 3;0 | 41 | **28** | - | **30** | /2/ |
| SLI3 | m | 3;1 | **33** | **31** | - | **30** | /2/ |
| SLI4 | m | 3;1 | **35** | **38** | - | **36** | /2/ |
| SLI5 | m | 3;1 | 44 | 55 | - | **39** | /2/ |
| SLI6 | f | 3;7 | **39** | 40 | - | **39** | /2/ |
| SLI7 | m | 3;7 | 45 | 41 | - | **39** | /2/ |
| SLI8 | m | 4;0 | 45 | - | **38** | 43 | 98 |
| SLI9 | m | 4;1 | 40 | - | **<20** | **26** | 98 |
| SLI10 | f | 4;1 | 50 | - | **37** | 47 | 106 |
| SLI11 | m | 4;1 | **31** | - | 58 | 48 | 106 |
| SLI12 | m | 4;3 | **37** | - | **39** | 48 | 100 |
| SLI13 | f | 4;8 | 45 | - | **31** | **31** | 88 |
| SLI14 | m | 4;10 | 41 | - | **32** | 48 | 90 |
| SLI15 | m | 5;0 | **39** | - | **37** | 41 | 90 |
| SLI16 | m | 5;0 | **39** | - | **33** | **37** | 85 |
| SLI17 | f | 5;1 | 43 | - | 42 | **36** | 96 |
| | **Mean** | **3;11** | **40.76** | **38.71** | **36.7** | **39.18** | |

*Note.* standardised SETK-scores with mean=50; SD=10
f=female; m=male
SC=Sentence Comprehension; ESR=Encoding of Semantic Relations;
SR=Sentence Repetition; MR=Application of Morphological Rules;
VMI=Developmental Test of Visual-Motor Integration, mean=100, SD=15

#### *4.3.3.2 Phonologically impaired control group*

Eleven boys and 5 girls served as the phonologically impaired control group (PI group) and were selected on the basis of their phonological problems. According to a recently administered PLAKSS and following the classification system of Fox (2004) and Dodd (1995), eight children were diagnosed as exhibiting a phonological delay with developmental phonological processes, and another eight children were found to have a phonological disorder with deviant phonological processes. Children were not excluded if they had additional articulation problems, such as a lisp. An isolated articulation disorder, however, did not qualify as phonological impairment.

The participants of the PI group also met the following criteria:

- Children performed above -1 standard deviation on the subtests SC, ESR and MR (3-year-olds) or SC, SR and MR (4- and 5-year-olds) of the SETK. Accordingly, none of the PI subjects met the criteria for SLI.
- Children demonstrated a score of 85 or above on the Developmental Test of Visual-Motor Integration (VMI) (Beery & Buktenica, 1989).
- Children had passed their routine hearing screening with the paediatrician and did not have a history of a long-term recurring otitis media as reported by the child’s parents.
- Children had no history of neurological dysfunction as reported by the child’s parents, nurse or paediatrician.
- Children did not present emotional or behavioural problems as reported by the child’s parents, nurse or paediatrician.
- Children did not have peripheral oral motor or sensory deficits or any oral facial anomalies as diagnosed by the speech and language therapist.

The children of the PI group were chronological age and gender matched to the children of the SLI group. Each child of the PI group was within 4 months of age of a gender-matched child in the SLI group. However, no matching boy was found for one child of the SLI group (SLI5). The consequence of having different numbers of participants in the groups is acknowledged in the statistical approaches adopted. Using the independent *t*-test for group comparison regarding the children's age, alpha (*p*-value) is high and indicates that the children do not differ in age ($t(31) = -.395$, $p=.696$, r=.07) and that a type II error is very unlikely ($p > .5$).

Similarly to the SLI group, all children were on the waiting list for intervention or had undergone a maximum of 5 session of their treatment programme when examined for this study.

Table 4 gives the children's individual scores for each SETK subtest as well as their score of the VMI. MLU and NDW values are presented in Table 7.

**Table 4. PI-Group: SETK- and VMI-scores**

| *PI-Group* | | | *SETK 3-5 / t-score* | | | | *VMI* |
|---|---|---|---|---|---|---|---|
| **Code** | **Gender** | **Age** | **SC** | **ESR** | **SR** | **MR** | |
| PI1 | m | 3;0 | 47 | 47 | - | 48 | /3/ |
| PI2 | f | 3;4 | 60 | 52 | - | 56 | /4/ |
| PI3 | m | 3;1 | 44 | 48 | - | 59 | /3/ |
| PI4 | m | 3;3 | 54 | 50 | - | 51 | /4/ |
| PI5 | f | 4;1 | 53 | - | 50 | 65 | 117 |
| PI6 | m | 3;7 | 59 | - | 47 | 48 | /3/ |
| PI7 | m | 3;8 | 70 | - | 51 | 59 | /4/ |
| PI8 | m | 4;5 | 47 | - | 43 | 45 | 100 |
| PI9 | f | 3;11 | 45 | - | 46 | 62 | 122 |
| PI10 | m | 3;10 | 45 | - | 58 | 49 | 93 |
| PI11 | m | 4;7 | 60 | - | 49 | 49 | 112 |
| PI12 | f | 4;11 | 60 | - | 52 | 63 | 102 |
| PI13 | m | 4;6 | 60 | - | 49 | 55 | 117 |
| PI14 | m | 5;2 | 65 | - | 57 | 53 | 112 |
| PI15 | m | 5;0 | 49 | - | 51 | 53 | 96 |
| PI16 | f | 5;1 | 43 | - | 43 | 51 | 102 |
| | **Mean** | **4;1** | **53.81** | **49.25** | **49.67** | **54.13** | |

*Note.* standardised SETK-scores with mean=50; SD=10
f=female; m=male
SC=Sentence Comprehension; ESR=Encoding of Semantic Relations;
SR=Sentence Repetition; MR=Application of Morphological Rules;
VMI=Developmental Test of Visual-Motor Integration, mean=100, SD=15

#### *4.3.3.3 Age-matched control group*

Twelve boys and five girls served as age matched controls (ND-A). These typically developing children were chronological age and gender matched to the children with SLI. All children performed within one standard deviation of the mean or above on the SETK 3-5 and demonstrated a score of 85 or above on the VMI. The children's chronological age fell in a range of 3 months from the age of the corresponding gender-matched child in the SLI group. Using the independent t-test for group comparison regarding the children's age, alpha is very high and indicates that the children do not differ in age
($t$ (32) = -.114, p = .910, r = .02) and that a type II error is very unlikely.

No speech assessment was administered to the children of the ND-A group. However, spontaneous language samples were transcribed for each child. None of the children had a speech delay or disorder, i.e. phonological processes present were within age expectations (Fox, 2004).

Table 5 gives the children's individual scores for each SETK subtest as well as their score of the VMI. MLU and NDW values are presented in Table 7.

**Table 5. ND-A-Group: SETK- and VMI-scores**

| *ND-A-Group* | | | *SETK 3-5 / t-score* | | | | *VMI* |
|---|---|---|---|---|---|---|---|
| **Code** | **Gender** | **Age** | **SC** | **ESR** | **SR** | **MR** | |
| ND-A1 | m | 3;0 | 56 | 62 | - | 53 | /3/ |
| ND-A2 | f | 3;1 | 67 | 59 | - | 56 | /3/ |
| ND-A3 | m | 2;11 | 51 | 50 | - | 62 | /2/ |
| ND-A4 | m | 3;2 | 74 | 66 | - | 72 | /3/ |
| ND-A5 | m | 3;4 | 64 | 65 | - | 62 | /3/ |
| ND-A6 | f | 3;7 | 47 | 53 | - | 44 | /6/ |
| ND-A7 | m | 3;9 | 64 | 79 | - | 72 | /6/ |
| ND-A8 | m | 3;11 | 40 | 53 | - | 51 | /3/ |
| ND-A9 | m | 4;0 | 69 | - | 69 | 72 | /6/ |
| ND-A10 | f | 3;11 | 70 | 71 | - | 72 | 112 |
| ND-A11 | m | 4;3 | 74 | - | 67 | 65 | 118 |
| ND-A12 | m | 4;3 | 57 | - | 67 | 56 | 107 |
| ND-A13 | f | 4;9 | 65 | - | 66 | 56 | 112 |
| ND-A14 | m | 4;7 | 62 | - | 51 | 55 | 112 |
| ND-A15 | m | 4;11 | 49 | - | 46 | 69 | 90 |
| ND-A16 | m | 4;11 | 49 | - | 56 | 46 | 85 |
| ND-A17 | f | 5;3 | 59 | | 59 | 63 | 97 |
| | **Mean** | **4;0** | **59.82** | **62.0** | **60.13** | **60.35** | |

*Note.* standardised SETK-scores with mean=50; SD=10
f=female; m=male
SC=Sentence Comprehension; ESR=Encoding of Semantic Relations;
SR=Sentence Repetition; MR=Application of Morphological Rules;
VMI=Developmental Test of Visual-Motor Integration, mean=100, SD=15

#### *4.3.3.4 Language-matched control group*

Eleven boys and five girls, developing language normally, served as language-matched controls (ND-L). The children were gender matched to the children with SLI and additionally matched on the basis of their performance on the subtest *sentence comprehension* (SC) of the SETK 3-5. Since the test of sentence comprehension does not provide age-equivalent scores it was decided to use the children's raw scores instead. All children's raw scores reflected exactly the raw score of their gender-matched counterpart of the SLI group. In total, 31 children had been tested on the *SC.* Sixteen children matched to one of the children with SLI regarding sentence comprehension and gender. For child SLI1, no language-matched boy could be found. This is acknowledged in the statistical approaches adopted. Using the independent t-test for group comparison regarding sentence comprehension, alpha is very high and indicates that the children do not differ in sentence comprehension ($t$ (31) = .080, $p$ = .937, $r$ = .01) and that a type II error is very unlikely.

The children's mean age was 3;3 years, with a range from 2;0 to 3;11. Using the one-tailed independent t-test, alpha indicates that the language-matched children are significantly younger than the children with SLI ($t$ (31) = 2.845, $p$ = .004) with a medium effect size ($r$ = .46). All children performed within normal range on the SETK 3-5 or, in the case of the 2-year-olds, obtained an age-appropriate score on the ELFRA-2 (Grimm & Doil, 2000). No speech assessment was administered to the children of the ND-L group. Nevertheless, spontaneous language samples were transcribed for each child. None of the children had a speech delay or disorder, i.e. phonological processes present were within age expectations (Fox, 2004). All children had a score of 85 or above on the Developmental Test of Visual-Motor Integration (VMI) (Beery & Buktenica, 1989). However, one child of this group refused to complete the VMI. Table 6 gives the children's individual scores for each SETK subtest as well as their score of the VMI. MLU and NDW values are presented in Table 7.

Five children of the ND-L group do not live in the same geographical area as the children with SLI, Hamburg in North-Germany. The five children were recruited from nurseries in the centre of Germany, in Marburg. Nevertheless, these children did not have a regional dialect that was different from the other participants.

The SC-subtest differs over the age groups. Consequently, six children were required to complete the version for 4- and 5-year-olds although they were still aged 3 years, one further child was only 2;11 when tested on this version of the test. Five 2-year-olds were required to complete the test version for 3-year-olds. Nevertheless, all children were able to understand the test instructions and obtained scores equivalent to the older children with SLI.

**Table 6.** ND-L-Group: SETK- , ELFRA- and VMI-scores

| *ND-L-Group* | | | *SETK 3-5 / t-score* | | | *ELFRA-2* | *VMI* |
|---|---|---|---|---|---|---|---|
| **Code** | **Gender** | **Age** | **SC** | **ESR** | **MR** | | |
| ND-L1 | f | 2;6 | 46 | - | - | 250/41/14 | /2/ |
| ND-L2 | m | 2;0 | 41 | - | - | 168/26/7 | - |
| ND-L3 | m | 2;10 | **33** | - | - | 236/44/12 | /2/ |
| ND-L4 | m | 2;11 | **35** | - | - | 230/47/15 | /3/ |
| ND-L5 | f | 2;5 | 44 | - | - | 187/42/15 | /3/ |
| ND-L6 | m | 3;3 | **39** | 59 | 59 | | /3/ |
| ND-L7 | m | 3;11 | 45 | 54 | 62 | | /7/ |
| ND-L8 | m | 2;11 | 45 | 53 | 51 | 227/44/13 | /2/ |
| ND-L9 | f | 3;6 | 40 | 56 | 62 | | /5/ |
| ND-L10 | m | 2;10 | 50 | 46 | 59 | 220/39/16 | /3/ |
| ND-L11 | m | 3;11 | **31** | 57 | 56 | | /5/ |
| ND-L12 | f | 3;6 | **37** | 61 | 72 | | /5/ |
| ND-L13 | m | 3;9 | 45 | 54 | 62 | | /3/ |
| ND-L14 | m | 3;11 | 41 | 51 | 51 | | /5/ |
| ND-L15 | m | 3;11 | **39** | 64 | 59 | | /4/ |
| ND-L16 | f | 3;11 | **39** | 63 | 53 | | /6/ |
| | **Mean** | **3;3** | **50.36** | **56.18** | **58.73** | **217/40/13** | |

*Note.* standardised SETK-scores with mean=50; SD=10
f=female; m=male
SC=Sentence Comprehension; ESR=Encoding of Semantic Relations;
ELFRA-2=Elternfragebogen für die Früherkennung von Risikokindern 2
VMI=Developmental Test of Visual-Motor Integration, mean=100, SD=15

### 4.3.4 Overview of participant groups

Table 7 gives an overview over the four participant groups. Between group comparisons were carried out using ANOVA regarding three measures obtained from the spontaneous language samples of the children: mean length of utterances in words (MLU-W), mean length of utterances in morphemes (MLU-M) and number of different word roots within the first 100 utterances of the language sample (NDW). The homogeneity of variance assumption was violated for all three variables. Therefore, Games Howell procedure was used as post-hoc test. The preliminary descriptive results of between group comparisons are presented in Table 7.

### 4.3.5 Preliminary descriptive results

Significant group effects were evident for all three measures used for between group comparisons:

MLU-W ($F$ (3, 62) = 7.74, $p$ = .000, $\omega^2$ = .23), MLU-M ($F$ (3, 62) = 6.61, $p$ = .001, $\omega^2$ = .20) and NDW ($F$ (3, 48.57) = 7.91, $p$ = .000, $\omega^2$ = .24). This revealed that the children with SLI had a significantly lower MLU-W as well as MLU-M than the age-matched controls developing language normally (MLU-W: $p$ = .002; MLU-M: $p$ = .002) and also than their age-matched peers with phonological problems (MLU-W:

$p$ = .014; MLU-M: $p$ = .040). Furthermore, the group of children with SLI showed a significant lower NDW compared to the age-matched controls ($p$ = .001) and the phonologically impaired children ($p$ = .037). The SLI group did not differ significantly on these basic measures compared to the language-matched control group. Regarding the second clinical group, children with PI did not differ significantly from the age-matched children developing language typically in MLU-W, MLU-M or NDW. They also did not differ from the language matched control group on these measures. Looking at the mean scores of these measures, the PI group fell in between both groups of typically developing children.

**Table 7.** Participant Groups

| ***Group*** | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|
| ***n*** | *17* | *16* | *17* | *16* |
| ***Age*** | 4;00 (0;91)<br>Range: 3;0-5;1 | 4;1 (0;88<br>Range: 3;0-5;1 | 4;0 (0;88)<br>Range: 2;11-5;3 | 3;3 (0;76)<br>Range: 2;0-3;11 |
| ***MLU-M*** | $2.87\ (0.95)_{a,b}$ | $3.71\ (0.77)_{a}$ | $3.93\ (0.46)_{b}$ | 3.39 (0.68) |
| ***MLU-W*** | $2.48\ (0.76)_{a,b}$ | $3.25\ (0.59)_{a}$ | $3.35\ (0.36)_{b}$ | 2.96 (0.51) |
| ***NDW*** | $82.24\ (33.35)_{a,b}$ | $110.07\ (20.58)_{a}$ | $121.41\ (16.48)_{b,c}$ | $101.75\ (23.10)\ _{c}$ |

*Note.* Standard deviations in brackets. Means in a row sharing subscripts are significantly different when analysed with ANOVA.

SLI = Specific Language Impairment Group; PI = Phonologically Impaired Group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group, n = Number of participants
MLU-M = Mean length of utterance in morphemes; MLU-W = Mean length of utterance in words
NDW = Number of different words in first 100 complete and intelligible utterances

## 4.4 General Procedure

### *4.4.1 Experimental measures*

Two fundamentally different procedures have been employed for data collection in the present study: spontaneous language samples and specifically designed elicitation tasks.

Traditionally in clinical marker research, the most common procedure has been the recording, transcription and analysis of spontaneous language. This procedure is known to provide the most typical picture of the child's language behaviour (Wren, 1985). Free play is a method often chosen to elicit spontaneous language, but others, such as interviewing the child or story telling, have been used too (Schöler et al., 1998b). Those methods, however, that are structured to a greater extend are more popular with older children than with pre-school children as young as in the present study. Young children may be restricted in task comprehension. Another possible problem with a structured and rather demanding communication situation is that children with language problems are often aware of their problems. Due to their awareness of the language deficit, language impaired children may not talk and answer questions as freely as typically developing, eloquent children. Such a lack of spontaneity, however, may lead to an unrepresentative overview of the language structures that the child typically produces. Comparisons of different techniques to elicit spontaneous language showed that free play situations tend to provide a higher number of utterances compared to retelling of stories (Southwood & Russell, 2004). It is important to distinguish in this respect between language impaired and typically developing children. Whereas data of typically developing children showed less complex language structures for free play than for narration (Southwood & Russell, 2004), conversational situations were reported to facilitate more complex verb forms than narrations in language impaired children (Wagner, Nettelbladt, Sahlén, & Nilholm, 2000).

Nevertheless, data collected by means of spontaneous language is not standardised and difficult to compare. Spontaneous language samples also lag behind in the number of obligatory contexts of many morphemes if they are compared to more structured elicitation tasks (Polite & Leonard, 2007). The researcher is often left with a gap between what language impaired children are capable of doing and what they spontaneously do.

More recently, therefore, investigations have employed specifically designed elicitation tasks as an alternative method to spontaneous language samples (Bedore & Leonard, 2001; Bortolini et al., 2006; Conti-Ramsden, 2003). These probe tasks usually aim to elicit a maximal production of the studied morphemes and consequently offer a maximum of obligatory contexts. Furthermore, due to the tasks' highly structured nature, results of elicitation tasks are easy to compare. Hence, this procedure overcomes two disadvantages of spontaneous language samples: the results' comparability and the risk of a low number of obligatory contexts. Yet, children's performance on highly structured tasks does not necessarily reflect the child's typical performance. For example, studies comparing language performance observed in spontaneous language and those produced during elicited imitation tasks showed that the highly structured elicitations can not present a consistent prediction of the children's spontaneous language use (Connell & Myles-Zitzer, 1982; Fujiki & Brinton, 1987; Kany & Schöler, 1998). The error frequency has been reported higher in elicited utterances than in spontaneous language (Kany & Schöler, 1998), but Fujiki and Brinton reported that the performance on two so different measures correlates only for some children. Even the choice of specifically designed elicitation tasks can therefore not always bridge the gap satisfactorily between the children's maximal ability and the children's typical production of morphemes.

According to Wren (1985), any researcher has to face these two general conflicts in language data collection. Firstly, an increase in the representativeness of the child's typical use of language will result in a decrease of the results' comparability. Secondly, typical language performance in natural contexts does not necessarily mirror maximal language performance. Both conflicts can be considered as bi-directional. For the purpose of the present study, two different procedures have been chosen in order to account for both typical and maximal performance. A morphology elicitation task was designed that targets the research variables in particular and provides obligatory contexts for these. On the other hand, a spontaneous language sample was recorded, as well, for analysis of the children's typical language use. An approximate length of 20 minutes was chosen as often suggested in the literature (e.g. Dunn, Flax, Sliwinski, & Aram, 1996; Gavin, Klee, & Membrino, 1993) and a most natural setting was preferred in order to allow the children to feel the least pressure to produce language.

### *4.4.2 Procedure*

The first contact between the normally developing children and the examiner took place during the children's kindergarten programme. The examiner participated in the programme and joined free play situations in order to make contact with the selected children before starting the testing. The participation of the children of the SLI and PI group involved greater effort of the parents as they had to turn up with the child to two extra appointments for testing. For this reason, children of the clinical groups did not have an additional session to familiarise with the examiner but the first minutes of the initial meeting were used for the child to settle in the situation before the testing started.

All children were seen in two sessions and were tested individually in a quiet room with the examiner and sometimes with one or both parents. Most clinical cases were seen in a therapy room, although the examiner visited a few children at home. The typically developing (TD) controls were tested at their kindergarten. Each session's duration varied between 30 and 45 minutes. All children were rewarded with stickers or could choose a little gift at the completion of each session.

During the first session, the SETK 3-5 subtests were administered in order to assess the children's language skills and to verify their participation and grouping in the study. Additionally, most of the children completed the VMI. If the last assessment on the PLAKSS dated more than 6 weeks back this test was administered, too, during the first session with children of the two clinical groups. The second session served for the remaining tasks of the study: the elicitation tasks and the spontaneous speech sample.

## 4.5 Elicitation Tasks

As illustrated above, morphemes were selected for the elicitation tasks which had been reported to be vulnerable in German-speaking children with SLI. Table 8 gives an overview over the grammatical structures targeted with the five elicitation tasks.

**Table 8.** Overview of the Morphology Probes Elicited

| *Grammatical function* | *Target structure* | *Procedure* | *No of probes* |
|---|---|---|---|
| *Present tense* | 2nd and 3rd person singular verb inflection | *Picture naming*<br>*Pantomime quiz* | *5* |
| *Case marking* | articles in nominative, accusative and dative | *Picture naming*<br>*Bingo game*<br>*Jigsaw* | *6 per case* |

***Note.*** *Number of probes does not include training items*

All tasks included a training item in order to ensure that the children knew what they were supposed to do. The procedure was audiorecorded using a SONY minidisk recorder MZ-R909 and notes were taken. Two versions of the full scoring sheet are attached in the Appendix 2 and 3: the German original and its translation to English.

### *4.5.1 Scoring*

#### *4.5.1.1 Categorical scoring*

For each grammatical morpheme, percentages of correct use in obligatory contexts were computed. Guidelines for the decision correct / incorrect were developed for each morpheme and applied during scoring. These guidelines are described in detail in the sections on the individual target structures. In general, correct answers were scored as such. Incorrect answers were categorised as one of three categories:

- Grammatical error
- Phonological error, if the individual child's speech pattern could account for the error
- No obligatory context (e.g. zero response, missing production of the word class necessary for the obligatory context (i.e. no verb produced in the elicitation of verb inflections))

Only the sum of correct answers and grammatical errors were considered as the total of obligatory contexts. Errors that could result from the children's speech impairment were excluded from the data analysis. In these cases, it could not be decided whether the child had a grammatical problem that impacted on the target structure or whether only the child's speech difficulties impacted on the correct production. Usually, speech difficulties were considered as relevant if the child showed a minimum of two instances of the phonological process that conflicted with the production of the respective grammatical morpheme. However, if the relevant phonological context occurred only three times or less in non-grammatical context, one occurrence of the phonological error was sufficient to categorise the error in grammatical context as phonological rather than grammatical error. The third category of the list above illustrates that the obligatory context was clearly defined for each morpheme. Thus, for verb inflections the production of a verb was necessary, and for case marking the production of a noun. Zero responses or responses missing these indicators of the obligatory context were considered as ambiguous. It remains arguable whether these responses represent a grammatical deficit because a lack of task comprehension or a pragmatic deficit that could have resulted in such responses, too.

Therefore, the percentage of correct use in obligatory context for each morpheme was calculated as follows:

$$\frac{\text{Correct productions}}{\text{Total of correct productions + grammatical errors}}$$

#### 4.5.1.2 Qualitative analysis

Additionally to a categorical analysis of the responses, error patterns were identified. Research so far indicated that some error patterns may be particularly characteristic for children with SLI only (e.g. Schöler et al., 1998b). Furthermore, in order to compare error patterns across groups, responses were included in the qualitative analysis that had been excluded from quantitative and therefore categorical analysis. Children with phonological impairment, for example, were considered to produce more phonological errors than the other three participant groups. Children with SLI as well as language matched controls were seen as more likely to show zero responses or responses missing the obligatory context for morphemes. A descriptive analysis of error pattern was consequently added. The relevant categories are illustrated in the sections on the individual target structures.

### 4.5.2 Statistical analysis

#### 4.5.2.1 Quantitative analysis

Arcsine transformations were applied to the percentage data[6] and analysed through analysis of variance (ANOVA). The assumption of homogeneity was checked with Levene's test. Accordingly, ANOVA was followed by Gabriel's procedure if variance was equal across groups or by Games-Howell post hoc test if the assumption of homogeneity of variance homogeneity was violated. Both post-hoc tests are suitable for data from groups with different sample sizes (Field, 2009, p.374).

***Children with SLI excluded from quantitative analysis***

The results from four children from the SLI group (SLI2, SLI3, SLI4, SLI9) were excluded from quantitative data analysis either due to their very low levels of expressive language (three children for whom the majority of utterances heard did not exceed one-word) or because they consistently scored zero on all elicitation tasks (the three children previously mentioned and an additional

[6] The distribution of percentages is binomial and arcsine transformation of data was used to achieve a normal distribution.

child). The data from these four weakest children were excluded from quantitative analysis for all elicitation tasks. The results of these children would have distorted the group mean considerably for the SLI group. The children's MLU as well as the children's performance on the elicitation tasks indicated that the children struggled with the general linguistic demands of the task rather than with the morphology per se. As a consequence, the children's zero score on the elicitation task would have represented rather the inability to complete the task appropriately than the inability to apply the elicited morphology. As the aim of the study was the identification of a clinical marker in the children's morphology, though, it would not have been meaningful to apply a zero score to these four children's responses but would have skewed the data. Although, the identification of a clinical marker would possibly have been facilitated by this distortion, such a marker would nevertheless not be representative for a marked grammatical deficit. Thus, the elicited data of the SLI group entering quantitative analysis comprised the responses of 13 children only. Table 54 in Appendix 9 gives an overview over the participant groups with this smaller SLI group.

***Composite scores***

A number of variables were combined to composite scores. Compositional scores have often been used in the literature to characterise the morphology of children with SLI (Bedore & Leonard, 1998; Bortolini et al., 2002; Fletcher & Peters, 1984; Rice & Wexler, 1996). Furthermore, composite scores may be more likely to result in significant group differences than scores for individual morphemes as they allow a larger heterogeneity within the clinical group. One possibility is to use logistic regression in order to identify the most promising combination of variables as clinical marker (e.g. Bortolini et al., 2006). Another possibility is to combine the scores of different variables featuring broader grammatical units such as tense (Rice & Wexler, 1996). In the present study, the data was combined into composite scores following the latter approach, complementary to the analysis of every variable separately. The larger amount of data for each composite variable may have a larger statistical power than the variables each taken on its own.

Two composites were computed: present tense marking and case marking. For this, obligatory contexts and correct productions were summed-up across the morphemes that belonged to each category and percentages correct were calculated. Arc-sine transformations were applied to the percentage data and between-group comparisons carried out using ANOVA.

#### *4.5.2.2 Qualitative analysis*

The qualitative analysis was carried out for each target structure separately. Proportions were calculated of the different response patterns, i.e. the proportion of correct responses and the proportion of all error patterns respectively. For this, the total number of items per target structure was calculated by multiplying the number of participants by the number of items. Next, the number of items was counted per group for each category of response pattern. This number of responses was divided by the total number of items for each target structure in order to calculate the proportion of the respective response pattern. No statistical between group comparisons were performed regarding error pattern because the available quantity of data per response category was too limited.

## 4.6 Language Samples

The second data set focussing on the children's morphology for the current study consisted of conversationally based language samples. Each child was provided with age-appropriate activities for free play. Usually, the examiner served as conversational partner (*n*=60). In three cases, however, the children interacted with their speech and language therapist. A further three children were very shy interacting with the examiner, who was less familiar to them. Therefore, the original language sample did not reflect typical language behaviour and one of the parents was asked to play with their son / daughter, either in the therapy room or at home. Sample activities included play with Lego, Playmobil, dolls house, animals or other toys the child was interested in. All sessions were audiorecorded using a SONY minidisk recorder MZ-R909, or in few cases a SONY audiotape recorder. The approximate length of the language samples was 20 minutes. All sessions were subsequently transcribed by the examiner using the computer software SALT (*Systematic Analysis of Language Transcripts* by Miller & Chapman), research version 7. This software produces quantitative reports regarding the frequency of particular morpheme, word or error codes in the transcript. The sample sizes averaged 202 complete and intelligible utterances per child. The children of both clinical groups produced on average 206 utterances (SLI: SD= 32.68, range 152-257; PI: SD=68, range 94-381), the control groups ND-A and ND-L 199 and 197 utterances respectively (ND-A: SD=53.74, range 141-324; ND-L: SD=37.62, range 131-282). Only one child produced less than 100 complete and intelligible utterances (94).

### *4.6.1 Transcription*

Since no conventions for the transcription of language samples with SALT exist for German, the researcher developed new conventions. The conventions for utterance segmentation follow in general the approach documented by Garman (1989) and Klee (1992), although some were adapted to the individual characteristics of German. The decision rules are summarised in Appendix 5. In order to benefit from the advantages of spontaneous language production, the transcription was not restricted to the morphemes investigated with elicitation tasks. As German has a rather rich morphology, the selection of morphemes represented only a subset of morphemes with potential as a clinical marker and that were feasible to elicit. In order to avoid a bias to this selection, all grammatical morphemes were coded in the transcripts of the spontaneous language samples.

#### *4.6.1.1 Correct and incorrect productions*

Overall, verb inflections in present tense, present perfect, past tense and conditional were coded as bound morphemes. Case and gender marking was highlighted in indefinite articles, pronouns, and strong and weak adjectives. Nouns were marked for case and morphemes indicating plural were coded. Adjectives were further marked if they signalled comparative or superlative. Verbs with particle were transcribed as two morphemes (e.g. *auf machen* (English: *to open / to do open*); *aus ziehen* (English: *to take off*)) although they are orthographically written as one word if they appear as infinitive. In sentences including an object, though, the particle and the verb appear in different positions (e.g. *Ich ziehe meine Schuhe aus.* (English*: I take my shoes off*)*.* Omitted bound morphemes were marked with an asterisk if the grammatical context was obligatory.

In addition to these bound morphemes, case and gender marking in definite articles was indicated by the insertion of word codes. These morphemes are separate grammatical morphemes and would not be picked up by the SALT programme if not coded. Moreover, correct use of pronouns in nominative, dative and accusative was highlighted, as well as the correct use of auxiliaries, modal verbs, and the copula *sein* (= to be). Furthermore, verb placement was marked by codes for utterance initial, second or final position. Besides the codes for correct morphology, error codes were inserted for all grammatical units mentioned above. The first initials of the error code refer to the target morpheme (e.g. EC=error case), the following letters specify the context and error further. As the PLAKSS was not carried out to all participant groups phonological errors were coded, too, and considered for the distinction between phonological and grammatical errors. Appendix 5 provides an overview of all SALT codes used.

#### *4.6.1.2 Qualitative analysis*

In order to identify different error patterns such as omissions and substitutions, error codes were usually completed with a code referring to the error pattern (e.g. EVIV = error verb: infinitive). However, errors could not always unambiguously be categorised. With respect to case marking, for example, interpretations were difficult because some articles are homonyms with different grammatical semantics (e.g. *der* is the correct form for the definite masculine article in nominative as well as the feminine article in dative). As a consequence, the production of an incorrect article could not always unambiguously be identified as gender or case error. Neither was it always possible to decide whether a morpheme had been produced correctly. For example, if the verb with 3$^{rd}$ person singular inflection was followed by a word with initial alveolar plosive it was, due to co-articulation, in most cases impossible to judge whether the verb inflection was actually produced or not (e.g. *Der Mann kauf**t d**as Buch*. – English: *The man buys the book*.). Generally, ambiguous cases were marked as such but did not enter quantitative or qualitative analysis. The complete list of applied transcription conventions can be found in the Appendix 5.

### *4.6.2 Analysis*

A quantitative and qualitative analysis of the data followed the transcription.

#### *4.6.2.1 Quantitative analysis*

Morphemes were selected for statistical analysis if at least eight children per group produced a minimum of two obligatory contexts. The data of any child entered analysis if the child produced a minimum of two obligatory contexts for the respective morpheme. The present study presents the data for present tense inflections and case marking. Percentages of correct use in obligatory context were calculated by dividing the number of correct productions of the morpheme by the number of correct productions plus the number of instances in which the morpheme was replaced by an incorrect inflection or omitted. As with the data from the elicitation tasks, only those errors were counted that could not be explained by the child's individual speech output. Ambiguous cases were excluded from for the calculation of percentages, too.

The four children who were excluded from analysis of the morphology screening were not deliberately excluded from analysis of the language sample. As, naturally, the four children produced obligatory contexts for few morphemes only a small proportion of their data entered analysis and could not be assumed to skew the results. This procedure was chosen based on the assumption that the minimum production of two obligatory contexts indicates the respective morpheme to lie within the scope of the child's language profile. Thus, the data of any child entered analysis only if the child produced two or more obligatory contexts for the respective morpheme.

Arc-sine transformations were applied to the percentage data and between-group comparisons performed using analysis of variance (ANOVA) followed by Gabriel's procedure. If Levene's test indicated that the assumption of homogeneity of variance was violated, the Games-Howell procedure was used as post hoc test instead (Field, 2009).

Next to the analysis of each morpheme that met the criterion for statistical analysis, compositional scores were derived. The data of the language samples provided the possibility for two different aspects reflected in the composites: present tense inflections and case marking. The above applied criteria for statistical analysis were now broadened to the total of inflections falling into a particular category. Thus, it was necessary that a minimum of eight children presented a minimum of two obligatory contexts each for present tense and case respectively. For each composite percentages correct were computed and between-group comparisons were carried out with ANOVA.

#### *4.6.2.2 Qualitative analysis*

The qualitative analysis was carried out for each morpheme in present tense and case marking that met the criterion for statistical analysis. Proportions were calculated of the different response pattern, i.e. the proportion of correct responses and the proportion of all error patterns respectively. For this, the total number of obligatory contexts was calculated by summing-up the number of obligatory contexts within each participant group. Per error category, the number of errors was then counted per group. This number of errors was divided by the total number of obligatory contexts for each target structure in order to calculate the proportion of the respective response pattern. No statistical between group comparisons were performed regarding error patterns because the available quantity of data per response category was too limited.

## 4.7 Further Exploration of the Data

***Individual children with SLI in relation to age-matched and language-matched children***

Besides the between-group comparisons, *t*-values were computed in order to allow the comparison of individual children's performance to the control groups. For this purpose, the children's individual data were referenced to the performance of the typically developing children using *t*-statistics as it is commonly used in single-case studies. Both control groups of typically developing children were used as reference group respectively. The results of the children with SLI were therefore transformed to *t*-values, firstly, based on the mean and standard deviation of the ND-A group, and secondly, based on the mean and standard deviation of the ND-L group. The modified *t*-test procedure by Sokal and Rohlf (as cited in Crawford & Howell, 1998) was preferred over the calculation of *z*-scores due to the small size of the reference sample. For the *t*-values relative to the ND-L group, the number of degrees of freedom (*n*-1) is 15 because the *t*-statistics are based on a control sample with 16 participants. For the *t*-values relative to the ND-A group, the number of degrees of freedom is 16. Crawford and Howell's method was further used to compute the percentile scores resulting from the *t*-values. This gives an estimate of the percentage of the control population that would perform below or above the score of the patient.

***Evaluation of the suitability of a clinical marker***

The existence of significant between-group relations between the clinical group and unaffected cases is a crucial pre-condition for a clinical marker. However, the suitability of a potential clinical marker is usually further assessed by calculating the diagnostic accuracy, the sensitivity and the specificity of such a marker. Additionally, the positive as well as negative likelihood ratio reflect the degree of confidence that an individual person is classified correctly as unaffected (LR-) or affected (LR+) independent of the prevalence of the tested condition (s. chapter 2 for review). Both procedures were employed in this study in order to evaluate the tested variables as potential clinical marker if significant between-group differences suggested this. Overall discrimination accuracy, sensitivity and specificity was calculated for every value of the independent variable, as well as the from these resulting likelihood ratios. The value resulting in the best diagnostic accuracy was chosen as best cut-off and reported along with the corresponding sensitivity and specificity rates as well as likelihood ratios. If the diagnostic accuracy was the same for several values the one was chosen resulting in the higher positive likelihood ratio. The overview over all scores from the SLI and ND-A group and the resulting sensitivity, specificity and likelihood ratios is provided in Appendix 8.

The following formulas were used to evaluate this cut-off:

Overall discrimination accuracy: total of correctly classified children from SLI and ND-A group / total of children in SLI and ND-a group for the respective variable

Sensitivity: true positives / (true positives + false negatives)
Specificity: true negatives / (true negatives + false positives)
LR+: sensitivity / (1- specificity)
LR-: (1 – sensitivity) / specificity

The confidence intervals for 95% confidence were computed using the 'CEBM Statistics Calculator' on the website of the Centre for Evidence-Based Medicine at the KT Clearinghouse in Toronto (2014, September 24) (Centre for Evidence-Based Medicine, 2000). The guidelines presented in chapter 2 were used for assessment of these LR, sensitivity and specificity values. Sensitivity as well as specificity levels between 80% and 90% were consequently considered as acceptable, and levels of 90% or higher as good (e.g. Bortolini et al., 2006; Plante & Vance, 1994). A LR+ value higher than 10 was judged as good (Klee et al., 2007) and as excellent if the value reached 20. The LR- value, on the other hand, was considered to be good if lower than 0.20 and as excellent if it was 0.10 or below.

## 4.8 Reliability

In order to evaluate the online transcription of the elicitation tasks seven children (10.6% of 66 children) were chosen randomly for each task (at least one per group) and transcribed again by the examiner using the audiorecordings. Point-to-point agreement ranged from 84% to 100% (Mean = 90 %). Kappa lies at 0.781 for the intrajudge agreement on the morphology screening with the 95% confidence interval between 0.695 and 0.868. The agreement can therefore be considered as substantial to almost perfect (Landis & Koch, 1977).

Furthermore, seven language samples were randomly selected (at least one per group) and independently transcribed by a second judge, a qualified and experienced German-speaking speech and language therapist. Word-by-word agreement between the two judges ranged from 84.3% to 99.4% (Mean: 94.7%); morpheme-by-morpheme agreement ranged from 73.6% to 97.2% (Mean: 92.1%). To evaluate interjudge reliability for coding of the language sample for the children’s use of grammatical morphemes in obligatory context, seven uncoded transcripts were randomly selected, although again at least one per group. These uncoded transcripts were given to an independent coder who noted omissions, correct and incorrect use of the grammatical structures of interest in obligatory context. Agreement between the first and second coder ranged from 89.7 to 99.8 (Mean: 96.5%).

## 4.9 Ethical Considerations

Before the beginning of data collection the parents were informed about the purpose of the study and all gave their written permission for testing. These information letters including the permission form are given in Appendix 6 and 7. The data of each child is used and stored anonymously.

An application for ethical approval was sent to the 'Ethikkommission der Ärztekammer Hamburg', the ethics committee responsible for clinical studies carried out in the federal state Hamburg. However, according to the ethics committee no such approval was required in Germany for the planned study.

All tasks were presented in form of a game, were non invasive and dependent on the cooperation of the child.

# 5 Verb Morphology in German SLI

## 5.1 Background

The investigation of verb inflections in present tense has shown promising results as potential clinical marker for SLI across several languages (e.g. Bedore & Leonard, 1998; Rice & Wexler, 1996). Also German-speaking children have been reported to show limitations in their verb morphology (e.g. Clahsen, 1989, 1991; Clahsen et al., 1997; Rice et al., 1997; Roberts & Leonard, 1997; Schöler et al., 1998). Rice, Noll and Grimm (1997) showed that 4-year-old, German-speaking children with SLI lagged behind younger MLU-matched controls in verb finiteness in present tense. This indicates that these deficits may be larger than would be predicted by the child's general language status. Such a pattern of a "delay within a delay", first identified by Rice (2003, p. 65), could make this error with present tense inflections an appropriate candidate for a clinical marker for SLI in German.

Difficulties were also reported in German present perfect (Clahsen, 1989; Clahsen & Rothweiler, 1992). Present perfect is commonly used in spoken German to refer to past events. Two studies investigated the production of present perfect. Clahsen (1989) reported the omission of auxiliaries as a frequent error pattern in SLI whereas the production of the past participle seemed to be unimpaired in his sample. Longitudinal data from a larger sample (n=19) aged between 3;1 and 7;11 confirmed generally high accuracy rates for present perfect inflections (mean >90%), although these were comparable to those of language-matched controls for both the past participle prefix and the past participle suffix (Clahsen & Rothweiler, 1992). Consequently, children with SLI may lag behind age-expectations but there is no evidence for a marked problem with the past participle standing out from the general language profile. Also error patterns reflected those of the younger, typically developing children. This indicates that present perfect inflections have little potential for the clinical identification of German-speaking with SLI. The current sub-study will therefore focus on present tense in German-speaking with SLI.

## 5.2 Present Tense in German

Verbs in German are inflected for person, number and tense. Each verb has an infinitive form which can be split into the verb stem and the suffix *-en*. As Figure 2 illustrates, morphemes that carry information for person, number and tense are suffixes that replace the infinitive ending. The example of *spielen* (Engl.: *to play*) below in Table 9 provides the inflections in present tense for a 'weak' verb. These verbs require only the suffixation of the inflection, whereas 'strong' verbs are conjugated irregularly and involve additionally a vowel change in the verb stem. German does not have present progressive. Thus, present tense is referring to present events, a state of being or an occurrence in the future.

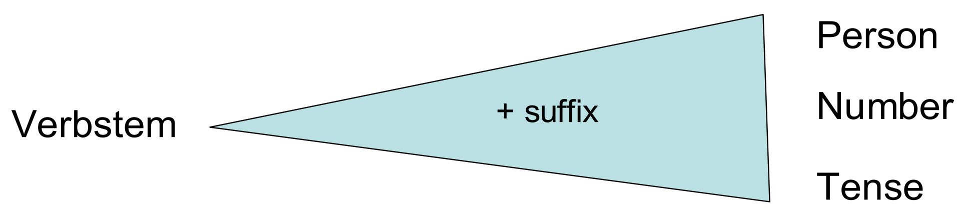



**Figure 2.** German verb phrase

**Table 9. Present Tense Inflections in German**

| Person | Singular | Plural | Infinitive |
|---|---|---|---|
| 1st | ich spiel**e** | wir spiel**en** | spiel**en** |
| 2nd | du spiel**st** | ihr spiel**t** | (to play) |
| 3rd | er / sie / es spiel**t** | sie spiel**en** | |

The conjugation of the copula *sein* (English: *to be*) is given in Table 10.
All forms are monosyllabic.

**Table 10.** Conjugation *sein*

| Person | Singular | Plural | Infinitive |
|---|---|---|---|
| **1st** | ich **bin** | wir **sind** | sein |
| **2nd** | du **bist** | ihr **seid** | (to be) |
| **3rd** | er / sie / es **ist** | sie **sind** | |

## 5.3 Present Tense in German SLI

Controversy remains as to the types of errors that German-speaking children with SLI make in marking the present tense. The predominant error types are thought to be verb infinitive or the production of the verb stem only (Clahsen, 1989, 1991; Clahsen et al., 1997; Kany & Schöler, 1998; Rice et al., 1997; Roberts & Leonard, 1997). This is of particular interest, because the production of the infinitive does not correspond to the omission of the verb inflection as in English children with SLI. Instead, the children 'add' the suffix *-en* to the verb stem. However, reports vary considerably with respect to the proportion of bare stems in the children's language. Whereas Roberts and Leonard (1997), for example, reported bare stems to be the most frequent error type, Rice et al. (1997) found such errors to account only for 6% of the errors. The children with SLI in this latter study produced significantly more verbs as infinitives than their language-matched controls. It is noteworthy that Rice and colleagues excluded those omissions of verb inflections from analysis that were likely to result from phonological deficits evident in the children's speech. This suggests that a large proportion of the bare stems reported by Roberts and Leonard can be accounted for by speech rather than grammatical problems of the children. Age could have influenced the differences between the two studies further: although both samples have comparable MLU, the children in the study by Rice et al. were approximately two years younger (aged 4;0-4;8 years) than those studied by Roberts and Leonard (aged 3;5-7;2 years). Consequently, the use of infinitives could be an error pattern especially characteristic for children with SLI below the age of five.

Where children with SLI do produce finite verbs in the present tense, again there is debate as to the pattern of errors in this production. For example, Roberts and Leonard (1997) reported that where children with SLI did produce finite verbs spontaneously then they were usually correct regarding subject-verb agreement. This is supported by spontaneous data presented by Rice et al. (1997). If the children produced verb inflections corresponding to 3$^{rd}$ or 2$^{nd}$ person singular in present tense, they appeared most of the time in the correct context. Clahsen et al. (1997), on the other hand, reported from a sample of six children (aged 5;8 – 7;11 years) that one third of verb errors in spontaneous language represented finite verbs with an incorrect inflection. Spontaneous data of ten children (aged 3;8-9;6 at first recording) with SLI presented earlier (Clahsen, 1989) suggested that 3rd person singular inflection (-*t*) was the preferred substitute, the data from 1997 presented 1st person singular (-*e*) and 3rd person singular as the most frequent substitutes.

The production of the copula *sein* (Engl.: *to be*) was investigated by Roberts & Leonard (Roberts & Leonard, 1997). The German copula shares the same surface features with the forms of the English copula *to be:* all forms are monosyllabic and, in connected speech, they appear usually unstressed. Roberts and Leonard demonstrated that the copula is indeed prone to omissions in children with SLI. However, the error rate of 25% in German-speaking children was much lower than that reported in English-speaking children (64%). This corresponds to an omission rate for *sein* reported by Rice et al. (1997). The initial omission rate of 29% at age 4;0 to 4;8 improved within one year to 9%. The production of the copula has repeatedly been used as part of a tense composite as clinical marker for English SLI (Bedore & Leonard, 1998; Rice & Wexler, 1996). Given the results for German-speaking children with SLI, it is unlikely that the production of the copula *sein* can be used alone to discriminate between typical development and SLI. However, as for English, this variable may bear some potential for a composite.

Overall, a look at the age ranges in the literature on German SLI indicates that these are rather large but that researchers treated the group of children with SLI nevertheless as one group. The impact of large age ranges had probably a considerable impact on the results as the sample sizes varied between 3 and 90 participants with sample sizes larger than 15 being an exception.

The investigation therefore addresses the following experimental questions:

1. Do German speaking children with SLI exhibit the pattern of a delay within a delay in their use of present tense verb inflections when compared to typically developing children and those with a phonological impairment (PI)?

2. What types of errors do children with SLI make and are these qualitatively different from those made by typically developing children and those with PI?

In order to answer these questions, two approaches were used for data collection: elicitation procedures and the recording of spontaneous language samples. Both procedures of data collection, including data analysis and results are presented below.

## 5.4 Morphology Probes on Present Tense

### *5.4.1 Participants*

Sixty-two monolingual German-speaking children participated in the experiment on case marking. Thirteen children met the criterion SLI (SLI group), 16 children were age and gender matched and had typically developing language but a phonological speech output impairment (PI group), 17 children were typically language developing and age and gender matched with the SLI group (ND-A group), and a further 16 children were typically language developing and matched with the SLI group on gender and language comprehension (ND-L). All groups are further described in the general method chapter 4.

### *5.4.2 Elicitation tasks*

Two present tense verb inflections were chosen for the elicitation task: 2nd and 3rd person singular. Third person singular is one of the inflections acquired early (age 1;6 – 1;11), whereas 2nd person singular is the latest acquired present tense verb inflection (2;6-2;11) (Schrey-Dern, 2006). Thus, the elicitation task involved two inflections of different age of acquisition.

First and 3rd person plural inflection were excluded in order to avoid ambiguity. Both inflections share the same surface pattern (*-en*) with the infinitive which is the predominant developmental error pattern (e.g. Clahsen, 1989, 1991). The second person plural inflection shares surface pattern with the 3rd person singular (*-t*) and was therefore also not considered for elicitation. The inflection for 1$^{st}$ person singular would have been interesting to investigate. However, its elicitation would have been pragmatically rather challenging and highly depending on a good language comprehension of the participating children. The obligatory contexts for 2$^{nd}$ and 3$^{rd}$ person singular, on the other hand, were easy to create and could be presented to the child in an interesting way. Furthermore, none of the excluded inflections have previously been reported to be particularly vulnerable in children with SLI.

### *5.4.2.1 Third person singular in present tense*

Third person singular verb inflection in present tense is signalled by a word final *-t* following the word stem (e.g. er komm**t** – he comes). The inflection was elicited through a simple picture describing task as it is used in a number of standardised language tests, such as the SETK 3-5 (Grimm, 2001) or the *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen (Christina Kauschke & Siegmüller, 2009)*.

Six picture cards, as in Picture 1, were presented to the participants showing drawings of children doing something, e.g. drinking, playing football, or feeding a cat. All verbs were regular, weak German verbs, representative for the vocabulary of young children[7] and easy to illustrate in a drawing: *trinken, schaukeln, spielen, füttern, kaufen, bauen*. The items included intransitive as well transitive verbs. The direct objects following the verb in the target sentence did not start with a stop consonant. A stop consonant would have resulted in co-articulation and the distinction between the inflectional morpheme and the initial consonant made impossible.

The examiner tried to elicit the third person singular verb inflection by asking:
“What is the girl / boy doing?  The girl / boy …”

<u>Exp.:</u>
*Examiner:* Was macht denn das Mädchen da? Das Mädchen …?
*Child:* … schaukel**t**. (Engl.: is swinging)
One sentence was used as practice item, five sentences were scored.

[7] The verbs reflected actions relevant for the daily routine of young children: general actions (*spielen* (to play)), object related actions (*kaufen* (to buy), *trinken* (to drink), *füttern* (to feed), *bauen* (to build) and actions involving movement (*schaukeln* (to swing). Four of the six verbs were also used in the vocabulary checklist for 2-year-olds ELFRA 2 (Grimm & Doil, 2000).

**Picture 1.** Example of picture for 3rd person singular task.

### *5.4.2.2 Second person singular in present tense*

The second person singular verb inflection in present tense is marked by word final -*st* added to the word stem (e.g. du komm**st** – you come). The inflection was elicited through a pantomime task in which the child should guess an activity that the examiner demonstrated. The author had successfully used this approach with children with SLI in therapeutic contexts and so knew that the task could offer a pragmatically appropriate stimulus for the second person singular inflection, whilst also being comprehensible and entertaining for children. All verbs were regular, weak German verbs, representative for the vocabulary of young children and easy to act out: *trinken, schreiben, weinen, schwimmen, lachen, telefonieren*. Since all verbs were intransitive, the phonological context enabled an unambigious decision whether the child produced the verb inflection correctly or not.

In order to elicit the target item, the examiner pantomimed each verb and asked the child to guess "What am I doing?".

If the child produced the infinitive instead of the 2nd person singular, the examiner tried to prompt the inflected form by pointing towards herself saying: "Yes … (infinitive verb). And now you can say to me *You*… / How does it go with *you*?".

If the child still did not produce the inflected form the examiner gave the example of the preceding sentence (or test item). E.g.: “You said *trinken* (Engl.: to drink) with *du* (Engl.: you) is *du trinkst* (Engl.: you are drinking). How is this called with *schreiben* (Engl.: to write)? *Du* …?” In the case that this prompt still did not help the child to produce the verb inflection, the infinitive was counted as the child’s response and thus as an error. Again, one sentence served as trial, five sentences were scored.

### *5.4.3 Documentation*

The procedure was audiorecorded using a SONY minidisk recorder MZ-R909 and online notes were taken. If the children produced the target structure the verb inflection was marked as correct; if the child gave a different answer to the target this was transcribed on-line. Section 4.3 gives intra- and interjudge agreement values for 10,6% of the transcripts.

### *5.4.4 Quantitive analysis*

#### *5.4.4.1 Scoring*

For both inflections, percentages of correct use were computed by dividing the number of correct productions by the number of obligatory contexts for each inflection. The following outlines for both inflections separately the approach how to decide whether a response was categorised as correct or incorrect:

<u>3rd person singular</u>

- Zero responses and responses which did not include a verb were removed from the number of obligatory contexts. It was in these cases ambiguous whether the child did not intend to produce the target item (possibly due to task comprehension or cooperation) or whether the child was not able to respond due to his linguistic limitations.

- Errors were categorised as grammatical or phonological error. This categorisation was based on the individual speech output of each participant and was relevant especially for two error patterns:
    - Deletion of word final consonants
    - Reduction of consonant clusters (relevant if the preceding verb stem ended in a consonant)

  Errors were categorised as phonological error if the child showed these processes involving word-final /t/ in at least two phonological contexts other than 3rd person singular tense marking, too.
- Phonological errors were excluded from analysis and the number of obligatory contexts.
- Percentage of correct use was calculated by dividing the number of correct productions by the number of obligatory contexts

2nd person singular

Considering research on the production of verb inflections in German-speaking children with SLI, detailed information is missing when verb inflections have been considered as correct. This is of special interest for the 2nd person singular inflection in present tense. While the phonemically correct form is the suffix –*st,* the realisation of –*s* only is already enough to indicate the grammatical contrast. Both conditions would be a possible criterion for the analysis of the correct production. One of the reasons why researchers did not mention this dilemma may be the fact that it is difficult to tease both conditions apart in data from spontaneous language. It seems useful to follow both approaches: for comparability to previous research, the phonemically correct production of the morpheme –*st* should be analysed. The realisation of the grammatical contrast, on the other hand, is most important in the conveyance of the meaning and remains noticeable in most phonological contexts. Moreover, the second approach seems especially relevant in the research of child language. The reduction of consonant clusters is a developmental phonological process present in 10-20% of German-speaking children up to the age of 4;4 (Fox, 2004). This process could result in the deletion of the final –t of the inflection –*st* whereas the grammatical contrast (-*s*) is maintained. Consequently, two variables were analysed regarding 2nd person singular: First, the phonetically correct production of the morpheme *–st* and second, the marking of the grammatical contrast through the suffix *–s.*

*Scoring 2nd person singular: -st realisation*

- Zero responses and responses which did not include a verb were removed from the number of obligatory contexts. It was in these cases ambiguous whether the child did not intend to produce the target item (possibly due to task comprehension or cooperation) or whether the child was not able to respond due to his linguistic limitations.
- Errors were categorised as grammatical or phonological error. This categorisation was based on the individual speech output of each participant and was relevant especially for two error patterns:
    - Deletion of word final consonants
    - Reduction of consonant clusters.

  Errors were categorised as phonological error if the child showed these processes involving word final /t/ or /st/ in phonological contexts other than 2$^{nd}$ person singular tense marking, too.
- Phonological errors were excluded from analysis and the number of obligatory contexts.
- Percentage of correct use was calculated by dividing the number of correct productions by the number of obligatory contexts

*Scoring 2nd person singular: grammatical contrast marked (-s or -st realisation)*

- Zero responses and responses that did not include a verb were removed from the number of obligatory contexts. It was in these cases ambiguous whether the child did not intend to produce the target item (possibly due to task comprehension or cooperation) or whether the child was not able to respond due to his linguistic limitations.
- Errors were categorised as grammatical or phonological error. This categorisation was based on the individual speech output of each participant and was relevant especially for three error patterns:
    - Deletion of word final consonants
    - Stopping of *–s*

  - Reduction of consonant clusters[8]

  Errors were categorised as phonological error if the child showed these processes involving the phoneme /s/ in at least two phonological contexts other than 2nd person singular tense marking, too.

- Phonological errors were excluded from analysis and the number of obligatory contexts.
- Percentage of correct use was calculated by dividing the number of correct productions by the number of obligatory contexts

Therefore, considering the children's speech output, the percentage of production accuracy was calculated follows:

$$\frac{\text{Correct productions}}{\text{Total of correct productions + grammatical errors}}$$

### *5.4.4.2 Statistical analysis*

**Dependant variables**

The four groups of children were compared in terms of their accuracy producing 2nd and 3rd person singular verb inflection in present tense. Five variables were compared: (a) the production accuracy of 3rd person singular *–t,* (b) the production accuracy of 2nd person singular *–st* and (c) the production accuracy of the grammatical contrast *–s* for 2nd person singular (both suffixes, *–st* and *–s*, considered as allomorphs), (d) a present tense composite combining the production accuracy for 3rd person singular and 2nd person singular *–st,* and (e) a present tense composite combining the production accuracy for 3rd person singular and the grammatical contrast *–s* for 2nd person singular. The two different composites were computed in order to compare the impact of different scoring criteria. Following the scoring guidelines above, Table 11 illustrates the number of items that were included and excluded in quantitative analysis for between group comparisons.

[8] relevant for three out of four items as the verb stem ended in a consonant

**Table 11.** Number of Items for Elicited Present Tense Inflections Entering Qualitative Analysis

| | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|
| Total no participants | 13 | 16 | 17 | 16 |
| Total no of items per group and per verb inflection | 65 | 80 | 85 | 80 |
| No of items excluded due to task refusal | | | | |
| 3rd person singular | 2 | 4 | 2 | - |
| 2nd person singular | 17 | 8 | - | 13 |
| No of items excluded due no verb context | | | | |
| 3rd person singular | 2 | - | - | 3 |
| 2nd person singular | 3 | 2 | 1 | 3 |
| No of items excluded due to phonology | | | | |
| 3rd person singular | - | 7 | - | - |
| 2nd person singular: –*st* | 2 | 16 | - | 5 |
| 2nd person singular: grammatical contrast –*s* | 1 | 10 | - | - |
| No of items for analysis (a) | 61 | 69 | 83 | 77 |
| No of items for analysis (b) | 43 | 54 | 84 | 59 |
| No of items for analysis (c) | 44 | 60 | 84 | 64 |
| No of items for analysis (d) | 104 | 123 | 167 | 136 |
| No of items for analysis (e) | 105 | 129 | 167 | 141 |

*Note.* No of items for analysis (d) is (a+b), no of items for analysis (e) is (a+c)

**Data preparation and analysis**

Arc-sine transformations were applied to the percentage data and analysed through analysis of variance (ANOVA). The assumption of homogeneity was checked with Levene's test. Accordingly, ANOVA was followed by Gabriel's procedure if variance was equal across groups or by Games-Howell post hoc test if the assumption of homogeneity of variance homogeneity was violated. If significant between group differences were found between the SLI group and the ND-A or ND-L group the variable was further explored as clinical marker by calculating sensitivity, specificity, LR+ and LR-.

The results of the children with SLI were transformed to *t*-values and percentiles relative to the children's age (i.e. ND-A data) and, secondly, relative to the children's language development (i.e. ND-L data).

### *5.4.5 Results of quantitative analysis*

The four groups of children were compared in terms of their accuracy producing 2nd and 3rd person singular verb inflection in present tense. The descriptives of the statistical analysis are given below in Table 12, providing the different group means, standard deviations and performance ranges for the correct production of the verb inflections.

All scores reflect the proportional accuracy in obligatory context; however, the scores were transformed to arcsine values for further statistical analysis. Group sizes differ because all data that may have been affected by the children's individual phonological processes was excluded. All relevant phonological processes were observed: final consonant deletion, reduction of consonant clusters and stopping of [s]. Furthermore, not all children produced obligatory contexts for each item. In present tense for 3rd person singular, three children (1x SLI, 2x ND-L) failed to produce a verb for 1-2 items. Eight children (2 x SLI, 4 x PI, 2x ND-A) did not respond to one of the items respectively, although all of them produced a minimum of one correct answer within the same task. In present tense for 2nd person singular, nine children (3x SLI, 2x PI, 1 x ND-A, 3x ND-L) did not produce a verb for one item respectively. Twelve children (6 x SLI, 3 x PI, 3x ND-L) did not respond to all of the 2nd person singular items, four of which failed to respond to any of the items (1x SLI, 1 x PI, x ND-L). The resulting number of subjects for which the data was analysed is presented in Table 12.

Separate ANOVAs were calculated for all three conditions and revealed main group effects for each of them. As the homogeneity of variance assumption was violated for the grammatical contrast *-s* (i.e. *-s* or *-st*), the Brown-Forsythe *F*-ratio is given for this variable. Furthermore, Games-Howell procedure was used as post hoc test for this variable whereas analysis of both remaining variables was followed by Gabriel's procedure as post hoc test.

**Table 12.** Elicitation Tasks: Proportional Scores of Correct Use of Present Tense Verb Inflections

| ***Inflection*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **3rd singular** | **M** | $0.62_a$ | 0.75 | $0.93_a$ | 0.76 |
| | **SD** | 0.34 | 0.30 | 0.13 | 0.31 |
| | **Range** | 0.0 – 1 | 0.0 – 1 | 0.6 – 1 | 0.0 – 1 |
| | **n** | 13 | 16 | 17 | 16 |
| **2nd singular** | **M** | 0.42 | 0.48 | 0.77 | 0.76 |
| | **SD** | 0.37 | 0.45 | 0.28 | 0.32 |
| | **Range** | 0.0 – 1 | 0.0 – 1 | 0.2 – 1 | 0.0 – 1 |
| | **n** | 11 | 11 | 15 | 13 |
| **2nd singular** | **M** | 0.66 | 0.78 | 0.99 | 0.99 |
| **-s / -st** | **SD** | 0.45 | 0.37 | 0.05 | 0.05 |
| | **Range** | 0.0 – 1 | 0.0 – 1 | 0.8 – 1 | 0.8 – 1 |
| | **n** | 11 | 15 | 17 | 14 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

The accuracy of 3rd person singular inflection differed significantly across the groups, $F$ (3, 58) = 3.34, $p$ = .025, $\omega^2$ = .10. Gabriel's procedure revealed that the children with SLI produced significantly less correct 3rd person singular inflections than the ND-A group, $p$ = .019. No further group differences reached the level of significance.

If used as cut-off in the production of elicited 3rd person singular in present tense, the production accuracy of 75% resulted in the best diagnostic accuracy as shown in Table 41 in Appendix 8. Applying this cut-off to the data, 77% of the children were correctly classified as SLI or ND-A. The corresponding sensitivity is 0.615 [0.355 - 0.823], the specificity 0.882 [0.657 - 0.967], the LR+ = 5.231 [1.328 - 20.604] and the LR- = 0.436 [0.214 -0.886].

For 2nd person singular, ANOVA showed a main group effect ($F$ (3, 46) = 3.06, $p$ = .037, $\omega^2$ = .11) but Gabriel's procedure did not identify significant differences between individual groups. The main effect remained when the grammatical contrast -*s* was analysed as correct, Brown-Forsythe $F$ (3, 23.58) = 3.93, $p$ = .021, $\omega^2$ = .16. Although in this condition, too, the post hoc Games-Howell procedure failed to identify significant between-group differences for individual groups. Figure 3 illustrates the findings providing the different group means of the proportional scores before arcsine transformations.

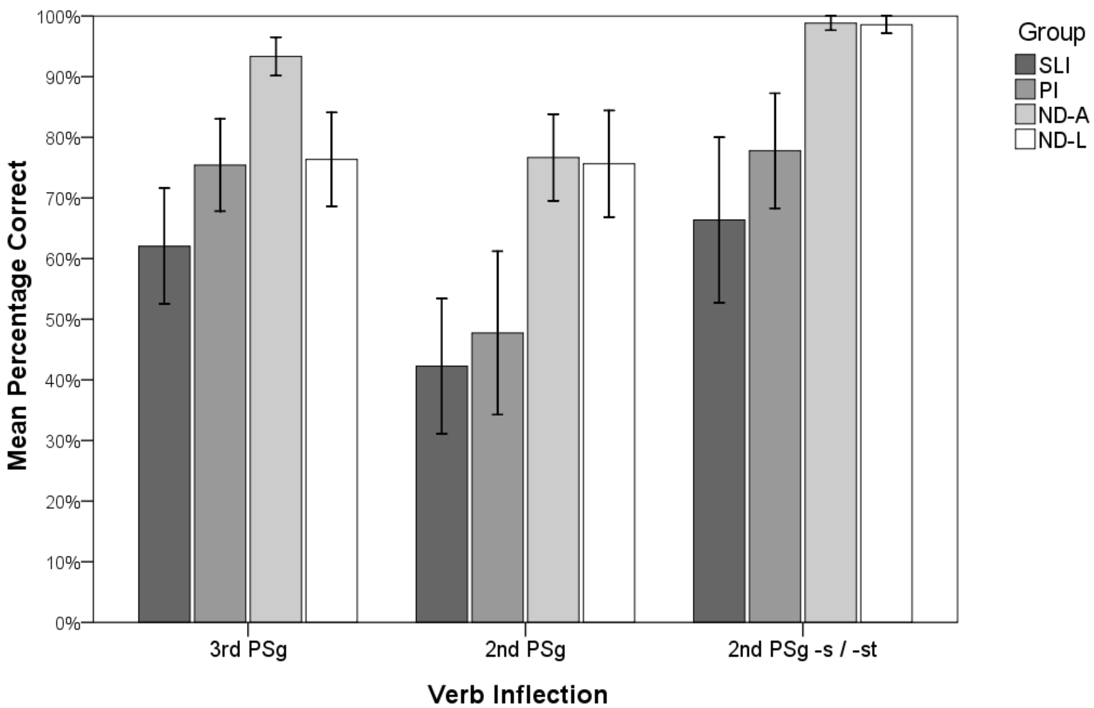



**Figure 3.** Mean percentages of use (± 1 SE) for the four groups of children for elicited third and second person singular verb inflection in present tense and for marking the grammatical contrast in second person singular (phonetically correct or incorrect).

**Present tense composite**

Two composite scores for present tense marking were derived from the elicitation tasks by combining the data of the studied inflections 2nd person singular and 3rd person singular in present tense. For the first composite, all correct 3rd person singular inflections and all correct 2nd person singular inflections were taken together and related to the total number of items requiring these inflections. For a second present tense composite, the percentage correct was calculated referring to 3rd person singular inflections and the grammatical contrast for 2nd person singular. Summary group data are presented below in Table 13.

**Table 13.** Elicitation Tasks: Compositional Scores Present Tense

| *Compositional Score* | | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|---|
| **Present tense** | **M** | $0.58_a$ | 0.62 | $0.85_a$ | 0.75 |
| | **SD** | 0.30 | 0.31 | 0.15 | 0.32 |
| | **Range** | 0 – 1 | 0 – 1 | 0.6 – 1 | 0 – 1 |
| | **n** | 13 | 16 | 17 | 16 |
| **Present tense *-s*** | **M** | $0.67_a$ | $0.72_b$ | $0.98_{a,b}$ | 0.81 |
| | **SD** | 0.31 | 0.29 | 0.05 | 0.26 |
| | **Range** | 0 – 1 | 0.2 – 1 | 0.88 – 1 | 0 – 1 |
| | **n** | 13 | 16 | 17 | 16 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants
present tense *-s* = 3rd person sing. + 2nd person sing. including allomorph *-s*

For further statistical analysis, all proportional scores were transformed to arcsine values. One-way ANOVA revealed for both present tense composites significant main group effects which were followed-up by post hoc procedures:

present tense *F* (3, 1.85) = 3.11, *p* = .033, $\omega^2$ = .09; present tense –*s* *F* (3, 46.60) = 7.00, *p* = .001, $\omega^2$ = .23. These post hoc tests demonstrated that the stricter criterion for 2nd person singular resulted in significant differences only regarding the SLI and the ND-A group (present tense SLI < ND-A (*p* = .05)) whereas the inclusion of the allomorph -*s* as correct marking of the grammatical contrast resulted in significant differences between both clinical groups respectively and the ND-A group (present tense –*s* SLI < ND-A (*p* = .001) and PI < ND-A (*p* = .007)). Figure 4 illustrates both proportional present tense composite scores for the four groups.

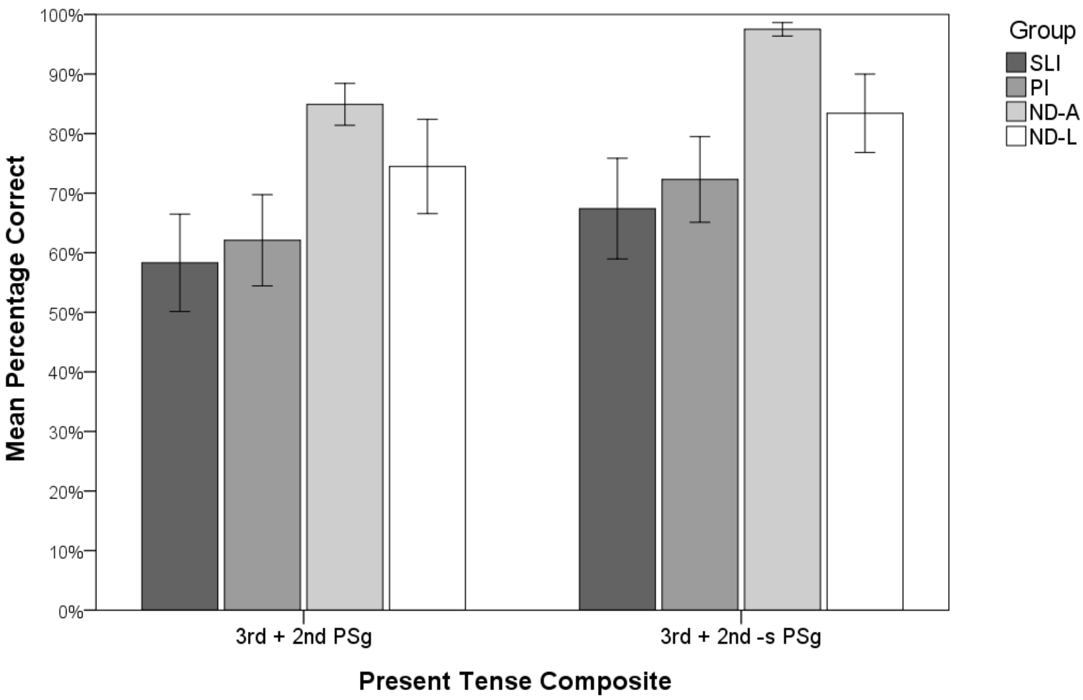



**Figure 4.** Mean percentages of use (± 1 SE) for present tense composites from elicited production.

If using the present tense composite as diagnostic tool, 56% was the most suitable cut-off as the overview over the data in Table 42 in Appendix 8 demonstrates. This cut-off resulted in a in a diagnostic accuracy of 77%, a sensitivity of 0.46, a specificity of 1, and LR- = 0.54. Due to the absence of false positives, it was not possible to calculate neither LR+ nor the confidence intervals. If the allomorph –*s* was considered as correct marking of 2nd person singular, the data suggest a cut-off of 88% as shown in Table 43. This resulted in an overall discrimination accuracy of 90%, a sensitivity of 0.846 [0.578 - 0.957], a specificity of 0.941 [0.73 - 0.99], LR+ = 14.385 [2.118 - 97.678] and LR- = 0.163 [0.045 - 0.588].

### *5.4.6 Qualitative analysis*

**Data coding**

In order to reveal trends in error patterns as reported in previous research (Clahsen, 1989, 1991; Clahsen et al., 1997; Kany & Schöler, 1998; Rice et al., 1997; Roberts & Leonard, 1997), a descriptive analysis was carried out of the response patterns for present tense inflections. For this, responses were coded into eight categories of production patterns for 3$^{rd}$ person singular inflections and nine categories of production patterns for 2$^{nd}$ person singular inflections:

- Correct production
- Production allomorph –*s* (2$^{nd}$ person singular only)
- Verb infinitive
- Verb stem (omission of inflection)
- Phonological error[9]
- Incorrect inflection
- Vowel change in verb stem
- No verb produced
- No attempt

**Data included**

For qualitative analysis, the data of all 62 children was included, including zero responses (category *no attempt*). This category was added to the qualitative analysis in order to account for the fact that children with very poor language skills may not have been able to attempt the task. Also the categories

[9] This error pattern appeared at the surface as production of the verb stem only or as the production of the allomorph –*s*. However, if the respective child showed a phonological process that accounted for this production pattern the error was counted as phonological error rather than as an error from another category.

*phonological error* and *no verb produced* comprised data that had been excluded from quantitative analysis. This approach could give an indication in how far the inclusion or exclusion of this data impacted on the outcome of the between group comparison.

**Data analysis**

The total number of items per group and per present tense verb inflection is given in Table 14 together with the proportion of each error type.

### *5.4.7 Results of qualitative analysis*

**Error patterns in present tense inflections**

Figure 5 illustrates the distribution of response types for each group for 3rd person singular present tense in the elicitation task. Similarly to the analysis with one-way, the four worst cases of the SLI group were excluded.

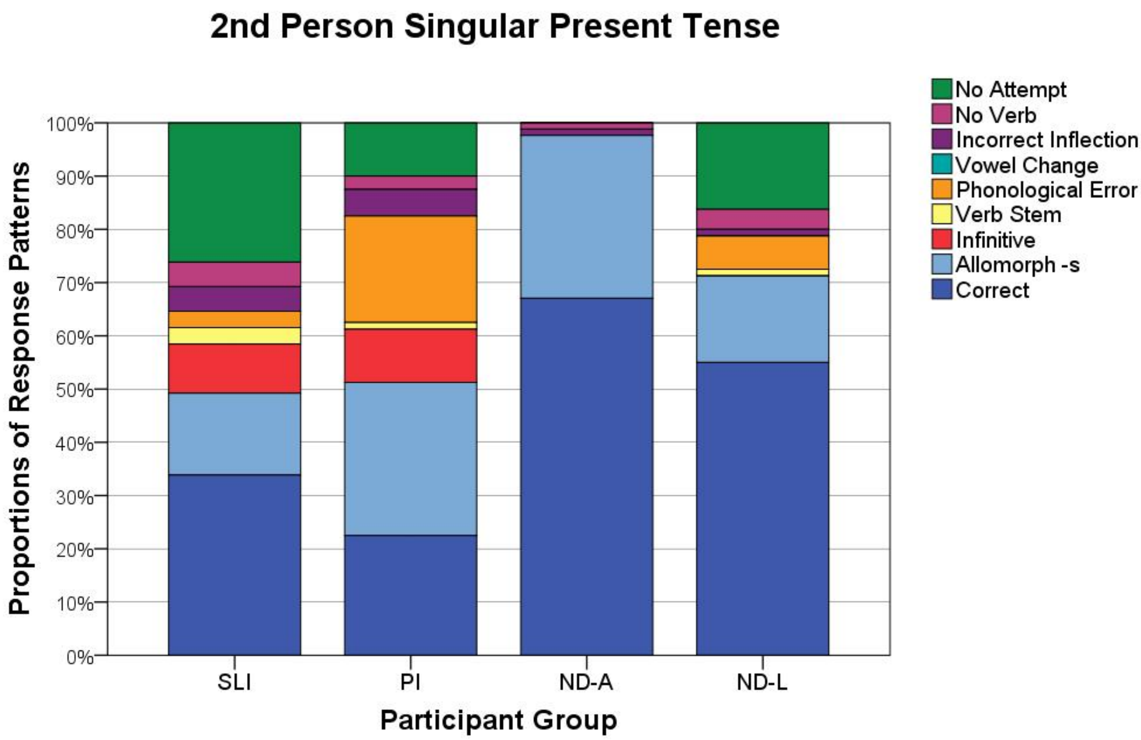



**Figure 5.** Production patterns for elicited 3rd person singular inflection in present tense.

Figure 6 gives the distribution of the response types for each group for 2nd person singular present tense in the elicitation task.

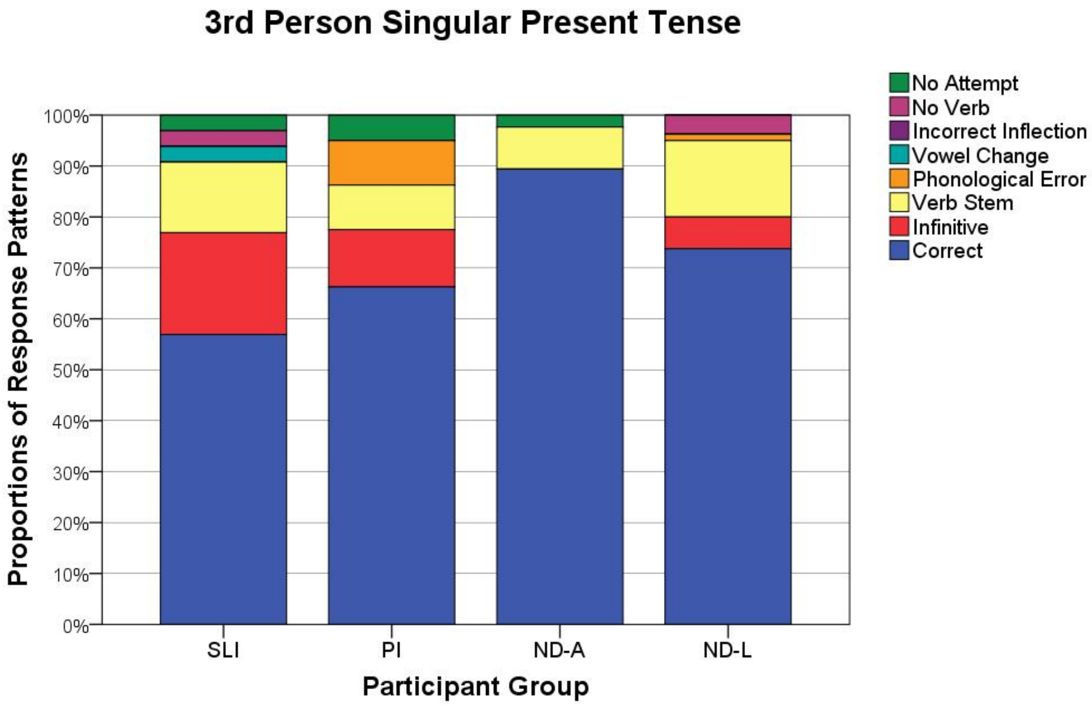



**Figure 6.** Production patterns for elicited 2nd person singular inflection in present tense.

Rounded proportions in relation to the total of productions are presented in Table 14 for both 3rd and 2nd person singular. Between group comparisons were not performed for the different production patterns. The available data per category was too limited.

**Table 14**. Elicitation Tasks: Production Patterns for Present Tense Inflections

| ***Production Pattern*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **3rd singular** | **Correct** | 57% | 66% | 89% | 73% |
| | **Infinitive** | 20% | 11% | 0% | 6% |
| | **Verb stem** | 14% | 9% | 8% | 15% |
| | **Phonological error** | 0% | 9% | 0% | 15% |
| | **Incorrect inflection** | 0% | 0% | 0% | 0% |
| | **Vowel change** | 3% | 0% | 0% | 0% |
| | **No verb** | 3% | 0% | 0% | 4% |
| | **No attempt** | 3% | 5% | 2% | 0% |
| **2nd singular** | **Correct** | 34% | 23% | 67% | 55% |
| | **Allomorph *-s*** | 15% | 29% | 31% | 16% |
| | **Infinitive** | 9% | 10% | 0% | 0% |
| | **Verb stem** | 3% | 1% | 0% | 1% |
| | **Phonological error** | 3% | 20% | 0% | 6% |
| | **Incorrect inflection** | 5% | 5% | 1% | 1% |
| | **Vowel change** | 0% | 0% | 0% | 0% |
| | **No verb** | 5% | 3% | 1% | 4% |
| | **No attempt** | 26% | 10% | 0% | 16% |

Note. Percentages are rounded.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

### *5.4.8 Interpretation*

Ideally, a variable qualifying as suitable clinical marker for SLI should stand out from the children's general language impairment (Rice, 2003). In the context of the present study, this would be represented by a significant poorer performance of the children with SLI not only relative to the ND-A group but also relative to the ND-L group. Such pattern was not apparent in the results for elicited present tense marking. There is consequently no straightforward candidate in elicited present tense marking as clinical marker for German SLI. However, children with SLI showed a delay regarding the production of 3$^{rd}$ person singular inflections, and both present tense composites as they lagged significantly behind the ND-A group on these three variables. These findings, as well as the results regarding the PI group, will be further discussed below, and linked to the data from spontaneous language samples.

## 5.5 Spontaneous Data

### 5.5.1 Participants

Sixty-six monolingual German-speaking children participated in the experiment on present tense. Seventeen children with a mean age of 3;11 met the criterion SLI (SLI group), 16 children were age and gender matched and had a phonological speech output impairment (PI group), 17 children were typically language developing and age and gender matched with the SLI group (ND-A group), and a further 16 children were typically language developing and matched with the SLI group on gender and language comprehension (ND-L). All groups are further described in the general method chapter 4. The numbers of participants in the SLI group differ from the elicited data analysis because no children were excluded as outliers prior to the analysis.

### 5.5.2 Procedure

Spontaneous language samples were recorded during free play with the examiner, a parent or the speech and language therapist. Average length of the transcripts was 204 complete and intelligible utterances. The samples were recorded using a SONY minidisk recorder MZ-R909 or a SONY audiotape recorder. The procedure is described in greater detail in chapter 4.

### 5.5.3 Transcription

The examiner transcribed all recordings of language samples using the computer software SALT (*Systematic Analysis of Language Transcripts* by Miller & Chapman), research version 7. Since no conventions for the transcription of language samples with SALT exist for German, the researcher developed new conventions. Section 4.8 in chapter 4 gives intra- and interjudge agreement values for 10,6% of the transcripts.

Suffixes representing present tense verb inflections were transcribed as bound morphemes, separated from the verbstem by a slash. Homomorphs were coded differently in order to identify person and number of the inflection according to the grammatical context in which they appeared (e.g. the inflection –*t* can mark 3$^{rd}$ person singular (transcribed as */3t*) as well as 2$^{nd}$ person plural present tense (transcribed as */2t*)). Nonfinite verbs were transcribed as two morphemes, too, as the infinitive is represented by the verb stem and the suffix *–en*. This allowed the distinction between the production of the verb stem and nonfinite verbs. Appendix 1 gives an overview of all codes used for bound morphemes. The transcription codes were used for lexical verbs only, because modal verbs, auxiliaries or the copula *sein* (Engl.: *to be*) have irregular forms which need to be learned on an item-by-item base. Furthermore, word codes were inserted following the copula *sein* (*CS*). Contracted forms of the 3rd person singular form of *sein* (*ist* → *'s)* were marked with the same word code as the other forms. However, it was often unclear whether the contracted form had been produced. These ambiguous cases were marked with a different word code (*AmCS*).

In order to analyse errors in present tense marking, inflections omitted from obligatory context were marked with an asterisk and error codes were inserted. However, if the phonotactic context of an inflection, or background noise did not allow an unambiguous decision whether the inflection had been produced correctly it was not transcribed as separate morpheme. Instead, a word code was inserted indicating the ambiguity. Incorrectly inflected verbs were followed by error codes. These provided information over the sort of error (e.g. *EVI* (wrong inflection), *EVT* (wrong tense), *EVIV* (nonfinite verb)) and the target inflection (e.g. *EVI-3PS*). The omission of the copula was indicated by error codes, too. A list of the error codes is presented in Appendix 1.

### *5.5.4 Quantitative analysis*

Obligatory contexts for an individual present tense inflection were computed by adding the total of correct productions[10], the number of omissions and the number of error codes referring to that particular inflection. The production of a verb was considered as pre-requisite for an obligatory context of verb inflections. Errors referring to the omission of the entire verb were therefore not included into the calculation of obligatory contexts.

[10] as indicated by the number of bound morpheme codes for the particular inflection excluding those followed by an error code

Inflections were selected for statistical between group comparison if at least eight children per group produced a minimum of two obligatory contexts. The data of any child entered analysis if the child produced a minimum of two obligatory contexts for the respective inflection. Percentages of correct use in obligatory context were calculated by dividing the number of correct productions of the inflection by the number of correct productions plus the number of grammatical errors. The following errors were considered as grammatical errors:

- nonfinite verb in the context of finite verb
- incorrect verb inflection
- verb stem (omission of the inflection)

Errors that could have resulted from the children's individual speech disorder were excluded from analysis. For this, the children's speech output was checked for relevant phonological processes that may have caused the occurring error patterns. Ambiguous cases were excluded from the calculation of percentages, too. Errors in the verb stem, i.e. an incorrect vowel change (e.g. a strong verb conjugated regularly), were not considered as error.

The production of the copula *sein* was not considered individually for between group comparisons but for a composite score only. The computation of a percentage correct followed the same guidelines as for the other present tense morphemes.

Arc-sine transformations were applied to the percentage data and between-group comparisons performed using analysis of variance (ANOVA) followed by Gabriel's procedure. If Levene's test indicated that the assumption of homogeneity of variance was violated, the Games-Howell procedure was used as post hoc test instead. If significant between group differences were found between the SLI group and the ND-A or ND-L group, the variable was further explored as clinical marker by calculation sensitivity, specificity, LR+ and LR-. The results of the children with SLI were transformed to *t*-values and percentiles relative to the children's age (i.e. ND-A data) and, secondly, relative to the children's language development (i.e. ND-L data).

### *5.5.5 Variables that meet the criterion*

Two verb inflections in present tense met the criterion for statistical analysis:

- 3rd person singular
- 3rd person plural

The children did not produce a sufficient number of obligatory contexts for the remaining verb inflections. However, as for the elicited data, present tense composite scores were computed for the spontaneous data, too. Three different composite scores were computed. In general, all correct productions of relevant inflections were taken together. This total score correct was then related to the total number of obligatory contexts of the inflections in order to calculate a general percentage correct for present tense marking. Again, errors explicable on the basis of the children's speech output as well as ambiguous cases were excluded from these analyses. Regarding 2nd person singular inflections, only the phonemically correct form was counted as correct.

The first present tense composite was derived from combining all present tense verb inflections in lexical verbs. Thus, all verb inflections were considered as long as the respective child showed a minimum total of two obligatory contexts for present tense inflections. For the second composite score, the production accuracy of the copula *sein* (Engl.: *to be*) in present tense was added for better comparability to composites suggested for English-speaking children. Finally, in order to allow a direct comparison between elicited and spontaneous measures, a composite was calculated for 2nd and 3rd person singular inflection only. The three following versions of a present tense composite were consequently investigated:

- Present tense: production accuracy of any present tense marking in lexical verbs
- Present tense + copula: production accuracy of any present tense marking in lexical verbs + the copula *sein* in present tense
- Production accuracy of 2nd + 3rd person singular present tense in lexical verbs.

### *5.5.6 Results of quantitative analysis*

#### *5.5.6.1 Verb inflections in present tense*

Table 15 gives the summary data for each group for 3rd person singular and plural inflections in present tense.

**Table 15.** Language Samples: Occurrence of Obligatory Contexts for Morphemes in Present Tense

| *Inflection* | | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|---|
| **3rd singular** | **M** | 10.62 | 8.85 | 8.94 | 9.69 |
| | **SD** | 6.01 | 6.61 | 4.59 | 5.78 |
| | **Range** | 3 – 24 | 2 – 24 | 2 – 19 | 3 – 21 |
| | **Total** | 138 | 115 | 152 | 155 |
| **3rd plural** | **M** | 3.3 | 4.59 | 5.7 | 3.4 |
| | **SD** | 0.95 | 2.91 | 4.40 | 2.46 |
| | **Range** | 2 – 5 | 2 – 11 | 3 – 17 | 2 – 10 |
| | **Total** | 33 | 55 | 57 | 34 |

*Note*.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation

Group sizes and the descriptives of the statistical analysis are given below in Table 16. These scores reflect the proportional accuracy in obligatory context; transformed arcsine values were entered for further statistical analysis.

**Table 16.** Language Samples: Proportional Scores of Correct Use of Present Tense Verb Inflections

| *Inflection* | | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|---|
| **3rd singular** | **M** | 0.71 | 0.84 | 0.91 | 0.87 |
| | **SD** | 0.27 | 0.29 | 0.13 | 0.18 |
| | **Range** | 0.08 – 1 | 0.0 – 1 | 0.64 – 1 | 0.47 – 1 |
| | **n** | 13 | 13 | 17 | 16 |
| **3rd plural** | **M** | 0.81 | 0.83 | 1 | 0.99 |
| | **SD** | 0.27 | 0.26 | 0.0 | 0.03 |
| | **Range** | 0.33 – 1 | 0.25 – 1 | - | 0.9 – 1 |
| | **n** | 10 | 12 | 10 | 10 |

*Note*. No significant between group differences when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

Levene's test showed that the assumption of homogeneity of variance was met only for 3rd person singular verb inflection. ANOVA did not show a main effect for the production of 3rd person singular. Since the ND-A group showed a ceiling effect in the use of 3rd person plural inflection, Brown-Forsythe *F*-ratio was not available for this morpheme. Post hoc procedures reflected the results above and did not reveal any significant difference for either of both morphemes. Results are illustrated in **Error! Reference source not found.**, presenting the different groups means of the proportional scores before arcsine transformations.

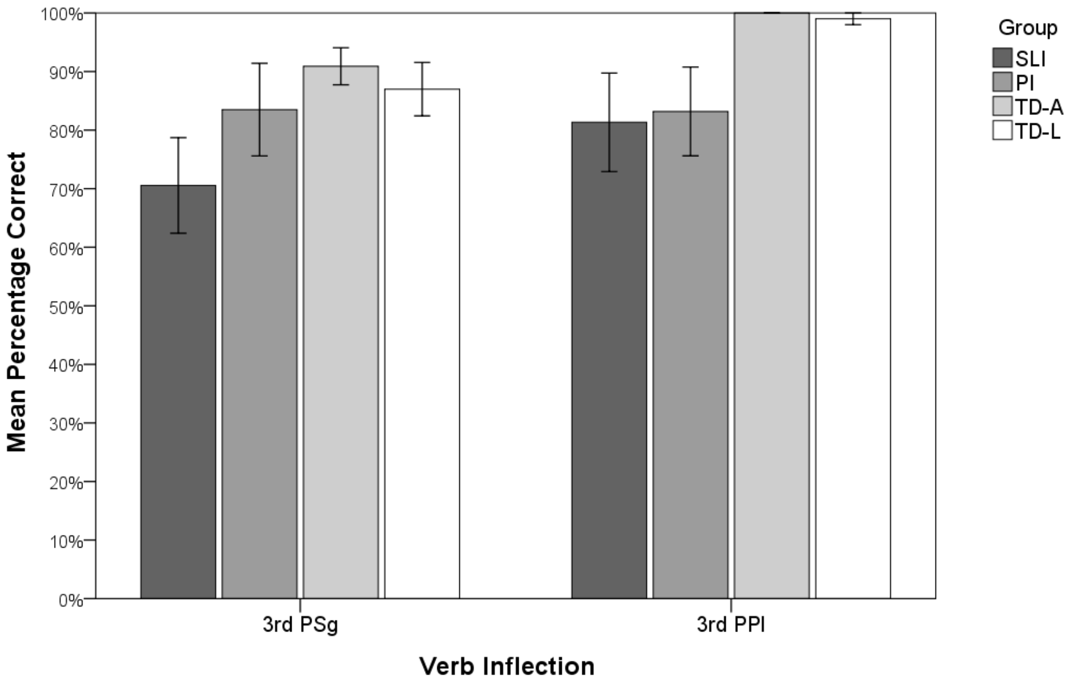



**Figure 7.** Mean percentages of use (± 1 SE) for spontaneous production of third person singular and third person plural verb inflection in present tense.

### 5.5.6.2 Present tense composite from spontaneous language data

Obligatory contexts for present tense marking were taken together for three different composites. Table 17 provides the descriptives for each of these composites.

**Table 17.** Language Samples: Compositional Scores Present Tense

| ***Compositional Score*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **2nd + 3rd person singular present tense** | **M** | 0.65 | 0.84 | 0.89 | 0.83 |
| | **SD** | 0.32 | 0.27 | 0.17 | 0.27 |
| | **Range** | 0 – 1 | 0 – 1 | 0.5 – 1 | 0 – 1 |
| | **n** | 14 | 15 | 17 | 16 |
| **Present tense** | **M** | $0.74_a$ | 0.82 | $0.93_a$ | 0.88 |
| | **SD** | 0.20 | 0.26 | 0.10 | 0.16 |
| | **Range** | 0.28 – 1 | 0.09 – 1 | 0.67 – 1 | 0.47 – 1 |
| | **n** | 14 | 14 | 17 | 16 |
| **Present tense + Cop** | **M** | $0.77_{a,b,c}$ | $0.92_a$ | $0.97_b$ | $0.94_c$ |
| | **SD** | 0.25 | 0.10 | 0.06 | 0.08 |
| | **Range** | 0 – 1 | 0.59 – 1 | 0.75 – 1 | 0.66 – 1 |
| | **n** | 15 | 16 | 17 | 16 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants
Cop = Copular *sein*

The production accuracy for the copula *sein* was included for the third composite. The SLI group (*n*=15) showed an overall accuracy of 80%, the PI group a mean accuracy of 95%, the ND-A group produced on average 99% of the copula correctly and the ND-L group 97%. As previous research did not suggest the copula alone as clinical marker, the data on this variable is not presented in greater detail.

For further statistical analysis, all proportional scores were transformed to arcsine values. Two present tense composites from spontaneous language showed a main group effect when analysed with one-way ANOVA: present tense ($F$ (3, 1.67) = 3.54; $p$ = .02, $\omega^2$ = .11) and present tense + copula ($F$ (3, 1.80) = 8.64; $p$ = .000, $\omega^2$ = .26). No significant group effect was evident for the composite 2$^{nd}$ + 3$^{rd}$ person singular. Individual between group comparisons with Gabriel's procedure demonstrated that the SLI group produced significantly less correct present tense inflections than the ND-A group ($p$ = .021). The composite including the copula (present tense + copula) revealed significant differences between the SLI and all three other groups (SLI > ND-A ($p$ = .000), SLI < ND-L ($p$ = .003) and SLI < PI ($p$ = .022). Figure 8 illustrates the three proportional present tense composite scores for all four groups.

If using the present tense composite from spontaneous language as diagnostic tool, a cut-off of 85% was the most suitable cut-off as the overview over the data in Table 44 demonstrates. This cut-off resulted in a in a diagnostic accuracy of 84%, a sensitivity of 0.786 [0.524 - 0.924], a specificity of 0.882 [0.657 - 0.967], LR+ = 6.679 [1.766 - 25.259] and LR- = 0.243 [0.088 -0.672]. If the production accuracy of the copula was included in the composite, the data suggest a cut-off of 92% as shown in Table 45. This resulted in an overall discrimination accuracy of 91%, a sensitivity of 0.867 [0.621-0.963], a specificity of 0.941 [0.73 - 0.99], LR+ = 14.733 [2.178 - 99.672] and LR- = 0.142 [0.039 - 0.518].

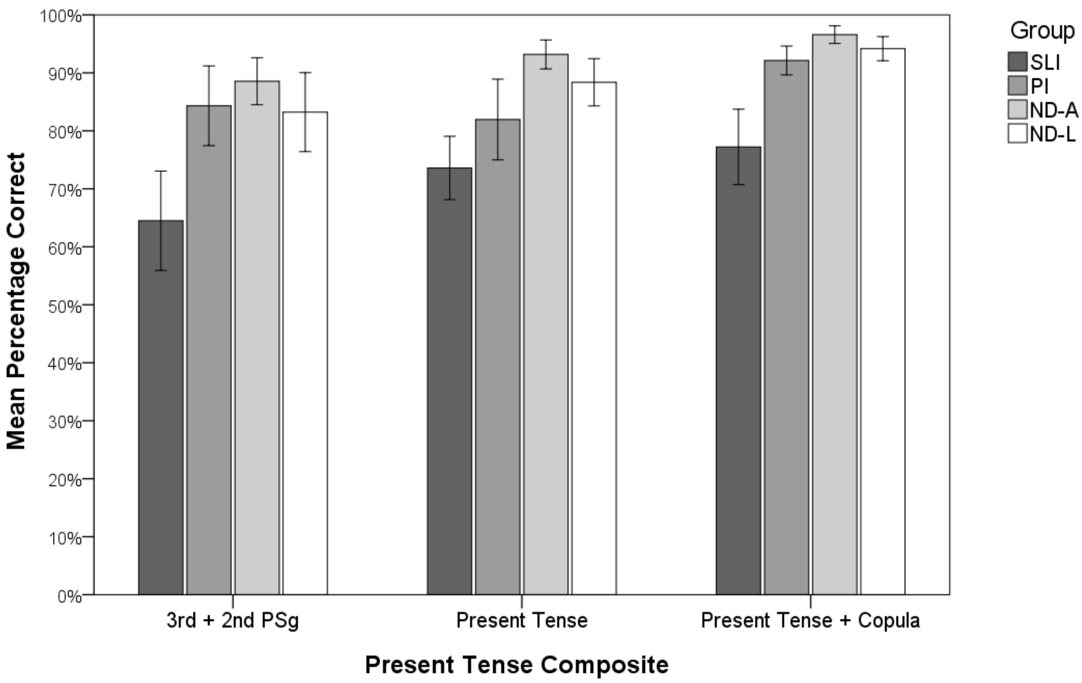



**Figure 8.** Mean percentages (± 1 SE) for present tense composites representing proportional correct production in spontaneous language.

### *5.5.7 Qualitative analysis*

A qualitative analysis was performed for the same variables that met the criterion for statistical analysis. The spontaneous production of 3rd person singular and 3rd person plural in present tense was therefore further analysed regarding production patterns. The total number of obligatory contexts is the basis for the relative frequency of each production pattern, i.e. the proportion of correct responses and the proportion of all error patterns respectively. The frequency of each production pattern per group was divided by the total number of obligatory contexts for each target inflection in order to calculate the proportion of each production pattern.

No statistical between group comparisons were performed regarding production pattern because the available quantity of data per category was too limited.

### *5.5.8 Results of qualitative analysis*

The following five production patterns occurred for both, 3rd person singular and 3rd plural inflection.

- Correct production
- Verb infinitive
- Verb stem (omission of inflection)
- Incorrect inflection
- Vowel change in verb stem

Errors that could be considered as resulting from phonological processes evident in the child's speech (category *phonological error*) occurred only in 3rd person singular. Thus, six categories of production patterns were considered for 3rd person singular inflection, five categories were considered for 3rd plural inflection. The infinitive ending and the 3rd person plural inflection share the same surface *–en.* Only if the ending *-en* occurred in sentence final position it was considered as infinitive. In other positions, it was considered as correct.[11].

[11] Several researchers showed an association between verb finiteness and verb placement in German SLI (Clahsen, 1989; Clahsen et al., 1997; Rice et al., 1997) .

Figure 9 illustrates the distribution of these production patterns for each group in the context of 3rd person singular present tense.

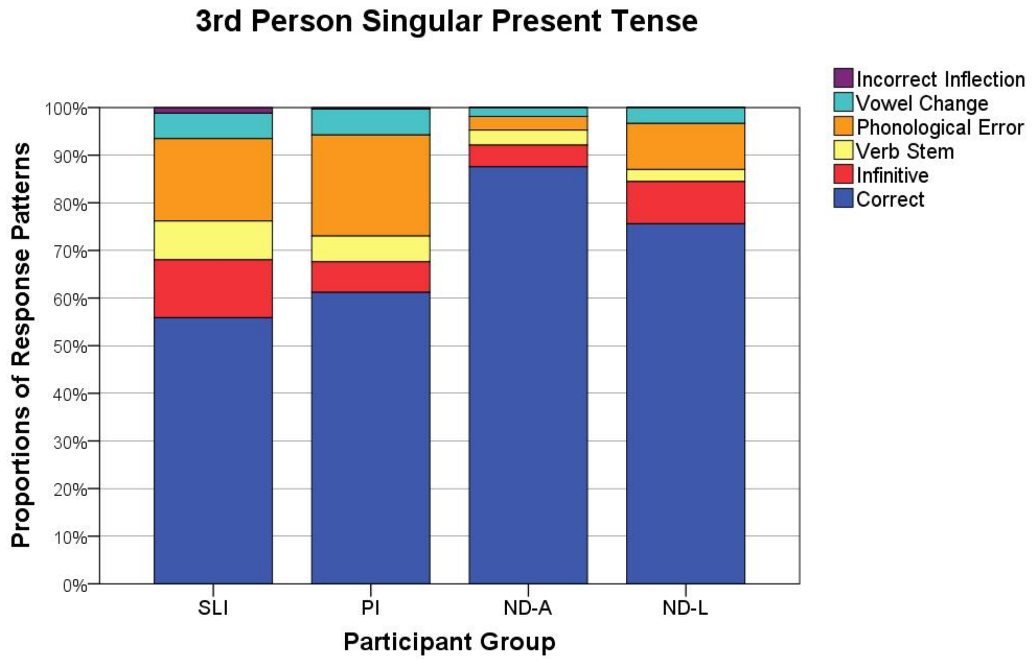



**Figure 9.** Spontaneous production patterns for 3rd person singular inflection in present tense.

Figure 10 gives the distribution of the production patterns in the context of 3rd person plural present tense.

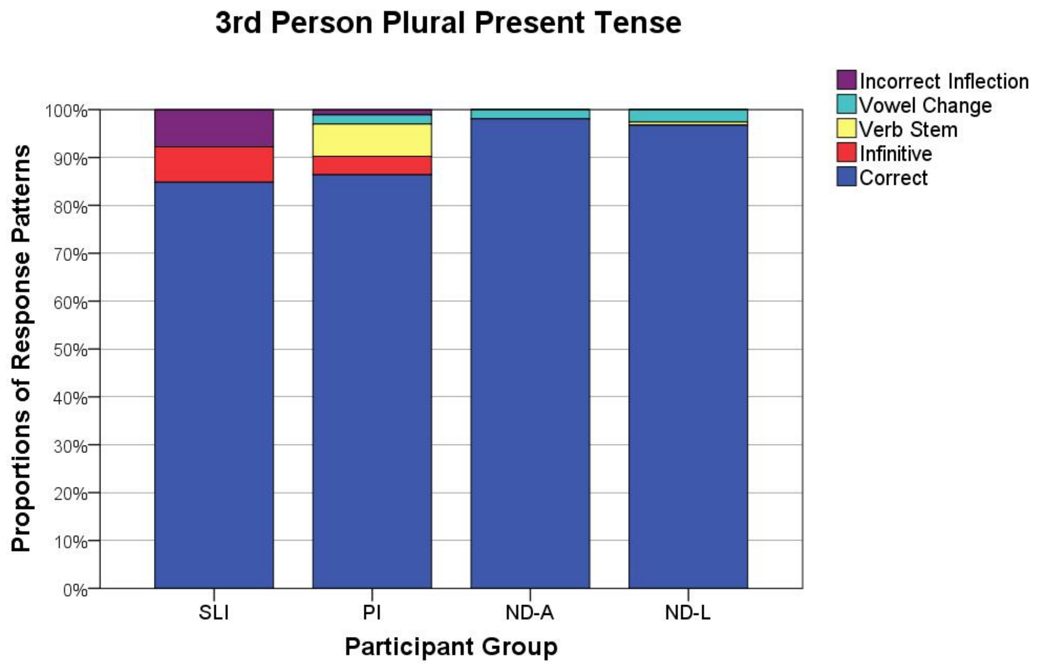



**Figure 10.** Spontaneous production patterns for 3rd person plural inflection in present tense.

Rounded proportions in relation to the total of obligatory contexts are presented in Table 18 for both 3rd person singular and 3rd person plural. Between group comparisons were not performed for the different production patterns. The available data was too limited for such statistical analysis as a consequence of the division into the different categories.

**Table 18.** Language Sample: Production Patterns for Present Tense Inflections

| ***Production Pattern*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **3rd singular** | **Correct** | 56% | 61% | 88% | 76% |
| | **Infinitive** | 12% | 6% | 5% | 9% |
| | **Verb stem** | 8% | 5% | 3% | 3% |
| | **Phonological Error** | 17% | 21% | 3% | 10% |
| | **Vowel change** | 5% | 5% | 2% | 3% |
| | **Incorrect inflection** | 1% | 0% | 0% | 0% |
| **3rd plural** | **Correct** | 85% | 86% | 98% | 97% |
| | **Infinitive** | 7% | 4% | 0% | 0% |
| | **Verb stem** | 0% | 7% | 0% | 1% |
| | **Incorrect inflection** | 8% | 2% | 0% | 0% |
| | **Vowel change** | 0% | 1% | 2% | 3% |

*Note.* Percentages are rounded.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

### *5.5.9 Interpretation*

The results of the spontaneous productions of present tense inflections presented a promising delay within delay pattern if all present tense inflections were combined to a composite score and complemented by the production accuracy for the copula *sein*. Furthermore, also the production accuracy of the PI group was significantly higher than that of the SLI group on this composite. Present tense marking including the copula *sein* may therefore be a suitable candidate for a clinical marker for German SLI. The composite shows potential in differentiating children with SLI not only from their age matched peers, but also from language matched children and children with PI. The analysis of individual present tense inflections, on the other hand, did not result in significant between group differences. A composite excluding the copula indicated a significant delay, but the children with SLI did neither lag behind the ND-L group nor behind the PI group. The findings will be discussed below further regarding their relevance for SLI, PI and the findings from the elicitation tasks.

## 5.6 Discussion

### *5.6.1 Verb morphology in present tense as clinical marker for German SLI?*

The primary intention of the presented study was to investigate the potential of verb inflections in present tense as clinical marker for SLI in German. Findings from two measures were presented above: morphology probes involving elicitation tasks for the 2nd and 3rd person singular inflection and spontaneous language samples. The results from between group comparisons across both measures are summarised in Table 19.

**Table 19.** Results From Between Group Comparisons for Present Tense

| | Data source | |
|---|---|---|
| ***Dependant variable*** | ***Elicitation tasks*** | ***Spontaneous language*** |
| 3rd person singular | SLI < ND-A | - |
| 2nd person singular | | |
| Phonetically correct form -*st* | main effect | n.a. |
| Grammatical contrast -*s* | main effect | n.a. |
| Present tense composite | | |
| 3rd + 2nd person singular | SLI < ND-A | - |
| 3rd + 2nd (-*s*) person singular | SLI < ND-A; PI < ND-A | not tested |
| Present tense (all inflections) | n.a. | SLI < ND-A |
| Present tense + copula | n.a. | SLI < PI; SLI < ND-A; SLI < ND-L |

*Note*. Only statistically significant group differences are presented ($p < 0.05$).
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

***Delay within delay***

If we follow the approach of Rice (2003), one of the first steps should be the identification of morphemes on which children with SLI perform below both age-matched and language-matched control children. This addresses the question whether we can find a specific delay in the children's morphology going beyond the children's general language problems. A look at the results from the present study shows that no more than one variable fulfils this criterion: the composite measure comprising all present tense inflections in spontaneous language and the production of the copula *sein.* Unfortunately, data on this particular composite score is available from only one of the measures. The elicitation tasks did not tap the production of the copula *sein* and, consequently, no conclusion can been drawn from this data source. The data from the language samples, however, indicates an advantage for both groups of typically developing children on this composite relative to the SLI group. Additionally, the composite raises hope for a clinical marker specific to SLI because between-group comparisons revealed a significant difference between the SLI and the PI group, too.

The discrimination accuracy of the present tense composite including the copula seems satisfactory, too. The sensitivity rate reaches 87%, the specificity even 94%. Also LR+ is larger than 10 and LR- lower than 0.2 which are the thresholds indicating a good discrimination accuracy according to Klee, Gavin & Stokes (2007).

The composite has nevertheless a flaw as potential marker: A look at the group means indicates a relatively high level of proficiency across all four groups. This is reflected in the very high cut-off for discrimination of 92% production accuracy for present tense inflections and the copula. An accuracy rate of 90% or above is often even considered as “adult like grammar” (e.g.Rice et al., 1997, p. 275).

In comparison to previous results, this high accuracy rate seems surprising. As mentioned above, Rice et al. (1997) reported a rate of 52% nonfinite verbs in the context of finite verbs in main clauses. The analysis of the present study differs somehow from that of Rice et al.. Rice et al. investigated four different variables in their data: first, the accuracy rate if 3rd and, second, 2nd person inflections appeared, third, the omission rate of the copula *sein* and fourth, the finiteness of verbs in main clauses with three or more constituents. The last variable resulted thus in a minimum error rate of 52%, i.e. the children with SLI inflected only 48% of the verbs requiring finiteness. The accuracy rate for the copula was 69% during the first sampling period. This time of measurement corresponds best to the age range of the current study and comparable results could therefore be expected. However, the children of the present study seem considerably more proficient in the production of verb inflections in present tense as well as the copula.

This is certainly true regarding the composite score including the copula as reference. Figure 8 illustrates, though, that the production accuracy is actually the highest for this composite score compared to the other two composites computed for the current data. Considering the correctness of 3rd and 2nd person singular inflections alone, the overall accuracy is lower and reaches only 65%. This lies still above the accuracy reported by Rice et al., but is closer. The fact that the SLI group falls significantly behind all three control groups on the composite including copula is due to the fact that the other groups perform all near ceiling level. It is also for this composite, the three control groups present the smallest variability across the group. The SLI group benefits thus the least of the inclusion of the copula to the composite score. This again explains that the group differences are statistically significant, rather than that a particularly low production accuracy in the SLI group can account for the results.

Looking at the children's individual performance, one child (SLI3) of the SLI group who produced obligatory contexts for the present tense composite differs from the rest of the sample included in the analysis. While this boy produced no obligatory context for any of the verb inflections in present tense he showed three obligatory contexts for the copula '*sein*' in his spontaneous language. He omitted the copula in all three contexts, though, and his percentage score was therefore zero, not only on the variable *copula* but also on the composite. In contrast, the number of obligatory contexts for a present tense composite varied between 19 and 53 across the remaining 14 children with SLI included in the analysis. Furthermore, the other children had an accuracy level from 49% to 100% on present tense marking including the copula. It seems therefore reasonable to consider this boy as an outlier and to exclude him from analysis. Posthoc group comparisons with this SLI subgroup shows that the remaining 14 children differ nevertheless from both groups of typically developing children. The PI group, though, does not perform any longer significantly better than the SLI group. Thus, the marker looses in its specificity. Additionally, as a matter of course, the mean percentage correct on the composite measure increases with the exclusion of the boy to 85%. This illustrates that the discriminative power of a clinical marker with a cut-off of 92% production accuracy cannot be ideal. This is reflected in a very large 95% confidence interval for the LR+. In fact, three children with SLI would not be identified by such a clinical marker and four further children have an accuracy level on this composite between 87 and 90%. This indicates that the children with SLI are overall too proficient on present tense marking in main verbs and the copula '*sein*'. Although between-group comparisons reveal promising differences, present tense marking can consequently be rejected as possible clinical marker due to a substantial overlap between the clinical and the age-matched typically developing children.

Nevertheless, the question remains whether this conclusion is representative for German-speaking children with SLI given that Rice et al. reported a higher error rate in verb finiteness. Rice et al. considered carefully the influence of individual speech errors and included only –*n* and –*en* endings as nonfinite, but no bare stems. Another factor that may well have contributed to the high error rate in Rice et al.'s data, though, is the fact that only those utterances entered analysis that consisted of a minimum of three constituents. It is possible that children's ability to mark verbs for subject-verb agreement is related to the verb argument structure within the utterance. Such an affect has been reported in relation with the omission of auxiliaries (Grela & Leonard, 2000; Pizzioli & Schelstraete, 2008). In the present data, the argument structure of the utterance did not influence the selection for analysis. This may have facilitated a higher performance.

The comparison to further research data on German SLI is difficult as most authors focused on older children. Thus, even if they reported higher accuracy rates in verb finiteness than Rice et al. (e.g.Hamann, Penner, & Lindner, 1998; Rothweiler, Chilla, & Clahsen, 2012) it may have been the age that accounts for this difference in findings. Furthermore, Rice et al.'s work is unfortunately the only study on German SLI mentioned above that incorporated a control group with language-matched, typically developing children. Recently, though, Ott and Höhle (2013) presented data of German speaking children with SLI (aged 4;1 – 5;1) in comparison to both age-matched and language-matched typically developing children. The children were asked to participate in a sentence completion task with nonce verbs that needed to be inflected for 3rd person singular in present tense. The overall accuracy rate was below 50%, too, but the children with SLI differed in this significantly only from the age-matched control children, and not to their language-matched peers. The children of the current study showed a lower performance in elicited subject-verb agreement than in spontaneous production. It is hence likely that the low performance accuracy presented by Ott and Höhle is partly due to the procedure and the fact that the children actually had to inflect nonce rather than real verbs. Although there is little further evidence, it seems therefore reasonable to consider the data presented in the current study as representative. It can consequently be concluded, that the high performance accuracy of children with SLI prevents the use of the composite present tense inflections including copula as clinical marker.

***Delay in present tense verb inflections as clinical marker?***

As summarised above, several authors reported for German-speaking children with SLI marked difficulties in verb inflections (e.g. Clahsen, 1989, 1991; Clahsen et al., 1997; Rice et al., 1997; Schöler, Fromm, & Kany, 1998a). Even if the present data does not suggest a delay relative to the children's other language problems, testing the production of verb inflections may have clinical potential for the distinction between children with SLI and typically developing children of same age. The variables' potential as clinical marker for SLI will be discussed in the following section. The individual grammatical forms will be discussed first, followed by the present tense composites from both measure tools.

**Individual morphemes**

A look at the mean accuracy scores for the individual verb inflections suggests a considerable deficit for the SLI group in comparison to the ND-A group across all present tense morphemes. This is true not only for the elicited data but also for 3$^{rd}$ person singular and plural in spontaneous speech, although the difference between the SLI and the ND-A group is less pronounced. Figure 3 and **Error! Reference source not found.** illustrate this. However, the between roup difference is only for elicited 3$^{rd}$ person singular statistically significant, although the numerical difference is similarly large for the two variables on elicited 2$^{nd}$ person singular: across all three variables from the probe data, the group difference extends 30%. The lack of significant results for the 2$^{nd}$ person singular variables is probably due to a large within group variability within the SLI group. Furthermore, also the group size was smaller for 2$^{nd}$ person singular as more children refused to attempt all tasks. It may therefore be possible that a larger data set can present more promising results not only for 3$^{rd}$ person singular but for the 2$^{nd}$ person singular inflection, too.

The potential of further research is questionable, though. As mentioned above, a significant delay could be expected according to previous results. The large variability in the present data, on the other hand, indicates that a possible clinical marker in present tense inflections would not apply to the general population of German-speaking children with SLI. Even though the within group variability may decrease in further research with more test items and a larger participant group, the present data illustrates the heterogeneity of the SLI group in this respect very clearly. For example, in elicited 3rd person singular, three out of thirteen children performed at ceiling level. For the 2nd person singular inflection, six out of eleven children produced the grammatical contrast *–s* always correctly. This heterogeneity of the population may be less evident in a larger data set but it would not disappear. Nevertheless, it would be useful to provide further evidence from German-speaking children with SLI for this claim. The evidence available so far has the disadvantage that several studies refer to the same set of data collected decades ago by Clahsen at al. (e.g. 1989; 1997; 1992). So far, it can be concluded that the delay in the production of individual present tense inflections is considerable in some but not all children with SLI. It is therefore no promising candidate as clinical marker.

**Present tense composite**

Regarding the present tense inflection composite, two versions exist for the morphology screening: One following the strict criterion for 2nd person singular (only *–st* correct) and another accepting *–s* as marking this verb inflection. Both versions result nevertheless in a considerable overlap between both groups. A marker following the stricter criterion in 2nd person singular and applying a cut-off of 56% cannot classify the children satisfactorily into children with or without SLI. If the more relaxed criterion is applied to 2nd person singular in the elicitation tasks, i.e. the allomorph *-s* considered as correct, four children with SLI perform even at an accuracy level of 87% or higher. Nevertheless, a cut-off of 88% accuracy would identify 85% of the children with SLI correctly and 94% of the ND-A children. Both LR values (LR+=14.39; LR-=0.16) reflect a good classification accuracy, too. Can we consequently conclude that a composite from 2nd and 3rd present tense inflection tested in elicitation tasks and scored according to phonological rather than phonetic accuracy is the clinical marker we were looking for?

It is worth looking at the results from the language sample. The present tense composite for the spontaneous data would suggest a cut-off of 85% proficiency as most suitable - thus, again a very high level of proficiency. Section 2.5.1

indicates that classification accuracy derived from the language sample data is not adequate, though. Since the present tense composite from spontaneous data comprises all present tense inflections rather than 2nd and 3rd person singular only, an additional score was computed representing the proficiency of only these two present tense inflections. ANOVA, however, does not even show a significant group effect for this score. This fact is little surprising given the fact that 3rd person singular in spontaneous language did not result in significant between-group differences either and only few children produced obligatory contexts for 2nd person singular. The spontaneous language data can consequently not support the choice of present tense marking as a clinical marker for SLI in German.

Overall, a present tense composite does not seem to be a suitable marker despite the highly promising results from the morphology screening. The data demonstrated clearly that even the children with SLI show a high level of proficiency on these morphemes (reflected in a cut-off of 92%). By no means, one could call this a marked deficit in present tense marking. Moreover, the analysis indicated a considerable inconsistency of the classification accuracy. If scoring criteria or data sources change, the marker looses immediately its strength. A suitable marker, though, should not only mark a distinct deficit but should furthermore be more robust to changes in data collection and scoring system.

### *5.6.2 Verb inflections in German PI*

The group of children with SLI was not only compared to typically developing children in this study but also to a group of age and gender matched children with PI. A clinical marker for SLI should be specific to this impairment only, but not to another developmental disorder of comparable prevalence[12]. The selection criteria for the PI group involved that the children performed within age expectations on a standardised language test. Given this, we should expect the children with PI showing similar levels of proficiency in present tense inflections as their typically developing age matched peers.

The data presented above, though, indicates almost across all variables that this is not the case. The only variable for which the PI group demonstrates an age equivalent proficiency level is the spontaneous production of the copula *sein* in present tense. This variable computed for the inclusion in a present tense composite is also the only variable for which the PI group is significantly ahead of the SLI group. Post hoc tests, on the other hand, indicated that the between group difference between the PI and the ND-A group reaches statistical significance only for one of the present tense composite derived from the probe data. For this variable, the PI group performs rather on a similar level as the SLI group. A look at the mean percentages correct indicates that this is also the case regarding the elicited 2nd person singular and the spontaneous production of 3rd person plural. In 3rd person singular marking, the PI group performs across both measures at a level comparable to the younger typically developing children. The question arises therefore whether the children of the PI group had been identified correctly. Are the children with PI indeed distinct from the children with SLI?

The qualitative analysis can provide some answers. Although phonological errors had been excluded for the statistical between group comparisons this category of errors was included in the analysis of error patterns. It is this error pattern that should be the predominant one among the children with PI. Grammatical errors, on the other hand, should be rare. The children with PI

---

[12] 3-10% of children show a speech output disorder (Fox, 2003; Shriberg et al., 1999)

presented on average indeed a higher rate of phonological errors than the SLI children, although no statistical analysis was carried out that could confirm the significance of these findings. The rate of phonological errors supports nevertheless the distinction between both groups strongly. This is especially the case for the probe data. For the 3rd person singular inflection, 9% of the responses entailed phonological rather than grammatical errors in the PI group, in the SLI group none. Regarding 2nd person singular, the rate of phonological errors lies at 20% for the PI group but at 3% only for the SLI group. Moreover, it seems more difficult for the children with PI to produce the phonologically correct ending *-st* for this inflection. The rate of producing merely the grammatical contrast *-s* instead of *-st* is double in the PI group compared to the SLI children. In spontaneous language, too, the rate of phonological errors is high (21%). However, even children with SLI presented an average rate of 17% phonological errors in spontaneous language. The distinction is thus less clear in spontaneous language. The elicitation tasks, on the other hand, seem to have facilitated the phonology of the children with SLI but not that of the children with PI. This is especially noteworthy given the fact that the SLI group entailed 12 children that were diagnosed with a co-occurring phonological delay or disorder (s. chapter 4). This indicates that the focus of the problem of the children with SLI lies indeed in the grammatical domain rather than in phonology.

While phonological errors were to be expected for the children with PI, grammatical errors were not. The qualitative analysis illustrates that this prediction was not entirely met. In 3rd person singular and plural, grammatical errors were still more frequent in the SLI than in the PI group, but also some children with PI produced infinitives, overgeneralisations or the verb stem only. In elicited 2nd person singular, even wrong inflections occurred and the general frequency of grammatical errors is as high in the PI as in the SLI group. How can these results be explained given that the children with PI scored within age-expectations on a language measure?

A look at the individual children’s data shows that these errors are not ascribable to a handful of children but that the grammatical errors are distributed across a number of children with PI. The possibility that the results for the PI group are distorted due to a few outliers only can therefore be rejected. The question remains, thus, whether the SETK 3-5 (Grimm, 2001) is sensitive enough to detect language difficulties or whether the criterion set for participant selection was too generous. The PI group presented on all subtests of the SETK 3-5 an average t-score of 49 or above. According to the SETK, the children should have an age equivalent language development. The ND-A group, on the other hand, performed notably better on the SETK 3-5. The ND-A group presented on average a minimum mean t-score of 60 across all sub-tests. If we assume that the SETK is sensitive enough to pick up language difficulties that are no longer age-equivalent we can conclude that it is the ND-A group that is slightly better than expected in their language skills rather than the PI group falling behind age expectations. That would further indicate that the grammatical errors observed in the PI group may still be age appropriate.

Comparing both groups regarding their grammatical error patterns, little differences occur in spontaneous language between the PI and the ND-A group. The error analysis for the probe data, though, shows more distinct differences. While the rate of verb stems is still comparable across both groups, the rate of infinitives is not. None of the children in the ND-A group used infinitives in the elicitation tasks. From the PI group, on the other hand, five children did so during the 3$^{rd}$ person singular task and three children during the 2$^{nd}$ person singular task. As noted above, only one child produced infinitives in both tasks, i.e. a total of seven out of sixteen children produced at least one infinitive instead of a finite verb during the present tense task. This is a considerable part of the group. Furthermore, it is unlikely that this can still be called age equivalent as suggested above, since the ND-L group showed much less instances of infinitives than the PI group. Although none of the children with PI produced the infinitive consequently across all tasks, this error pattern is therefore surprising.

Moreover, the omission of an inflection (i.e. the production of the verb stem) would rather be an error pattern to be expected from children with PI than the insertion of the infinitive ending *-en.* It seems likely that children with a phonological problem experience difficulties to process word-final consonant or consonant cluster efficiently if they occur only in certain contexts. As suggested by Chiat in her phonological theory (2001), children may therefore need longer to acquire the grammatical rule accompanying the verb inflection. The present data, however, presents that some children with PI do not omit the inflection but use the infinitive instead which consists of one syllable more than the finite verb. One of the possibilities is that the children did not process the complete pragmatic context of the task. The verb inflections were prompted in the probes by showing a picture and asking for example "What is the girl doing? The girl … ." If the children answered to the question only but did not intend to complete the following sentence, the infinitive of the verb would have been the correct answer. It is therefore possible that a larger proportion of children of the PI group were too quick answering to the question instead of to the probe. Findings of Liiva and Cleave (2005), for example, would support this suggestion as they found positive correlations of pragmatic difficulties with the severity of expressive speech and language limitations. The results from the spontaneous language samples support this further. The children with PI produced on average fewer infinitives when they constructed their utterances themselves than when they answered to the prompts of the elicitation tasks. Moreover, the fact that none of the children showed a general inability to produce a finite verb indicates that it was not the grammatical rule they lacked. Consequently, the present findings of the PI group do not necessarily result from inaccuracies in the participant selection.

Findings from Ott and Höhle (2013) support this. As mentioned above, Ott and Höhle found that children with SLI experienced more difficulties in inflecting nonce verbs correctly for 3rd person singular than age matched typically developing children. Yet, in contrast to the present study, the children included in the authors SLI group were mainly characterised by lexical and / or phonological impairments. A phonological disorder was thus an explicit inclusion criterion, although it is not clear how large the proportion of children with an exclusive speech disorder was in the study by Ott and Höhle. Nevertheless, the largest proportion of error types was, even for this SLI group, the reproduction of the infinite verb. This supports the conclusion that this error type is not uncharacteristic for children with PI. Another finding of Ott and Höhle may provide an explanation. The authors found an effect of phonotactic frequency for the SLI group but not for typically developing children neither of similar chronological age nor of similar language age. Ott and Höhle suggested furthermore that syllable complexity, too, may have an impact on the SLI children's ability to produce verb inflections. Both factors may have played a role in the present data. Three of the five 3rd person singular tasks resulted in word final consonant clusters. The 2nd person singular task included four of five items that resulted in word final consonant clusters of three consonants (CCC). None of these CCCs occurs in other contexts than inflected verbs. Consequently, phonological complexity as well as syllable frequency might have had an impact on the children's production accuracy. This, however, could only be investigated further by re-analysing the spontaneous language productions of inflections according to their phonological context. Even then, the scope of such an analysis would be restricted as few obligatory contexts occurred in spontaneous language for 2nd person singular. It would therefore be the best to carry out further research in order to follow up the impact of phonological complexity and syllable frequency on real verbs.

### *5.6.3 Patterns in present tense verb inflections*

An alternative approach to the identification of children with SLI could be the search for deviant error patterns as clinical marker rather than the search for a significant delay or delay within delay. Such error analysis can also stimulate the search for underlying causes for the language difficulties, especially in a cross-linguistic perspective. The finding that German-speaking children with SLI do not only omit verb inflections but rather tend to produce the verb infinitive ending in -*en* (e.g. Clahsen, 1989; Clahsen et al., 1997; Rice et al., 1997) stimulated the development of different theories on SLI such as the missing agreement account by Clahsen (1989, 1991) or the extended optional infinitive hypothesis by Rice and colleagues (1997).

The error analyses of the present study allow comparisons between the SLI group and the typically developing children regarding their error patterns. It is important to note, though, that all conclusions rely on descriptive results only as qualitative group differences were not compared statistically. Conclusions can therefore be of only preliminary nature.

A look at the qualitative analyses illustrates that there are five response patterns that seem to occur more frequently in language impaired than typically developing children: the use of verb infinitives, the production of the verb stem, the production of an incorrect verb inflection, phonological errors and no attempt. The frequency of these response patterns varies partly across measures and verb inflections, though. Each response pattern will be discussed below.

**Infinitives**

The production of verb infinitives is one of the most frequent error patterns in German-speaking children with SLI according to a number of previous studies (Ott & Höhle, 2013; Rice et al., 1997). This was also the case in the present study. Although the production of the verb infinitive is typical in the early acquisition of German this pattern usually disappears around an age of 3 years

(Schrey-Dern, 2006). The use of verb infinitives above that age could therefore have potential as clinical marker.

In the present study, rates of verb infinitives differed considerably across the intended verb inflections. The highest rates were observed in the context of 3rd person singular. In the probe data, children with SLI produced on average 20% of the items as infinitive, and in spontaneous language about 12%. If the phonological errors are excluded, infinitive is thus the most frequent error pattern in 3rd person singular for the SLI group. The rates for elicited 2nd person singular (9%) and 3rd person plural (7%) were lower. However, the low number of infinitives in the context of 2nd person singular is probably due to a large number of *no attempts* for this verb inflection. The 3rd person plural verb inflection, on the other hand, may be easier as it has the same surface as the infinitive ending. Another possibility that may account for the lower infinitive rate in 3rd person plural may be that the criterion of verb second position was not an entirely valid indicator of verb finiteness. In this case, the number of verb infinitives would be higher for 3rd person plural, too. The typically developing children, on the other hand, produced few infinitives and in the context of 3rd person singular only. Within the ND-A group, infinitives occurred only in spontaneous language (5%). The ND-L group did not present this error pattern frequently, either, but in both measures (on average 6% and 9% respectively). The occurrence of verb infinitives in the context of finite verbs seems thus most characteristic for children with SLI but not for typically developing children. Frequencies of this error pattern, however, do not seem high enough in order to use this as clinical marker for SLI. Nevertheless, verb finiteness seems to be vulnerable in German SLI and the occurrence of verb infinitives in the context of finite verbs should always raise the attention of clinicians.

In comparison to previous research, a maximum of 20% verb infinitives in finite context seems rather low, though. Infinitive rates reported previously were considerably higher. Rice et al. (1997) presented a rate of 52% infinitives in the context of finite verbs for 3;9-4;8 year-old German-speaking children with SLI and a rate of 12% when the children were about one year older. The children’s age of the first measurement by Rice and colleagues is most comparable to the age of the current SLI group. Rothweiler, Chilla and Clahsen (2012) presented recently longitudinal data from seven German-speaking children aged between 3;1 and 7;11. Rates of nonfinite verbs varied between 11.4 and 49% in contexts of verb inflections relevant for the present study. However, the data analysis by Rothweiler and colleagues included verb stems. If the figures presented by the authors are re-calculated to a percentage score for infinitives only, the rate drops to 10% only. This rate is far more comparable to the infinitive rate present in the current SLI group. The children of the present study are much younger, though, than the children investigated by Rothweiler and colleagues. A higher rate of verb infinitives could therefore be expected in comparison to both studies illustrated above. Hence, the earlier results question the representativeness of the current SLI group.

The data of the children with SLI were not collected by the authors themselves, as both Rice et al. (1997) and Rothweiler et al. (2012) report. In fact, both studies rely on data corpora providing spontaneous language data collected more than two decades ago. Rothweiler and colleagues used data of seven children first published in 1992 (Clahsen & Rothweiler, 1992). Rice and colleagues retrieved their data from a data corpus first presented in 1990 and 1993 (Grimm, 1993; Grimm & Weinert, 1990). The inclusion criteria for these studies varied considerably from those applied in the present study. While the children with SLI of the present study were selected on the basis of a norm-referenced language test covering different language areas, the children participating in the previous studies were selected on the basis of language information focussing on syntactic and morphological skills in particular. Clahsen and Rothweiler selected only children who were diagnosed by clinicians to have a syntactic and morphological impairment. Grimm and Weinert selected eight children whose parents reported that the children did not produce 2-word utterances before the age of 2;6 ($n$ = 4) or 3;0 ($n$ = 4). Given the fact that both studies refer to children with major difficulties in grammar, it seems less surprising that both groups of children show high percentage rates of nonfinite verbs. The second infinitive rate reported by Rice and colleagues, one year after initial assessment, indicates further that the children participating in their study were especially delayed in their acquisition of verb finiteness but that they caught up considerably. Moreover, the characteristic of using verb infinitives rather than finite verbs does not appear stable over time for children with SLI. The results of the current study seem therefore representative for a group of children with SLI who present a more general language impairment than those children reported in previous research. It is likely that some children of the present SLI group remained still at the stage of verb infinitives while a number of other participants had already moved on. This is further supported by a second strength of the present study: the results of the current study are based on two different measures, spontaneous language and probe data. Although the percentages vary across the measures, they sketch nevertheless a coherent picture of the children's error patterns. It can therefore be concluded that German-speaking children with SLI are likely to show a larger percentage of verb infinitives than typically developing children. Nevertheless, this symptom does not seem robust enough within the whole population of children with SLI and neither over time to use it as clinical marker.

**Verb stem**

Reports on the frequency of verb stems in the language of German-speaking children with SLI vary greatly. While Roberts and Leonard (1997) presented this as the most frequent error pattern, Rice et al. (1997) found only 6% of the errors representing verb stems. One of the major methodological differences between both studies is the fact that Rice and colleagues accounted for phonological errors in their count of verb stems. The present study is thus more in line with the study by Rice et al.. In spontaneous language, the children with SLI of the current study produced on average a rate of 17% of phonological errors in 3rd person singular contexts. This highlights that the exclusion of such errors can have a major impact on the results.

With all phonological errors excluded from the count of verb stems in the present study, too, the children with SLI presented the highest rate of bare stems in elicited 3rd person singular contexts (14%). However, this was still comparable to the rate of verb stems observed in the ND-L group. In spontaneous 3rd person singular, slight differences appeared between the SLI (8%) and typically developing children (both groups 3%). Across the groups, verb stems did not occur in 3rd person plural and were rare in 2nd person singular. It is likely that the surface pattern of the inflections can account for this difference between different verb inflections. The 3rd person singular inflection -*t* is short and more difficult to perceive in connected speech than the syllabic inflection *-en.* Also the 2nd person singular inflection seemed easier for all children, including the children with SLI. This may be due to the fricative -*s* in the inflection. This is supported by the fact that a number of responses to the 2nd person singular task comprised the fricative only but not the phonetically correct version of the inflection -*st*. Consequently, the omission of verb inflections seems to be affected by the surface pattern of the target inflection. Data indicates further that the production of the verb stem is a developmental error pattern. If individual speech errors were accounted for the frequency of verb stems in elicited production of the SLI group reflected that of the ND-L group. Thus, the occurrence of verb stems in the language of children with SLI is therefore not suitable as clinical marker for SLI.

**Incorrect inflection**

The typically developing children of the present study produced hardly any incorrect verb inflection. This is in line with data on the acquisition of German. Usually, German-speaking children demonstrate a very high proficiency in subject verb agreement by the age of 3 years and show very few examples of overgeneralisations of inflections (s. Christina Kauschke, 2012 for overview). Consequently, the occurrence of such errors could have some potential for a marker of SLI. However, even the children with SLI did not produce many wrong verb inflections either. This is in line with findings of several previous studies (Ott & Höhle, 2013; Rice et al., 1997; Roberts & Leonard, 1997). Clahsen et al. (1997), on the other hand, reported that about one third of errors in their SLI group were incorrect inflections. Even if the number of incorrect inflections is related to the total of contexts for finite verbs in Clahsen et al.’s data, the SLI group presented still an average rate of 7.6% incorrect inflections. This rate seems clearly higher than that found in the SLI group of the present study.

The error analysis of the present study allows comparisons across different grammatical contexts. While incorrect inflections occurred barely in the context of 3rd person singular across both measures, the average rates for incorrect inflections were slightly higher in the context of elicited 2nd person singular and of spontaneous 3rd person plural, 5% and 8% respectively. These rates get slightly closer to the rate reported by Clahsen et al. although the total average of the present study would lie much lower due to the high proficiency in 3rd person singular – the inflection that occurred also the most frequently. The current findings illustrate that the rate of incorrect inflections seems highly dependent on the grammatical context regarding person and number. This may account in parts for the conflicting findings reported above.

Third person singular is one of the earliest acquired verb inflections in German; second person singular on the other hand is the last inflection acquired in typical German (Clahsen, 1986; Christina Kauschke, 2012; Schrey-Dern, 2006). It is quite probable that the age of acquisition has an influence on the vulnerability of the respective verb inflection. The latest acquired verb inflections offer the context to overgeneralise earlier acquired inflections temporarily. Kauschke (2012) reports indeed that the 2nd person singular inflection *-st* is usually not overgeneralised in typical development while overgeneralisations occur for the earlier inflections. Incorrect inflections in the context of 3rd person singular, on the other hand, seem unlikely. If this inflection is the first to acquire (Clahsen, 1986) there is little choice of other inflections that could be overgeneralised to the context of 3rd person singular. However, the findings for 3rd person plural remain surprising. This inflection has the same surface pattern as the infinitive ending *-en*. The ending *-en* emerges usually before or as early as the 3rd person singular ending (Clahsen, 1986; Schrey-Dern, 2006). In order to differentiate between infinitive and intended 3rd person plural, the verb position was considered in the present study, too. In developmental charts, though, such as the ASAS (Aachener Screeningverfahren zur Analyse von Spontansprache) by Schrey-Dern (2006), the verb position is not considered. Ambiguities can consequently not be ruled out and any information derived from such charts refers rather to the emergence of the verb ending *-en* than to the age of acquisition of 3rd person plural. The ending *-en* can reflect the verb infinitive, 1st or 3rd person plural (see Table 9). In fact, single case studies of children acquiring standard or Austrian German suggest that plural inflections are acquired later than singular inflections and that the 3rd person plural inflection can therefore be considered as an inflection acquired rather late (Bittner, 2003; Klampfer, 2003; Köhler & Bruyère, 1996). The higher rate of incorrect inflections in the context of 3rd person plural is therefore no longer surprising.

In conclusion, the incorrect inflections observed in the present study reflect a developmental pattern. Incorrect inflections appeared only in the context of late acquired verb inflections, independent of the measure tool. Although they barely occurred in the data derived from typically developing children, these errors seem to persist somewhat longer in German-speaking children with SLI. The occurrence of incorrect verb inflections above the age of 3 should therefore be reason enough for a detailed language assessment. Nevertheless, the low frequency of these errors makes this error pattern an inappropriate clinical marker.

**Phonological errors**

Phonological errors of noteworthy frequency occurred within the SLI group in spontaneous language only. In spontaneous production, two verb inflections were analysed: 3rd person singular *-t* and 3rd person singular *-en*. While no phonological errors occurred for 3rd person plural, on average 17% of the contexts for 3rd person singular were considered to result from the children's phonological impairment. In comparison, the ND-A group showed a phonological error rate of only 3% and the ND-L group an average rate of 10%. As the rate is even higher for the PI group (21%), a high rate of phonological errors cannot be considered as suitable clinical marker for SLI. Moreover, the category *phonological error* was excluded from between group comparisons since these errors are likely to result from the children's speech rather than language problems. By definition, thus, phonological errors should be excluded rather than included in discrimination analyses. Otherwise children would be identified on grounds of their speech impairment as experiencing SLI (see chapter 2 for discussion of the differential diagnoses PI / SLI).

An increased rate of phonological errors in the spontaneous data as opposed to the elicited data was observed across the groups, not only within the SLI group. Possibly, the children's utterance length may have had an impact as the elicited responses consisted of a maximum of two words whereas the spontaneous utterances were likely to be as long or longer and demanded therefore a higher processing capacity. It is furthermore little surprising that the children with SLI did not show any phonological errors in the production of 3rd person plural. This inflection is syllabic, and therefore less vulnerable according to the surface account, and none of the participants showed phonological processes that were relevant in the production of word final *-en*.

The high rate of phonological errors for 3rd person singular in the SLI group, on the other hand, emphasises the importance to consider this error pattern carefully in data analysis. If these errors had not been categorised as such they would have been entered into data analyses as the realisation of the verb stem only. As such they would have been considered as grammatical errors and the difference between the SLI group and the typically developing children would have been larger than in the present data analysis. Consequently, a clinical marker may have been found in the production of the 3rd person singular inflection because a large proportion of the errors by the SLI group would have been misinterpreted. The qualitative data analysis of the spontaneous data demonstrates therefore that the exclusion of phonological errors cannot be neglected and may have a major effect on the outcome of the study.

**No attempt**

For the analysis of the morphology probes, one of the categories of response patterns was *no attempt.* While the rate of no attempts was very low for the 3rd person singular task across all groups, the 2nd person singular task caused more difficulties, especially in the SLI group. On average, the children with SLI did not respond to about one fourth of the items of the 2nd person singular task. Hence, the inclusion of these zero responses would probably have had a considerable affect on the results. The rate of no attempts increased also for the ND-L and PI group in 2nd person singular relative to 3rd person singular, although to a lesser extent. This is probably due to the task design. The elicitation of 2nd person singular is pragmatically more difficult than the elicitation of 3rd person singular. Children with SLI may find this especially challenging as their development of pragmatic skills can be affected by the language difficulties, too (Bishop, 2000). Moreover, 2nd person singular is one of the latest acquired verb inflections. This was probably more relevant for the SLI group than for the remaining groups and is further reflected in the very limited data set for this inflection that could be obtained from the spontaneous language samples of the SLI group. Although no attempts should always result in alternative language assessment, this response pattern is very vague. It is therefore not suitable as clinical marker for SLI.

## 5.7 Conclusion

The evidence presented in this chapter does not suggest a suitable marker in present tense inflections for German SLI. The children did present a significant delay within delay pattern in a present tense composite including the copula *sein* but the children's production was characterised by a very high accuracy rate. Regarding individual inflections, some children showed marked difficulties but these were not typical for the whole group. Further research may identify that these error patterns are a developmental characteristic for children with SLI, but it is unlikely that present tense marking can be used as marker for SLI that is stable over time. The lower frequency of errors in verb inflections presented in this study relative to English-speaking children with SLI is in line with the morphological richness account. It is likely that the higher number of different verb inflections and their semantic importance can explain the higher performance accuracy of German-speaking children.

Longitudinal research would further be useful in order to investigate the role of verb infinitives. The current study could not confirm that this is a robust hallmark for German SLI. However, there may be a timeframe during which this is an error pattern distinct from the language of typically developing children. The relatively high frequency of phonological errors highlights the importance to treat these with caution, especially with regards to verb stems. The occurrence of incorrect verb inflections in children aged 3 or older should raise the clinician's awareness and should result in detailed language diagnostics. However, no error pattern could be identified that has potential to discriminate children with SLI reliably from their typically developing peers.

# 6 Case Marking in German SLI

## 6.1. Background

Case marking in German appears in nouns, determiners, adjectives and pronouns. Schöler, Fromm and Kany (Schöler et al., 1998b) suggest that the morphology of German-speaking children with SLI is especially characterised by weaknesses in case marking.

## 6.2. Case Marking in German Noun Phrase

The noun phrase in German consists in general of a determiner, an adjective (optional) and the noun. Figure 11 illustrates these constituents. The determiner is usually obligatory apart from a few exceptions[13].

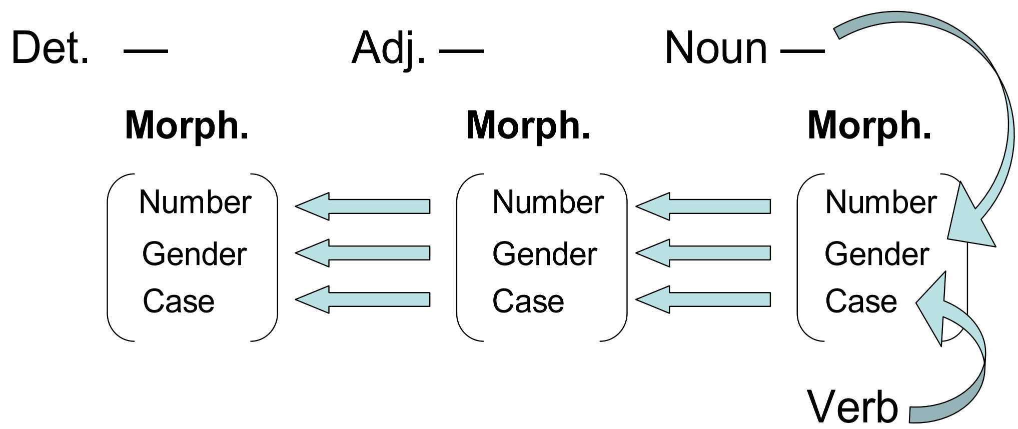



**Figure 11.** German noun phrase

[13] e.g. in noun phrases with an indefinite plural noun or a noun that refers to an uncountable mass (e.g. *Obst* (=fruit))

The noun phrase can be regarded as a single unit and it is marked for number, gender and case as indicated in Figure 11. German nouns take one of three genders: masculine, feminine or neuter. Figure 11 illustrates how the noun defines the number and gender of the noun phrase and the verb determines the case. Gender marking is always co-occurring with case marking in German. While the present study focuses on case marking, gender marking is therefore also included.

All constituents of the noun phrase are inflected into four grammatical cases in German: nominative, accusative, dative and genitive. The case of a particular noun phrase is defined by the verbs argument structure or a preposition and depends on the grammatical function of the noun phrase within the sentence.

- Nominative: the subject of a sentence
- Accusative: the direct object of a sentence or the object following certain prepositions
- Dative: the indirect object of a sentence or the object following certain prepositions
- Genitive: the possessor of something or the object following certain prepositions

All constituents of the noun phrase can be inflected by the suffixation of a morpheme according to number and case; overt marking for gender appears only in the determiner and the suffix of the adjective.

Table 20 gives an overview of definite and indefinite articles in German. The noun *Hund* (English: *dog*), for example, is masculine and requires in nominative case *der* as definite article whereas this article changes to *den* in accusative or *dem* in dative. The definite article for plural nouns changes only across cases but not across genders. It should be pointed out, however, that genitive is often replaced by dative in colloquial language (Sick, 2004). Furthermore, consideration of all possible forms of German illustrates that they are not exclusive in their grammatical meaning. For example, *der* can refer to a masculine noun in nominative or alternatively to a feminine noun in dative; *ein* can precede a masculine noun in nominative but also a neuter noun in nominative or accusative.

**Table 20.** Definite and Indefinite Articles of German

| | **Singular** | | | | | | **Plural** |
|---|---|---|---|---|---|---|---|
| | **Masculine** | | **Feminine** | | **Neuter** | | |
| **Case** | **Def.** | **Indef.** | **Def.** | **Indef.** | **Def.** | **Indef.** | **Def.** |
| **Nominative** | der | ein | die | eine | das | ein | die |
| **Accusative** | den | einen | die | eine | das | ein | die |
| **Dative** | dem | einem | der | einer | dem | einem | den |
| **Genitive** | des | eines | der | einer | des | eines | der |

*Note*. Def. = definite; Indef. = indefinite

The declension of adjectives is dependent on number, gender and case of the noun which the adjective refers to. Additionally, the declension of adjectives varies according to the preceding determiner. Adjectives that are preceded by a definite article follow a weak declension and receive the suffixes *-e* or *-en* only. Adjectives that are preceded by no article or by an indefinite article, on the other hand, need to carry information on all three grammatical features number, gender and case (i.e. strong / mixed declension). The strong declension results in the suffixation of *-e, -en, -er, -es,* or *-em.* Example (a) gives an instance of a noun phrase with a definite article. Example (b) illustrates the agreement between adjective and noun regarding number, gender and case if the article is indefinite.

a. das schwarz**e** Schaf
b. ein schwarz**es** Schaf

Nouns are inflected according to number and case. Nouns undergo case marking in particular phonotactic contexts through suffixes being attached to the noun (e.g. nominative: die Hunde (plural); accusative: den Hunde**n**).

## 6.3. Case Marking in German Pronouns

Sentence constituents can be represented by pronouns. In these cases, the pronouns reflect the same gender, number and case as the noun phrase they refer to. Thus, personal as well as possessive pronouns differ according to number, case and gender. Table 21 provides an overview for personal pronouns in nominative, accusative and dative.

**Table 21.** Personal Pronouns in German

| | **Person** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Singular** | | | | | **Plural** | | | **Formal** [14] |
| | **1st** | **2nd** | | **3rd** | | **1st** | **2nd** | **3rd** | |
| **Nominative** | ich | du | er | sie | es | wir | ihr | sie | Sie |
| **Accusative** | mich | dich | ihn | sie | es | uns | euch | sie | Sie |
| **Dative** | mir | dir | ihm | ihr | ihm | uns | euch | ihnen | Ihnen |

[14] In formal German, *Sie* is used as second person personal pronoun. The form of formal 2nd person singular and formal 2nd person plural is identical.

## 6.4 Case Marking in German SLI

Kany and Schöler (1998) investigated case marking in children with SLI (age 7;6-12;0 years) across all relevant word classes using a number of elicitation tasks (story telling, cloze procedures and sentence repetition) and analysis of spontaneous language samples. Results indicated that case marking errors are the most dominant error patterns in children with SLI. A story telling task revealed that general case marking errors disappeared in the language of typically developing children around an age of 7 or 8 years. For children with SLI aged 7-8 years, on the other hand, case errors still represented the largest proportion of errors (22%), followed by errors with verb inflections (14%), gender errors (10%), errors with prepositions (8%) and number errors (4%). The case errors disappeared thus more slowly in the spontaneous language of children with SLI. Nevertheless, at an average age of 10, the children failed only in 3.5% of contexts to mark case correctly. Kany and Schöler suggest therefore that case errors are also likely to disappear with increasing age in the spontaneous language of children with SLI.

In a sentence repetition task, however, case errors were the most frequent errors in sentence repetition across the age range studied (7;6-12;0), suggesting that they may remain a ‘fragile’ area of language knowledge for children with SLI which may emerge under high levels of processing demands, such as a complex argument structure. Across word classes, nominative was generally unaffected and thus correctly reproduced. Particularly challenging was the repetition of possessive pronouns in sentences of which only 36% were correctly reproduced. In general, any other form than the default case nominative (i.e. accusative or dative in this task) was challenging for the children. The children with SLI, but not the controls, substituted dative occasionally for accusative forms.

With regards to case marking in adjectives, Kany and Schöler (1998) presented data from two different tasks: cloze test and sentence repetition. In one task, the investigators asked the children to insert missing inflections in a cloze text. Eight of the item sentences involved the insertion of adjective morphology. Kany and Schöler report that the children with SLI predominantly repeated the adjective without the insertion of any inflection. Besides this, the children produced only very few incorrect inflections and no effects were apparent of gender, case, weak or strong declension. However, a few children with SLI overgeneralised the strong dative declension *-em*, an error pattern that was absent in typically developing children. Results from the sentence repetition task, too, show an error rate between 9 and 23% for the production of the correct adjective. This is a substantially higher error rate than for nouns or verbs.

Considering different word classes, most research on case marking in German SLI refers to the production of articles. Kany and Schöler (1998) also investigated case and gender marking in articles in the study described above. In the cloze procedure, children with SLI succeeded in only one third (34%) of the contexts for articles. In contrast, the typically developing controls inserted 79% of the articles correctly. The most common error patterns of the children with SLI were the repetition of the article presented in the task sentence and the insertion of an incorrectly inflected article. Similar to the adjectives, the overgeneralisation of an indefinite dative article in genitive or accusative context was observed in children with SLI only.

The production of articles has also been reported to be vulnerable by other researchers. Clahsen (1989) reported an overall omission rate of 55% in spontaneous data of children with SLI aged between 3;2 and 9;6. Roberts and Leonard (1997) found a comparable omission rate (49%) in spontaneous language samples of eight children aged between 4;8 and 7;2. The findings by Clahsen also confirm the results by Kany and Schöler (1998) regarding inflectional errors for case and gender marking in articles. Clahsen reported difficulties with correct accusative and dative marking (1989, 1991). Most commonly, the children overapplied nominative forms to other cases. However, in contrast to Schöler et al., neither Roberts and Leonard nor Clahsen related their findings to data from typically developing children. Eisenbeiss, Bartke and Clahsen (Eisenbeiss, Bartke, & Clahsen, 2005, 2006), on the other hand, studied case marking in the spontaneous language of five children with SLI (age

5;9 – 7;11years) and compared them to typically developing children matched for MLU. The children with SLI showed a similar pattern to their MLU-matched controls. For both groups, case marking was only difficult if the required case was an exception from the structural case[15] and had to be learned for a particular lexical item only (lexical case). Consequently, Eisenbeiss et al. could not identify a particular weakness in case marking for children with SLI.

These conflicting findings could result from several methodological differences. First of all, only Eisenbeiss and colleagues (2005, 2006) distinguished between structural and lexical case. Secondly, Eisenbeiss et al. investigated the language of older children than Clahsen (1989, 1991) or Roberts and Leonard (1997) (who used data collected earlier by Clahsen and Rothweiler (1992)). It could therefore be possible that a marked deficit is characteristic of younger children with SLI but disappears over time. However, Kany and Schöler (1998) investigated children of similar age as Eisenbeiss et al. and found impaired case marking in children with SLI, too. Only at an age of 10 year, Kany and Schöler identified a large improvement in case marking. It remains unclear whether these long standing difficulties in 68 children tested by authors could solely be triggered by items requiring lexical rather than structural case marking. A third explanation could be the procedure of data collection. While Kany and Schöler used a number of different procedures including sentence repetition and story telling, Eisenbeiss and colleagues analysed spontaneous language recordings of an hour each. The more difficult tasks by Kany and Schöler could have impacted on case marking or could have elicited a greater proportion of lexical case marking. The available data at present could therefore be interpreted as an indication of a particular weakness in case marking in young children with SLI which is overcome around school-age in spontaneous language and later in more structured tasks. Nevertheless, Eisenbeiss and colleagues based their conclusion on data from a rather small sample since they recorded only five children with SLI. Also the data reported by Clahsen,

[15] In structural case marking, the case is assigned in a particular structural position of the sentence (e.g. nominative subjects, accusative direct objects, dative indirect objects). Lexical case marking, on the other hand, is specific to a particular lexical item. This lexeme requires another case than the structural case would predict.

and Roberts and Leonard is based on the language of only 18 children in total. It would consequently be useful to investigate case marking in articles further in young children with SLI and to compare the children's production to language matched control children.

Little data is available on case marking in nouns. Kany and Schöler (1998) report from their cloze test that noun inflections were rarely inserted by children with SLI. However, typically developing children had major difficulties with this task, too. Clahsen (1989,1991) noted that his spontaneous language data did not present any example of obligatory context for case marking in nouns.

Following the purpose of the present study to identify a potential clinical marker for German SLI in 3-6-year-old children, the production of articles seems to be an appropriate starting point and this was tested using elicitation tasks as well as spontaneous language samples. Case marking in adjectives may also present a special weakness but elicitation and insertion procedures such as those used by Schöler and colleagues (1998) are inappropriate for the young children studied here. The review of the literature suggests that case marking in pronouns and nouns is less likely to show promising results in discriminating children with SLI from typically developing children. Nevertheless, the analysis of spontaneous language in this study offers the opportunity to also look for instances of case marking in adjectives, nouns and pronouns additionally to the analysis of articles.

The following investigation therefore addresses the following experimental questions:

1. Do German speaking children with SLI exhibit the pattern of a delay within a delay in their use of case marking in articles when compared to typically developing children and those with a phonological impairment (PI)?
2. What types of errors do children with SLI make and are these qualitatively different from those made by typically developing children and those with PI?

In order to answer these questions, two approaches were used for data collection: elicitation procedures and the recording of spontaneous language samples. Both procedures including data collection, data analysis and results are presented below.

## 6.5 Morphology Probes on Case Marking

### *6.5.1 Participants*

Sixty-two monolingual German-speaking children participated in the experiment on case marking. Thirteen children met the criterion for SLI (SLI group, mean age 4;2), 16 children were age and gender matched and had typically developing language but had a phonological speech output impairment (PI group, mean age 4;1), 17 children were typically language developing and age and gender matched with the SLI group (ND-A group, mean age 4;0), and a further 16 children were typically language developing and matched with the SLI group on gender and language comprehension (ND-L group, mean age 3;3). All groups are further described in the general method in Chapter 4.

### *6.5.2 Elicitation tasks*

Three morphology probe tasks were designed which aimed for eliciting definite articles in German. Table 22 gives the target structures: the definite article of all three genders in nominative, dative and accusative. The elicitation tasks aimed for definite articles as these morphemes are monosyllabic and should be more vulnerable than the partly two-syllabic indefinite article according to the surface account (e.g. Leonard, 1989; see also Leonard, 1998). Two versions of the scoring sheet are attached in Appendix 2 and 3: the German original and its translation to English.

**Table 22.** Case Marking in German Definite Articles

| Gender | Nominative | Dative | Accusative |
|---|---|---|---|
| Masculine | der | dem | den |
| Feminine | die | der | die |
| Neuter | das | dem | das |

**Nominative articles**

The nominative task took place during the introduction of a bingo game. Picture 2 shows the material consisting of a board with eight pictures and eight little cards mirroring the pictures on the board. The examiner named all small cards when presenting them first in order to create the context for definite (rather than indefinite) articles for subsequent naming.

Introducing the game, the examiner pointed to each of the pictures on the board asking: "What goes here?"

This situation represented obligatory context for definite articles because all target items existed only once as card and had been introduced earlier. If the child used the indefinite article the examiner tried to elicit the definite correspondence by modelling the trial examples and then asking "*Which* picture goes here?".

Two pictures served as trial examples and six were scored, two for each gender. The test items represented animals (snail, pig, and elephant), clothes (shoe), furniture (bed) and the sun.

**Picture 2.** Bingo game for the nominative and accusative task of the morphology probes

**Accusative articles**

The bingo game as introduced for the nominative article was used again. A hand puppet called *Fips* took part in the game, and the child was encouraged to ask the puppet for the small picture cards.

Exp.:
*Examiner:* “Now we need to put the little cards onto the board. Fips can give them to us. But we need to tell him what to give. Fips, give me….”
*Child:* “… the moon.” (German: “… den Mond.”)

Two items were training items. In order to elicit these specific items the examiner pointed to the respective picture while giving the prompt. For the other six experimental items the child could choose a picture. The six nouns were distributed evenly across the three grammatical genders. The test items were masculine because this is the only gender involving a definite accusative article that is exclusive to this grammatical context (see Table 22). In this way, the trials were thought to be likely to stimulate accusative. The verb *geben* (to give) that triggered a direct object and therefore accusative in this task is an early acquired verb (Grimm & Doil, 2000). The obligatory context for definite articles was enforced through the uniqueness of the items within the situation.

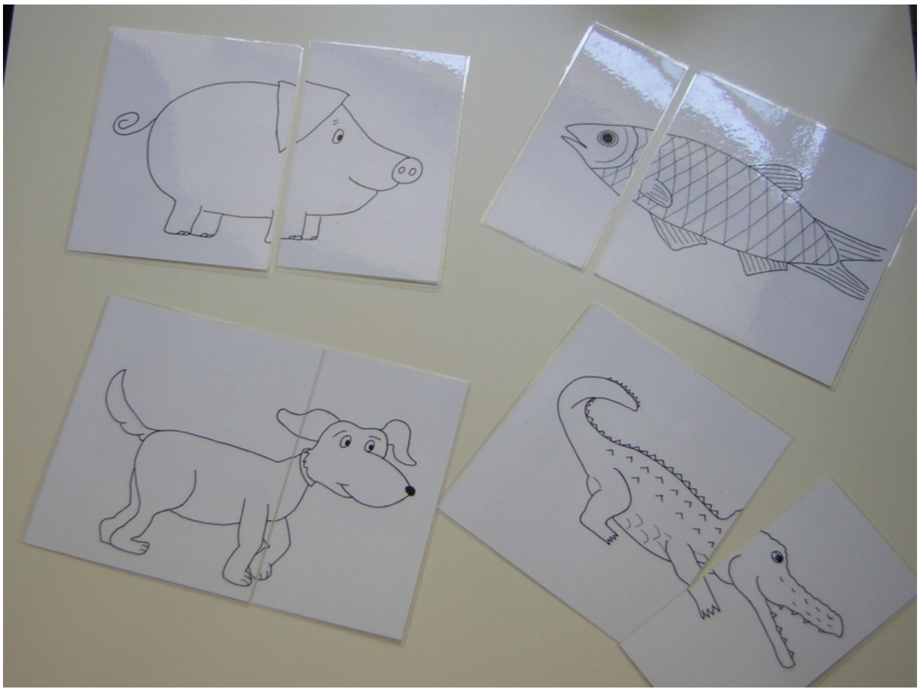
**Dative articles**

Picture cards with drawings of animals (see Picture 3) were presented to the child. Each animal consisted of two pieces: head and body. The child was shown a head and had to select the respective body for this animal from a choice of four bodies. While giving the animal’s head to the child the examiner asked:

“Who does this head belong to? – This head belongs to …?”
The child was supposed to name the appropriate animal with its dative definite article.

Exp.: *Examiner:* “Wem gehört denn der Kopf? – Der Kopf gehört …?“
*Child:* „…dem Krokodil.“ (Engl.: ‘… to the crocodile’.)

The verb *gehören* requires a dative object (lexical case). Although they designed a different task, Kauschke and Siegmüller (2002) also use the context of *gehören* in order to elicit definite articles in their language test battery *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen*. This test battery had not been standardised at the time of data collection for the present study, but normative data has been presented in a second edition in 2009 for children aged from 2;0 – 6;11.



**Picture 3.** Picture cards for the dative task of the morphology probes

Two items were considered as trial whereas six items were scored. The items that were scored were evenly distributed across genders with two masculine, two feminine and two neuter nouns. The two trial items were a masculine and a neuter noun, both resulting in the definite article which does not appear in other grammatical contexts. The feminine dative article has the same surface structure as the masculine nominative article and would therefore have been an ambiguous trial item. Four of the scored items represented pets or farm animals (cat, dog, sheep, and cow), one item was a toy (teddy bear) and one a wild animal (crocodile).

### *6.5.3 Documentation*

The procedure was audiorecorded using a SONY minidisk recorder MZ-R909 and online notes were taken. If the children produced the target structure the article was marked as correct; if the child gave a different answer to the target this was transcribed on-line. Section 4.8 in the methods chapter gives interjudge agreement values for 10.6% of the transcripts.

### *6.5.4 Quantitive analysis*

#### *6.5.4.1 Scoring*

For all three cases, percentages of correct use were computed by dividing the number of correct productions by the number of obligatory context for each case. The following outlines the approach applied in the decision whether a response was categorised as correct or incorrect:

- Productions of the indefinite rather than the targeted definite article were scored as correct if they were correctly marked for case. This was decided because the choice of the indefinite article in the context of a definite article represents a pragmatic issue, but not a deficit in case marking.
- In dative, articles were in few cases combined with the preposition *zu* (English: *to*) as explained in Table 23. These productions were scored as correct because the grammatical contrast was marked.
- The omission of an article, gender errors as well as the use of a filler word was scored as incorrect.
- Zero responses were excluded from the number of obligatory contexts. It was in these cases ambiguous whether the child did not try to produce the target item (possibly due to task comprehension or cooperation) or whether the child was not able to respond due to his linguistic limitations.
- Errors were categorised as grammatical or phonological error and phonological errors were excluded from analysis and the number of obligatory contexts. This categorisation was based on the individual speech output of each participant and was relevant especially for two error patterns:

- Production of [də]: If the child produced this as filler word across genders and cases this error was scored as incorrect. However, this error was considered as phonological error and excluded from analysis if, firstly, this error occurred only in contexts in which word final consonant deletion would result in such a response (i.e. exclusively *dem* and *den*) and secondly, if the child showed word final consonant deletion involving /m/ and /n/ in a minimum of two other speech contexts, too.
- Masculine dative article *dem* realised as *den*: One common error in the production of the masculine dative article *dem* is the overgeneralisation of the accusative article *den.* This could result from a phonological difficulty signalling the contrast between word final [m] and [n]. If children showed this error pattern the data was checked for two indicators whether this was a grammatical or a phonological error. Firstly, the child's data from all assessments was checked for a minimum of two instances of phonological processes involving the word final contrast [m] - [n]. Such errors in the child's speech output would cause immediate classification as phonological error and data would have been excluded from quantitative analysis. Secondly, all responses of the dative task were considered involving the articles of other genders with greater phonological salience. If the children showed errors in the feminine and/or neuter dative articles, too, it was assumed that the children had grammatical rather than phonological limitations causing the overgeneralisation of the accusative article in masculine dative articles. Thus, these errors were scored as incorrect and remained in the data set for quantitative analysis. If no examples of phonological errors were found *and* the child produced *dem* → *den* substitutions as only errors in dative, these remained included in the analysis set. This error has been reported as one of the most frequent error patterns in typically developing children (Szagun, 2004).

**Table 23.** Compound Forms of Preposition *zu* (English: *to*) and Definite Dative Article

| Gender | Preposition and article | Compound form |
|---|---|---|
| feminine | zu der | zur |
| masculine | zu dem | zum |
| neuter | zu dem | zum |

Therefore, the percentage of production accuracy was calculated for each case as follows:

$$\frac{\text{Correct productions}}{\text{Total of correct productions + grammatical errors}}$$

### *6.5.4.2 Statistical analysis*

**Dependant variables**

The four groups of children were compared in terms of their accuracy in case marking in articles. Three variables were compared: (a) the production accuracy of nominative articles*,* (b) the production accuracy of accusative articles*,* (c) the production accuracy of dative articles, and (d) a case composite combining the production accuracy for all three cases. Following the scoring guidelines above, Table 24 illustrates the number of items that were included and excluded in quantitative analysis for between-group comparison.

**Data preparation and analysis**

Arcsine transformations were applied to the percentage data and analysed through analysis of variance (ANOVA). The assumption of homogeneity was checked with Levene's test. Accordingly, ANOVA was followed by Gabriel's procedure if variance was equal across groups or by Games-Howell post hoc test if the assumption of homogeneity of variance homogeneity was violated. If significant between group differences were found between the SLI group and the ND-A or ND-L group the variable was further explored as clinical marker by calculating sensitivity, specificity, LR+ and LR-.

The results of the children with SLI were transformed to *t*-values and percentiles relative to the children's age (i.e. ND-A data) and, secondly, relative to the children's language development (i.e. ND-L data).

Furthermore, Pearson's correlation was computed in order to investigate post-hoc a relationship between the children's age in months and their production accuracy of elicited nominative articles. For this, the arcsine transformed data was used for nominative articles.

**Table 24.** Number of Items for Elicited Case Marking Entering Qualitative Analysis

| | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|
| Total no participants | 13 | 16 | 17 | 16 |
| Total no of items per group and per grammatical case | 78 | 96 | 102 | 96 |
| No of items excluded due to task refusal | | | | |
| nominative | 2 | - | - | 3 |
| accusative | - | 1 | - | 6 |
| dative | - | - | - | 2 |
| No of items excluded due to phonology | - | - | - | - |
| No of items for analysis (a) | 76 | 96 | 102 | 93 |
| No of items for analysis (b) | 78 | 95 | 102 | 90 |
| No of items for analysis (c) | 78 | 96 | 102 | 94 |
| No of items for analysis (d) | 232 | 287 | 306 | 277 |

*Note.* Number of items for analysis (d) is (a+b+c)

### *6.5.5 Results of quantitative analysis*

Case marking was first examined in elicited definite articles in nominative, accusative and dative. In nominative and accusative, some children produced the indefinite instead of the definite article. Both clinical groups showed more instances of indefinite articles than the control groups with typically developing children. In nominative, six children of each clinical group produced one or more indefinite articles compared to only three children of the ND-A group and two children of the ND-L group. In accusative, one child of the SLI group, five children of the PI group, and two children of the ND-L group produced indefinite articles. No child of any group used indefinite articles in dative.

Regarding phonological errors, no data was excluded from analysis. None of the phonological processes presented by the children would directly have resulted in any of the observed error patterns. The substitution of the masculine accusative article *den* for the masculine as well as the neuter dative article *dem* occurred 12 times in the SLI group, 14 times in the PI-group, 21 times in the ND-A group and 15 times in the ND-L group. Of the 21 children who showed this error pattern, five children presented phonological processes which impact on word-final consonants such as inconsistent word-final consonant deletion. However, these children presented additional errors in dative marking which did not involve the distinction between word-final [m] and [n]. It was therefore assumed that the children's errors did not result directly from the children's speech output, especially since no child demonstrated substitution errors involving [m] and [n] in their speech. Table 25 presents a summary of group data.

**Table 25.** Elicitation Tasks: Proportional Scores of Correct Use of Case Marking in Articles

| *Case* | | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|---|
| **Nominative** | M | 0.73 | 0.79 | 0.94 | 0.95 |
| | SD | 0.35 | 0.32 | 0.17 | 0.10 |
| | Range | 0.0 – 1 | 0.0 – 1 | 0.34 – 1 | 0.67 – 1 |
| | n | 13 | 16 | 17 | 16 |
| **Accusative** | M | $0.60_a$ | 0.74 | $0.91_a$ | 0.74 |
| | SD | 0.33 | 0.30 | 0.15 | 0.30 |
| | Range | 0.0 – 1 | 0.17 – 1 | 0.5 – 1 | 0.0 – 1 |
| | n | 13 | 16 | 17 | 16 |
| **Dative** | M | $0.08_a$ | 0.13 | $0.46_{a,b}$ | $0.14_b$ |
| | SD | 0.16 | 0.25 | 0.38 | 0.19 |
| | Range | 0.0 -0.5 | 0.0 -1 | 0.0 – 1 | 0.0 – 0.5 |
| | n | 13 | 16 | 17 | 16 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

When arcsine transformations were performed on the proportional scores for case marking, the homogeneity of variance assumption was nevertheless violated in nominative and dative. Brown-Forsythe F-ratio was therefore considered for both variables and Games-Howell procedure was chosen as post hoc procedure. For accusative, Gabriel's procedure was chosen instead as a post-hoc test. Main group effects were evident for accusative and dative articles, but not for nominative. The main group effect in accusative ($F$ (3, 30.72) = 4.65, $p$ = .009, $\omega^2$ = .12) can be explained by the fact that the children with SLI showed significantly more difficulties in producing correct accusative articles than the ND-A group ($p$ = .010). The main group effect in the production of dative articles ($F$ (3, 34.66) = 6.74, $p$ = .001, $\omega^2$ = .20) resulted from significant higher proportional scores in the ND-A group in comparison to both the SLI-group ($p$ = .013) and the ND-L group ($p$ = .031). The results for all three cases are illustrated in Figure 12.

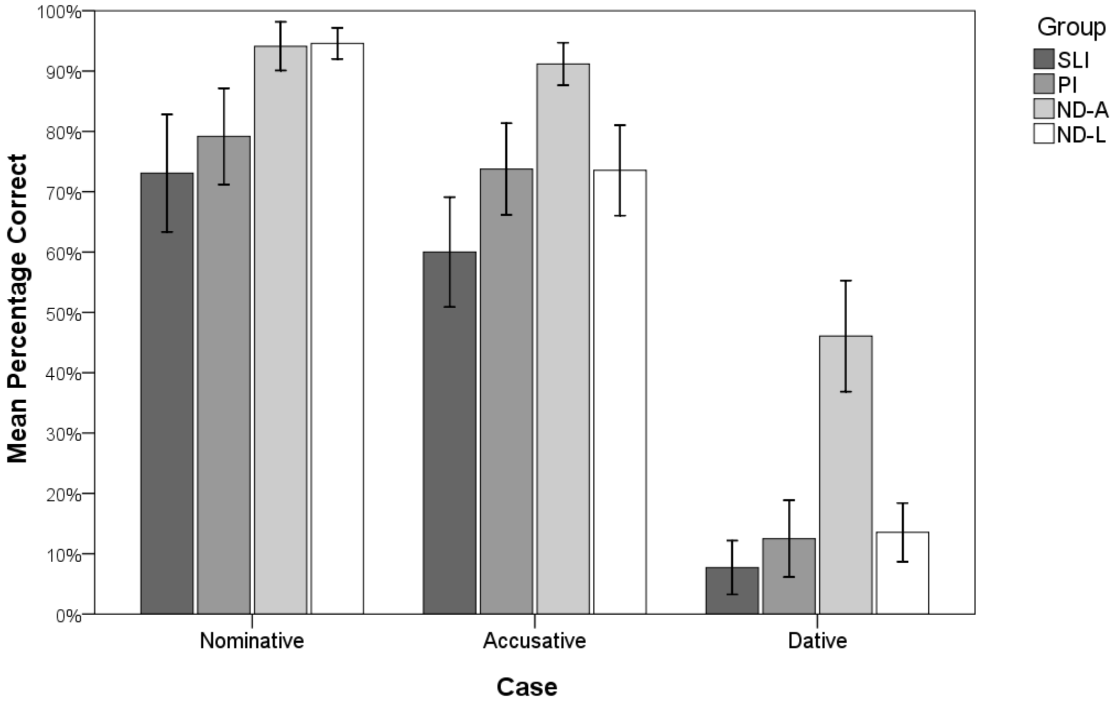



**Figure 12.** Mean percentages of use (± 1 SE) for nominative, accusative and dative case marking in elicited articles for the four groups of children.

If using the production accuracy of elicited accusative articles as diagnostic tool, a cut-off of 80% was the most suitable cut-off as the overview over the data in Table 46 demonstrates. This cut-off resulted in an overall discrimination accuracy of 80%, a sensitivity of 0.692 [0.424 - 0.873], a specificity of 0.882 [0.657 - 0.967], LR+ = 5.885 [1.524 - 22.73] and LR- = 0.349 [0.152 - 0.803]. For dative articles, the most suitable cut-off was at 0% correct dative productions as in Table 47 presented. Eighty percent of the children could be classified correctly as SLI or ND-A respectively. The sensitivity was 0.769 [0.497 - 0.918], specificity was 0.824 [0.59 - 0.938], LR+ = 4.359 [1.496 - 12.698] and LR- = 0.28 [0.101 - 0.774].

Percentiles relative to the ND-A group are given in Appendix 10.

Percentiles of the performance of the SLI group relative to the ND-L group are presented in Appendix 11.

The analysis was carried further by correlating the children's age in months with their percentage correct of elicited nominative articles. There was a trend towards a medium positive relationship between age and accuracy in nominative articles, $r = .226$, $p$ (one-tailed) = .034.

**Case Composite**

A composite score for case marking in articles was derived from the elicitation tasks by combining the data of all three cases. For this, all articles marking case correctly were taken together across nominative, accusative and dative and related to the total number of items requiring such case marking. The mean of correct case marking in articles, standard deviation, range and number of participants are given in Table 26.

**Table 26.** Elicitation Tasks: Compositional Scores Case

| ***Compositional score*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Case composite** | **M** | $0.47_{a}$ | $0.55_{b}$ | $0.77_{a,b,c}$ | $0.62_{c}$ |
| | **SD** | 0.20 | 0.20 | 0.18 | 0.13 |
| | **Range** | 0 – 0.78 | 0.06 – 0.78 | 0.39 – 1 | 0.29 – 0.82 |
| | **n** | 13 | 16 | 17 | 16 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

All proportional scores were transformed to arcsine values for between-group comparisons. Nevertheless, Levene's test indicated that the assumption of homogeneity was violated for the case composite. Brown-Forsythe $F$-ratio showed a main group effect ($F$ (3, 29.51) = 2.90; $p$ = .047, $\omega^2$ = .29) and Games-Howell post hoc test revealed that the ND-A group performed significantly better on case marking in general than the SLI group ($p$ = .001), the PI group ($p$ = .007) and the ND-L group ($p$ = .034). Figure 13 illustrates the mean proportional scores for the case composite across the four participant groups.

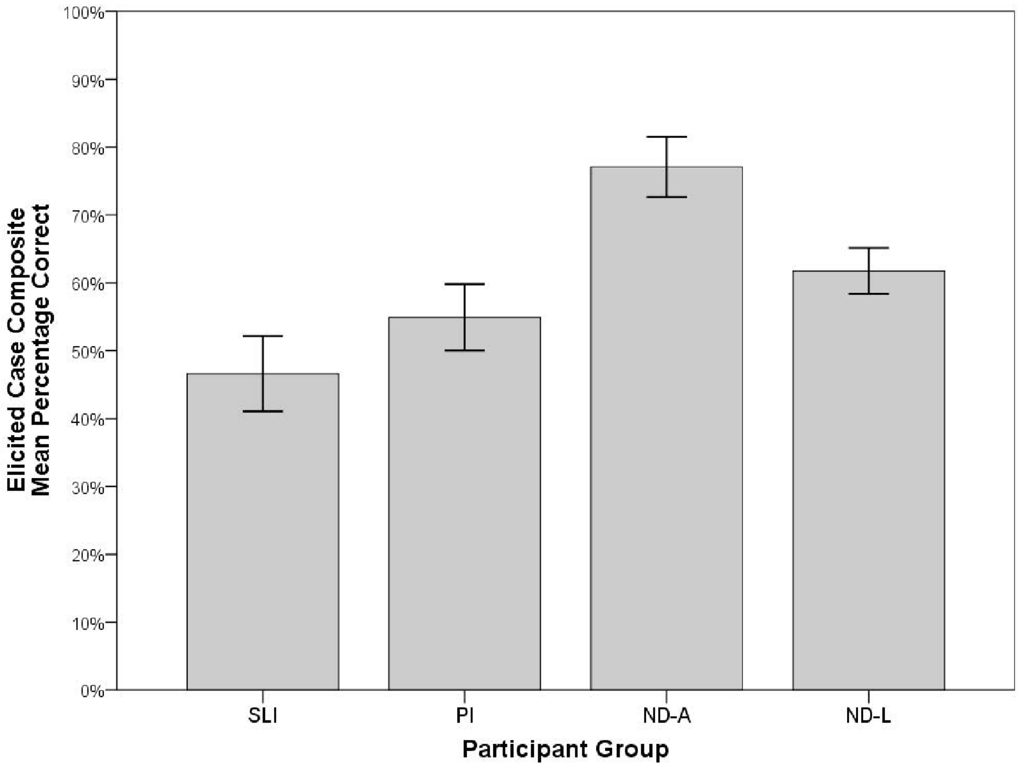



**Figure 13.** Mean percentages of use (± 1 SE) for case composite from elicited production.

Table 48 indicates 59% as the most suitable cut-off for the case composite for elicited articles. If this cut-off was applied, 83% of the children were correctly classified and a sensitivity of 0.769 [0.497 - 0.918] was obtained, a specificity of 0.882 [0.657 to 0.967], LR+ = 6.538 [1.72 to 24.857] and LR- = 0.262 [0.095 to 0.716].

Percentiles relative to the ND-A group are given in Appendix 10.
Percentiles of the performance of the SLI group relative to the ND-L group are presented in Appendix 11.

### 6.5.6 Qualitative analysis

**Data coding**

In order to reveal trends in error patterns as reported in previous research (e.g. Clahsen, 1989, 1991; Schöler et al., 1998), a descriptive analysis of the response pattern for case marking in definite articles was carried out. For this, responses were coded into seven categories of production patterns for nominative and accusative:

- Correct case marking
- Indefinite instead of definite article
- Case error
- Gender error
- Filler word (e.g. [də] across cases and / or genders).
- Omission of the article
- No attempt

In dative articles, three additional error types were identified:

- Gender *and* case error: e.g. *die Schaf* [feminine + nominative / accusative] instead of *dem Schaf* [neuter + dative]
- Overgeneralisation accusative in feminine and neuter articles: If the child produced accusative instead of dative across *a* minimum of three items this was treated as overgeneralisation.
- Overgeneralisation of the masculine accusative article *den* in masculine, feminine and / or neuter dative articles

**Data included**

For qualitative analysis, the data of all 62 children was included, including zero responses (category *no attempt*). This category was added to the qualitative analysis in order to account for the fact that children with very poor language skills may not have been able to attempt the task. This approach could give an indication in how far the inclusion or exclusion of this data impacted on the outcome of the between group comparison. The category *phonological error* does not appear as the children did not make any errors resulting from a speech output disorder.

**Data analysis**

The total number of items per group and per grammatical case is given in Table 24. The proportion of each response type was calculated per grammatical case.

### 6.5.7 Results of qualitative analysis

**Error patterns in case marking**

Figure 14 presents an overview over the response patterns across the groups for nominative articles, Figure 15 illustrates the response patterns of accusative articles and Figure 16 presents the response patterns for dative articles. Table 27 provides the percentages corresponding to the pie charts.

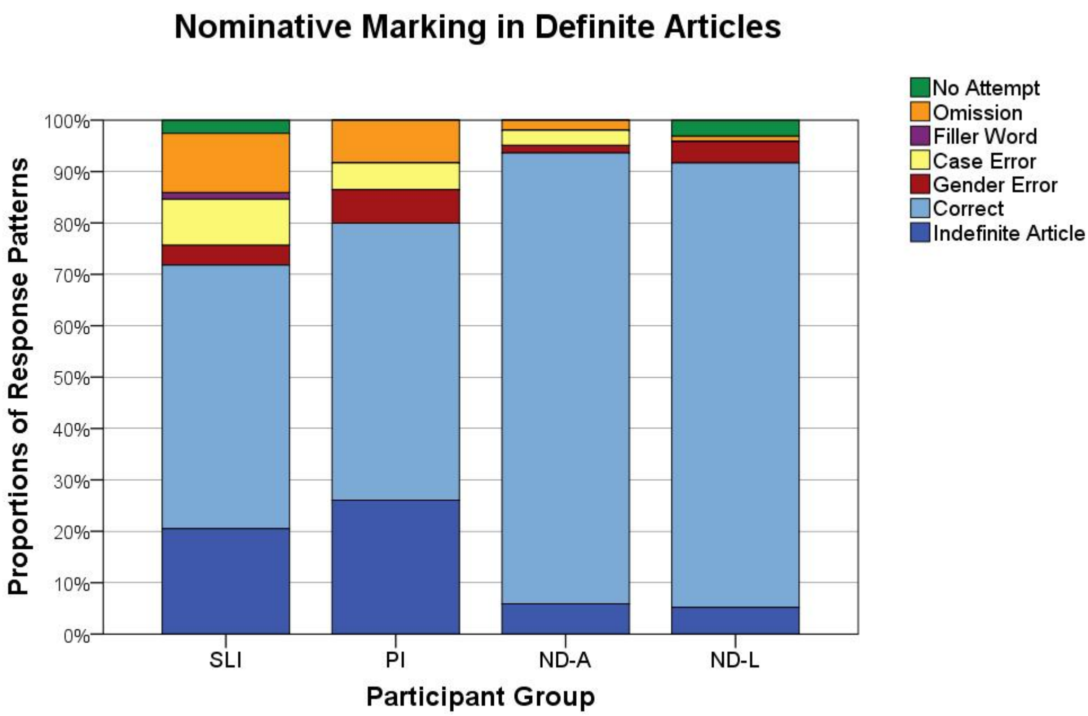



**Figure 14.** Production patterns for elicited definite articles in nominative.

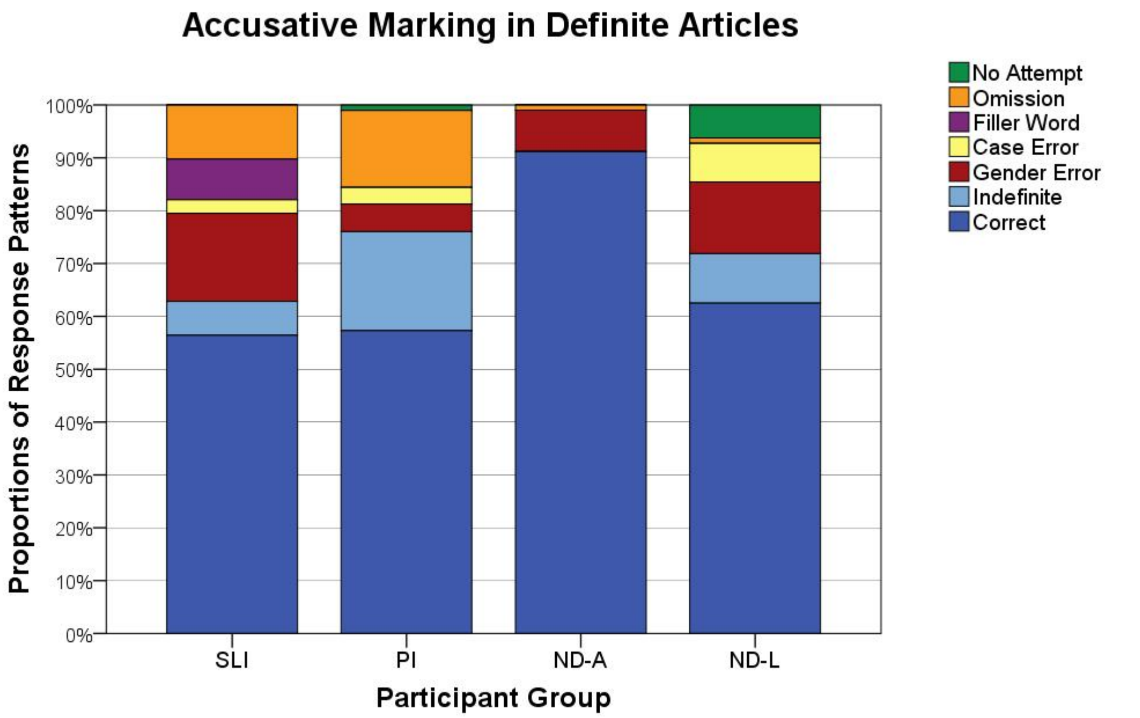



**Figure 15.** Production patterns for elicited definite articles in accusative.

No statistical between group comparisons were performed regarding error patterns in case marking. The available data per category was too limited.

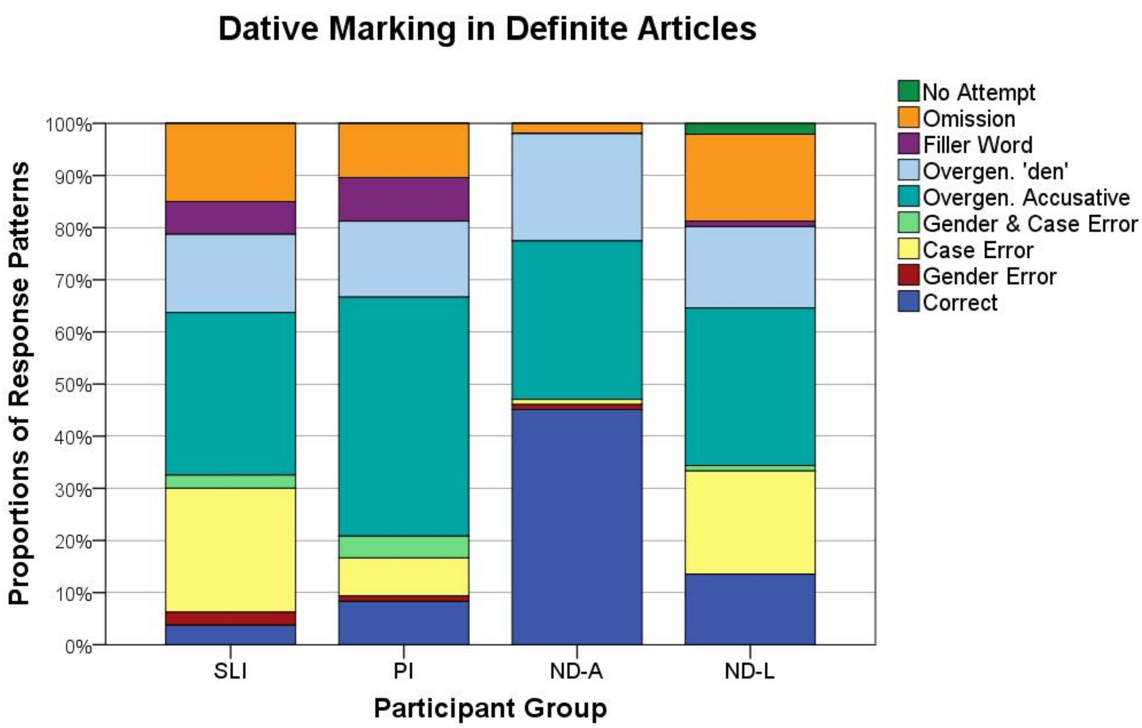



**Figure 16.** Production patterns for elicited definite articles in dative.

**Table 27.** Elicitation Tasks: Production Patterns for Case Marking in Definite Articles

| | ***Production Pattern*** | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative** | **Definite article correct** | 51% | 54% | 88% | 86% |
| | **Indefinite article** | 21% | 26% | 6% | 5% |
| | **Gender error** | 4% | 7% | 1% | 4% |
| | **Case error** | 9% | 5% | 3% | 0% |
| | **Filler word** | 1% | 0% | 0% | 0% |
| | **Omission article** | 12% | 8% | 2% | 1% |
| | **No attempt** | 3% | 0% | 0% | 3% |
| **Accusative** | **Definite article correct** | 56% | 57% | 91% | 63% |
| | **Indefinite article** | 6% | 19% | 0% | 9% |
| | **Gender error** | 17% | 5% | 8% | 14% |
| | **Case error** | 3% | 3% | 0% | 7% |
| | **Filler word** | 8% | 0% | 0% | 0% |
| | **Omission article** | 10% | 15% | 1% | 1% |
| | **No attempt** | 0% | 1% | 0% | 6% |
| **Dative** | **Definite article correct** | 4% | 8% | 45% | 14% |
| | **Gender error** | 3% | 1% | 1% | 0% |
| | **Case error** | 24% | 7% | 1% | 20% |
| | **Gender & case error** | 3% | 4% | 0% | 1% |
| | **Overgeneralisation accusative** | 32% | 46% | 30% | 30% |
| | **Overgeneralisation *den*** | 15% | 15% | 21% | 16% |
| | **Filler word** | 6% | 8% | 0% | 1% |
| | **Omission article** | 15% | 10% | 2% | 17% |
| | **No attempt** | 0% | 0% | 0% | 2% |

*Note*. Percentages are rounded.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

### *Interpretation*

Ideally, a variable qualifying as suitable clinical marker for SLI should stand out from the children's general language impairment (Rice, 2003). We could thus expect a delay-within-delay pattern for a potential clinical marker, measured by a performance of the SLI group that falls even behind that of language matched control children. Looking at the data from the morphology probes, eliciting case marking in articles, no such pattern can be identified. There is consequently no straightforward candidate in case marking as clinical marker for German SLI. In comparison to typically developing children of similar age, on the other hand, the children with SLI experienced significantly more difficulties with articles in two grammatical cases: accusative and dative. Regarding a case composite derived from all three grammatical cases, too, children with SLI lagged behind the ND-A group. These findings, as well as the results regarding the PI group, will be further discussed below in Section 6.7, and linked to the data from spontaneous language samples.

## 6.6 Spontaneous Data

### 6.6.1 Participants

Sixty-six monolingual German-speaking children participated in the experiment on case marking. Seventeen children with a mean age of 3;11 met the criterion SLI (SLI group), 16 children were age and gender matched and had a phonological speech output impairment (PI group), 17 children were typically language developing and age and gender matched with the SLI group (ND-A group), and a further 16 children were typically language developing and matched with the SLI group on gender and language comprehension (ND-L). All groups are further described in the general method chapter 4. The numbers of participants in the SLI group differ from the elicited data analysis because no children were excluded as outliers prior to the analysis.

### 6.6.2 Procedure

Spontaneous language samples were recorded during free play with the examiner, a parent or the speech and language therapist. Average length of the transcripts was 204 complete and intelligible utterances. The samples were recorded using a SONY minidisk recorder MZ-R909 or a SONY audiotape recorder. The procedure is described in greater detail in chapter 4.

### 6.6.3 Transcription

All recordings of language samples were transcribed by the examiner using the computer software SALT (*Systematic Analysis of Language Transcripts* by Miller & Chapman), research version 7. New conventions were developed by the researcher for German. Section 4.8 in the methods chapter gives interjudge agreement values for 10.6% of the transcripts.

Case marking was highlighted in all structures marked for case: definite and indefinite articles, adjectives and nouns. Possessive pronouns were coded in the same way as adjectives. In this way, data analysis was not restricted to articles as in the elicitation tasks. Kany and Schöler (1998) reported difficulties in case marking also for adjectives and nouns. Although these findings referred to elicited production, it seemed important not to exclude these grammatical contexts from analysis in order to avoid any bias. They had not been considered in the elicitation task of the present study due to the children's young age.

Suffixes marking case were transcribed as bound morphemes. This was the case for indefinite articles, adjectives and nouns. The morphemes were coded differently with respect to gender, even though some morphemes may share the same surface (e.g. the dative article *einem* was transcribed as *ein/emDM* preceding masculine nouns, and *ein/emDN* preceding neuter nouns). Case marking in definite articles and in personal pronouns is monomorphemic (neuter definite article: *das*$_{\text{NOM}}$, *das*$_{\text{ACC}}$, *dem*$_{\text{DAT}}$; personal pronoun 2$^{\text{nd}}$ person singular: *du*$_{\text{NOM}}$, *dich*$_{\text{ACC}}$, *dir*$_{\text{DAT}}$). Therefore, word codes were inserted following these morphemes. The codes provide information over case and word class, and where necessary word-class (e.g. *CN-M* (nominative: masculine definite article); *CAP* (accusative personal pronoun)). See Appendix 5 for a full list of bound morphemes and word codes.

In order to analyse errors in case marking, omitted bound morphemes were marked with an asterisk, and error codes were inserted. Error codes followed the erroneous structure and provided information over the target case, the wrong case that was selected instead and the gender in which the error occurred (e.g. *ECDA-M*: Dative masculine article substituted with accusative). Error codes differed for definite and indefinite articles and also for adjectives. Nouns did not require error codes because errors occurred only as omission and were therefore marked with asterisk preceding the respective morpheme. A list of the error codes is presented in Appendix 5.

Substitution errors could not always unambiguously be categorised, though. As some morphemes are homonyms with different grammatical semantics it was not always obvious which underlying form the child had chosen (e.g. *der* is the correct form for the definite masculine article in nominative as well as the feminine article in dative; the suffix –*em* can mark dative in masculine as well as neuter indefinite articles). The production of an incorrect article could consequently not always unambiguously be identified as gender, case or case and gender error. Neither was it always possible to decide whether a morpheme had been produced correctly because phonological differences were not always perceivable from the recording. For example, it could be difficult to decide whether a child had produced a dative (*einem / dem*) or an accusative article (*einen / den*). Indefinite articles were frequently contracted and only the ending was produced. In nominative masculine articles, the morpheme would thus be shortened to *‘n* (instead if *ein*); accusative and dative articles shortened for example to *‘nen* [nən] and *‘nem* [nəm]. Sometimes, differences between these three contracted forms were hard to perceive or indeed whether the morpheme had even been produced. Such cases were all marked as ambiguous. Contracted forms, though, that could unambiguously be categorised were coded according to case and gender.

### *6.6.4 Quantitative analysis*

Morphemes were selected for statistical analysis if at least eight children per group produced a minimum of two obligatory contexts. Percentages of correct use in obligatory context were calculated by dividing the number of correct productions of the morpheme by the number of correct productions plus the number of grammatical errors. The following errors were considered as grammatical errors:

- Gender error
- Case error
- Gender and case error
- Filler word
- Omission of the article

It was planned that any phonological error, i.e. an error that could result from the children’s individual speech output, would have been excluded from analysis. However, none of the children presented the relevant phonological processes regarding word final /m/ and /n/ and case marking errors went beyond this phonological contrast (see results elicitation tasks). Morphemes coded as ambiguous were excluded from the calculation of percentages, as errors as well as obligatory contexts.

Arcsine transformations were applied to the percentage data and between-group comparisons performed using analysis of variance (ANOVA) followed by Gabriel's procedure. If Levene's test indicated that the assumption of homogeneity of variance was violated Games-Howell procedure was used as post hoc test instead. If significant between group differences were found between the SLI group and the ND-A or ND-L group the variable was further explored as clinical marker by calculation sensitivity, specificity, LR+ and LR-. The results of the children with SLI were transformed to *t*-values relative to the children's age (i.e. ND-A data) and, secondly, relative to the children's language development (i.e. ND-L data).

Furthermore, Pearson's correlation was computed to the arcsine transformed data on nominative articles and the children's age in months.

### *6.6.5 Variables meeting the criterion*

Six variables regarding case marking met the criterion for statistical analysis:

- definite articles in nominative
- definite articles in accusative
- definite articles in dative
- indefinite articles in nominative
- indefinite articles in accusative
- personal pronouns in nominative

Pronominal use of articles was disregarded in order to ensure that the gender of the noun phrase could be unambiguously identified. The number of obligatory contexts for dative indefinite articles did not reach the threshold for statistical analysis. Furthermore, too few obligatory contexts were recorded for case marking in adjectives, in nouns and in accusative or dative personal pronouns as well as in genitive marking.

Definite and indefinite articles in the language sample data were first considered separately. However, in order to increase the comparability with the elicited data definite and indefinite articles were then combined to scores reflecting percentages of correct production of nominative articles, accusative articles and dative articles in general. Thus, the combined scores for nominative, accusative and dative reflect general case marking in articles, independent of definite or indefinite.

As for the elicited data, a case composite was computed for the spontaneous data. All articles marking case correctly were taken together across nominative, accusative and dative irrespective of their definiteness. This was related to the total number of obligatory contexts for articles required in order to calculate a general percentage correct for case marking in articles.

### 6.6.6 Results of quantitative analysis

#### 6.6.6.1 Case Marking in Articles

Table 28 gives the summary data for each group for the total of definite and indefinite articles in nominative, accusative and dative. Although the total number of obligatory contexts varies considerably across the groups in accusative and dative, these differences do not reach significance level when compared using one-way ANOVA.

**Table 28.** Language Samples: Occurrence of Obligatory Contexts for Case Marking in Articles

| ***Case*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative** | **M** | 13.47 | 18.44 | 15.12 | 15.63 |
| | **SD** | 8.49 | 10.60 | 8.57 | 11.60 |
| | **Range** | 3 – 30 | 2 – 38 | 3 – 32 | 3 – 47 |
| | **Total** | 229 | 295 | 257 | 250 |
| **Accusative** | **M** | 10.23 | 13.75 | 14.35 | 10.5 |
| | **SD** | 5.70 | 5.34 | 7.19 | 4.32 |
| | **Range** | 3 – 20 | 7 – 25 | 5 – 29 | 4 – 20 |
| | **Total** | 133 | 220 | 244 | 168 |
| **Dative** | **M** | 3.92 | 5.13 | 6.75 | 5.14 |
| | **SD** | 2.57 | 3.16 | 4.36 | 2.77 |
| | **Range** | 2 – 10 | 2 – 13 | 3 – 19 | 2 – 11 |
| | **Total** | 47 | 77 | 108 | 72 |

*Note*.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation

All four groups' mean for proportional correct use, standard deviations and performance ranges are given in Table 29, together with the number of participants of each group who produced a minimum of two obligatory contexts for the respective case marking articles. Proportional scores were transformed to arcsine values for analysis with one-way ANOVA.

**Table 29.** Language Samples: Proportional Scores of Correct Use of Case Marking in Articles

| ***Case*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative** | **M** | $0.70_{a,b}$ | 0.90 | $0.95_{a}$ | $0.91_{b}$ |
| | **SD** | 0.35 | 0.10 | 0.10 | 0.14 |
| | **Range** | 0.0 – 1 | 0.67 – 1 | 0.67 – 1 | 0.47 – 1 |
| | **n** | 17 | 16 | 17 | 16 |
| **Accusative** | **M** | 0.80 | 0.82 | 0.86 | 0.76 |
| | **SD** | 0.23 | 0.17 | 0.12 | 0.20 |
| | **Range** | 0.17 – 1 | 0.44 – 1 | 0.67 – 1 | 0.29 – 1 |
| | **n** | 13 | 16 | 17 | 16 |
| **Dative** | **M** | $0.21_{a}$ | $0.33_{b}$ | $0.60_{a,b,c}$ | $0.34_{c}$ |
| | **SD** | 0.24 | 0.25 | 0.19 | 0.20 |
| | **Range** | 0.0 – 0.67 | 0.0 – 0.80 | 0.33 – 0.92 | 0.0 – 0.57 |
| | **n** | 12 | 15 | 16 | 14 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

One-way ANOVA indicated a significant main effect for nominative and for dative case marking. The main group effect for nominative marking in articles ($F$ (3, 62) = 5.33, $p$ = .002, $\omega^2$ = .16) was further explored by the Gabriel's procedure which revealed a significant weaker production of nominative articles in the SLI group compared to the ND-A group ($p$ = .002) and to the ND-L group ($p$ = .025). Gabriel's procedure indicated that the significant main effect in dative articles ($F$ (3, 53) = 8.59, $p$ = .000, $\omega^2$ = .26) was due to a significant lower performance of the SLI group ($p$ = .000), the PI group ($p$ = .005) as well as the ND-L group ($p$ = .006) on dative marking in comparison to the ND-A group. No main effect or significant group differences were found for accusative marking in articles. The results are illustrated in Figure 17 showing group means of the proportional scores of correct case marking.

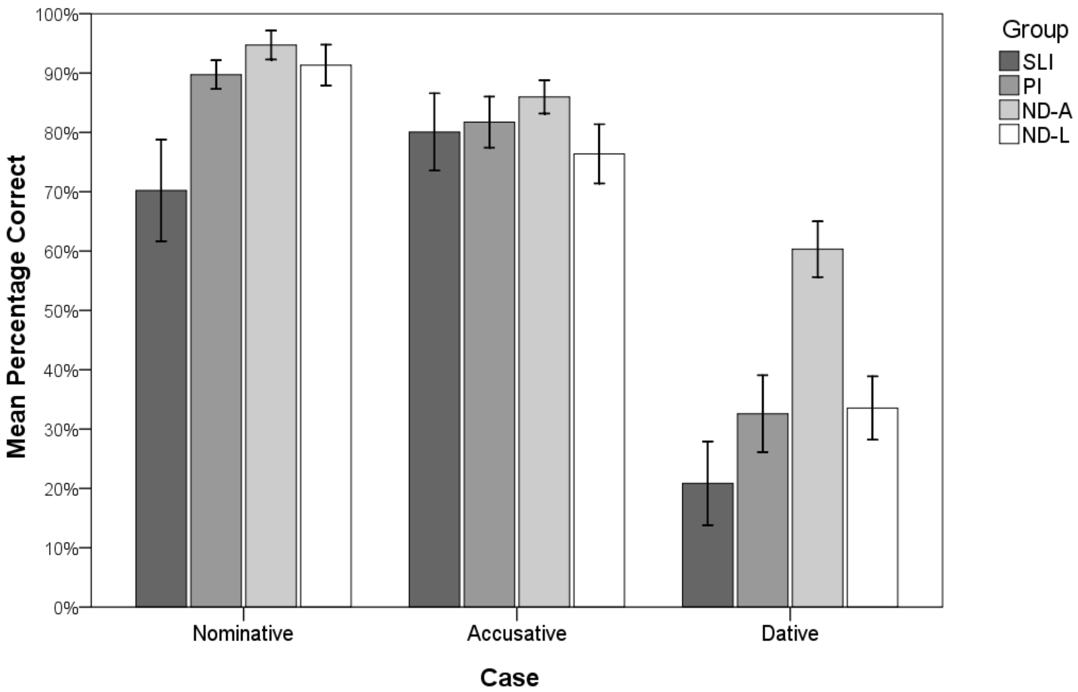



**Figure 17.** Mean percentages of use (± 1 SE) for nominative, accusative and dative case marking in spontaneous language for the four participant groups.

Table 49 in the Appendix 8 shows an accuracy level of 90.5% as the most promising cut-off for spontaneous nominative articles, resulting in an overall discrimination accuracy of 76%, a sensitivity of 0.706 [0.469 to 0.867] was obtained, a specificity of 0.824 [0.59 - 0.938], LR+ = 4 [1.37 to 11.682] and LR- = 0.357 [0.166 to 0.779. For dative articles, the best cut-off appeared at a production accuracy of 30% (s. Table 50), resulting in the correct classification of 89% of the children as SLI or ND-A, a sensitivity of 0.75, specificity of 1, and

LR- = 0.25. Due to the absence of false positives, it was not possible to calculate neither LR+ nor the confidence intervals.

Percentiles relative to the ND-A group are given in Appendix 10. Percentiles of the performance of the SLI group relative to the ND-L group are presented in Appendix 11.

The analysis of nominative articles was carried further by correlating the children's age in months with their percentage correct of nominative articles. There was a trend to a medium positive relationship between age and nominative accuracy, $r$ = .198, $p$ (one-tailed) = .056.

#### *6.6.6.2 Case marking in definite and indefinite articles*

While the section above gives the results for combined scores for definite and indefinite articles, both contexts were also considered separately. Table 30 provides the summary data. As before, cases in which the article was contracted with the preceding preposition were regarded as obligatory context if the article was clearly marked.

Group sizes and the descriptives of the statistical analysis are given in Table 31. In nominative, three children of the SLI group produced two or more obligatory contexts of indefinite nominative articles but less than two obligatory contexts for definite articles. Three other children produced enough obligatory contexts for analysis of definite nominative articles but insufficient for the analysis of indefinite articles. The same applied to two children of the SLI group respectively with the production of accusative articles. Transformed arcsine values were entered for between-group comparison with one-way ANOVA.

**Table 30.** Language Samples: Occurrence of Obligatory Contexts for Case Marking in Definite and Indefinite Articles

| *Case* | | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|---|
| **Nominative:** | **M** | 7.50 | 10.60 | 8.41 | 8.53 |
| **Definite articles** | **SD** | 4.72 | 6.06 | 6.16 | 6.67 |
| | **Range** | 2 – 17 | 4 – 23 | 2 – 22 | 2 – 27 |
| | **Total** | 106 | 160 | 143 | 129 |
| **Nominative:** | **M** | 8.57 | 10.31 | 7.47 | 7.93 |
| **Indefinite articles** | **SD** | 4.75 | 8.01 | 2.75 | 5.60 |
| | **Range** | 3 – 16 | 2– 27 | 3 – 14 | 2 –21 |
| | **Total** | 123 | 135 | 114 | 119 |
| **Accusative:** | **M** | 5.91 | 7.5 | 6.63 | 6.47 |
| **Definite articles** | **SD** | 2.84 | 3.16 | 3.32 | 3.27 |
| | **Range** | 2 – 10 | 3 – 14 | 2 – 14 | 2 – 12 |
| | **Total** | 69 | 120 | 106 | 98 |
| **Accusative:** | **M** | 5.91 | 6.60 | 8.56 | 4.67 |
| **Indefinite articles** | **SD** | 3.83 | 4.17 | 4.97 | 2.26 |
| | **Range** | 2 – 12 | 2 – 16 | 2 – 19 | 2 – 8 |
| | **Total** | 67 | 100 | 138 | 70 |
| **Dative:** | **M** | 4.00 | 4.71 | 6.25 | 4.79 |
| **Definite articles** | **SD** | 2.87 | 2.92 | 4.23 | 2.42 |
| | **Range** | 2 – 10 | 2 – 11 | 3 – 19 | 2 – 10 |
| | **Total** | 39 | 66 | 100 | 68 |

*Note*.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation

In nominative, considering definite and indefinite articles separately, ANOVA shows a significant group effect only for definite articles, $F$ (3, 27.06) = 3.03, $p$ = .046, $\omega^2$ =.09 (Brown-Forsythe $F$-ratio given because of significant results of test of homogeneity of variance). However, the following post hoc procedure using Games-Howell did not present significant differences between particular groups.

**Table 31.** Language Samples: Proportional Scores of Correct Use of Case Marking in Definite and Indefinite Articles

| ***Case*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative:** | **M** | 0.73 | 0.89 | 0.95 | 0.89 |
| **Definite articles** | **SD** | 0.34 | 0.14 | 0.11 | 0.17 |
| | **Range** | 0.0 – 1 | 0.60 – 1 | 0.60 – 1 | 0.50 – 1 |
| | **n** | 14 | 15 | 17 | 15 |
| **Nominative:** | **M** | 0.79 | 0.91 | 0.94 | 0.92 |
| **Indefinite articles** | **SD** | 0.32 | 0.13 | 0.11 | 0.17 |
| | **Range** | 0.0 – 1 | 0.60 – 1 | 0.67 – 1 | 0.33 – 1 |
| | **n** | 14 | 13 | 15 | 15 |
| **Accusative:** | **M** | 0.69 | 0.78 | 0.85 | 0.73 |
| **Definite articles** | **SD** | 0.32 | 0.19 | 0.17 | 0.17 |
| | **Range** | 0.0 – 1 | 0.43 – 1 | 0.50 – 1 | 0.29 – 1 |
| | **n** | 11 | 16 | 16 | 15 |
| **Accusative:** | **M** | 0.95 | 0.89 | 0.88 | 0.85 |
| **Indefinite articles** | **SD** | 0.11 | 0.20 | 0.16 | 0.15 |
| | **Range** | 0.67 – 1 | 0.33 – 1 | 0.50 – 1 | 0.50 – 1 |
| | **n** | 11 | 15 | 16 | 15 |
| **Dative:** | **M** | $0.29_{a}$ | $0.38_{b}$ | $0.68_{a,b,c}$ | $0.35_{c}$ |
| **Definite articles** | **SD** | 0.27 | 0.25 | 0.21 | 0.21 |
| | **Range** | 0.0 – 0.8 | 0.0 – 0.8 | 0.33 – 1 | 0.0 – 0.57 |
| | **n** | 9 | 14 | 16 | 14 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

Results in both accusative and dative reflect those reported above for combined scores of definite and indefinite articles. Main effects for accusative marking were not revealed in either definite or in indefinite articles. In dative definite articles, ANOVA showed a highly significant main effect, $F$ (3, 49) = 7.65, $p$ = .000, $\omega^2$ =.29. Gabriel's procedure identified the same significant group differences as for the combined case marking scores. The ND-A group produced significantly more correct definite dative articles than the three other groups: SLI ($p$ = .001), PI ($p$ = .004) and ND-L ($p$ = .001). The group means of the proportional scores are displayed in Figure 18 for all five variables.

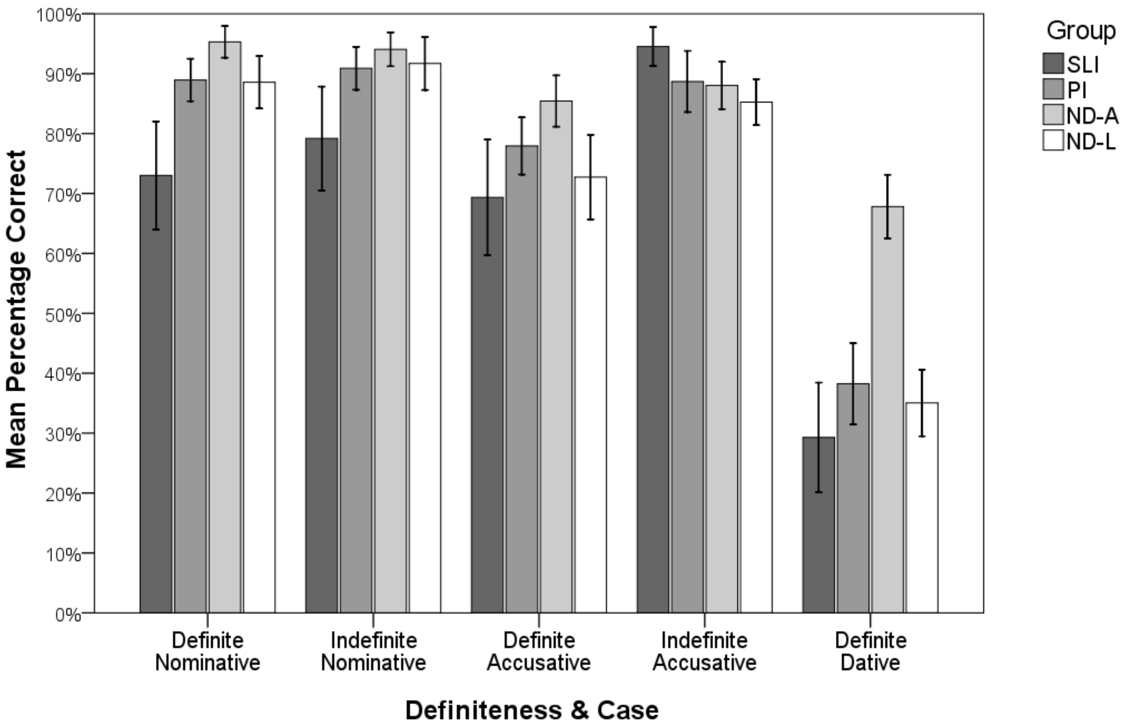



**Figure** 18**.** Mean percentages of use (± 1 SE) for nominative, accusative and dative definite and indefinite articles in spontaneous language for the four participant groups.

Table 51 in Appendix 8 suggests 30% as the most suitable cut-off for definite dative articles in spontaneous language. Applied to the present data, this cut-off resulted in the correct classification of 88% of the children a sensitivity of 0.67, the specificity of 1, and LR- = 0.33. Due to the absence of false positives, it was not possible to calculate neither LR+ nor the confidence intervals.

**Case composite from spontaneous language data**

Obligatory contexts for articles marking case and all produced articles were taken together for the case composite. The summary data for correct case marking in articles is given in Table 32.

**Table 32.** Language Samples: Compositional Score Case

| ***Compositional Score*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Case** | **M** | $0.62_{a}$ | 0.80 | $0.85_{a,b}$ | $0.75_{b}$ |
| | **SD** | 0.31 | 0.09 | 0.09 | 0.13 |
| | **Range** | 0 – 0.94 | 0.63 – 0.93 | 0.67 – 0.99 | 0.42 – 0.93 |
| | **n** | 17 | 16 | 17 | 16 |

*Note.* Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

Arcsine transformed values were entered for statistical analysis with one-way ANOVA. Levene’s test showed that the assumption of homogeneity was not met for the case composite. Brown-Forsythe F-ratio showed a main group effect ($F$ (3, 33.33) = 5.58; $p$ = .003, ω2 = .17) and Games-Howell post hoc test revealed that two groups preformed significantly below the ND-A group: the SLI group ($p$ = .016) and the ND-L group ($p$ = .047). Figure 19 illustrates the mean proportional scores in case marking across the four participant groups.

If using the composite for spontaneous case marking in articles as diagnostic tool, 79.2% was the most suitable cut-off as the overview over the data in Table 52 demonstrates. This cut-off resulted in a in a diagnostic accuracy of 76%, a sensitivity of 0.706 [0.469 - 0.867], a specificity of 0.824 [0.59 - 0.938], LR+ = 4 [1.37 - 11.682] and LR- = 0.357 [0.166 - 0.77].

Percentiles relative to the ND-A group are given in Appendix 10.
Percentiles of the performance of the SLI group relative to the ND-L group are presented in Appendix 11.

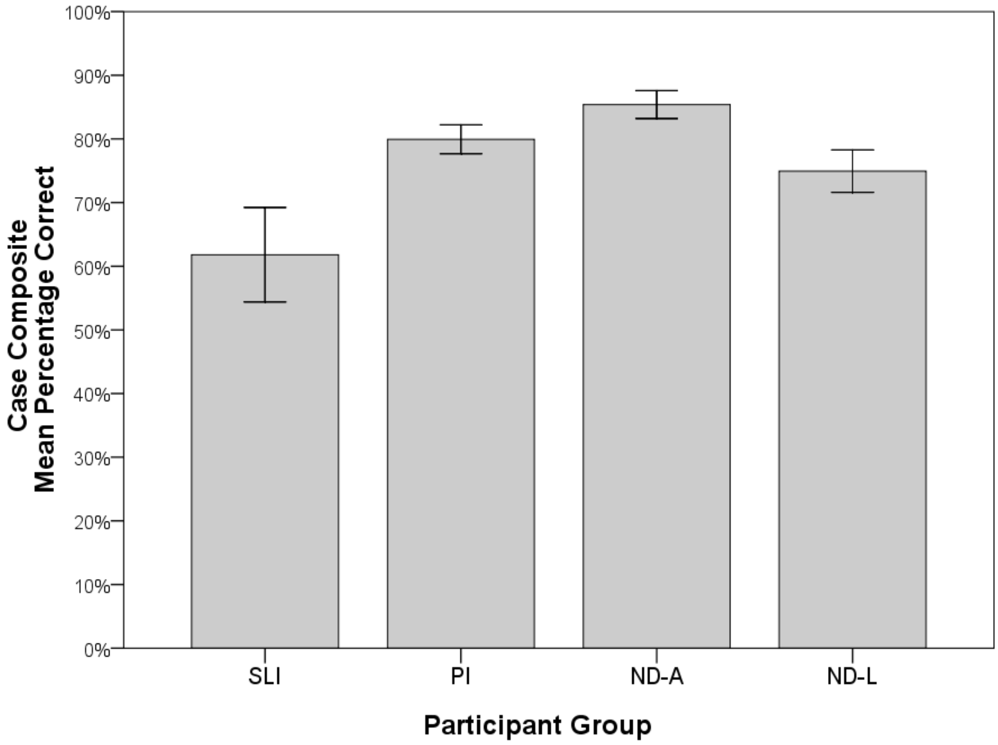



**Figure 19.** Mean percentages (± 1 SE) for case composite representing proportional correct production in spontaneous language.

### *6.6.6.4 Case marking in personal pronouns*

The use of personal pronouns met the criterion for statistical analysis in nominative[16]. Table 33 gives an overview of the number of obligatory contexts for each group. There was a significant main effect for group regarding the number of obligatory contexts for personal pronouns in nominative, $F$ (3, 61) = 3.63, $p$ = .018, $\omega^2$ = .11. The SLI group produced almost only half as much obligatory contexts than the PI and ND-A group. Gabriel's procedure indicated, that the first of these between group differences was significant, $p$ = .020.

Table 34 presents the descriptive results, together with the number of participants of each group who produced a minimum of two obligatory contexts for personal pronouns in nominative.

[16] See Section 4.6.2.1: Morphemes were selected for statistical analysis if at least eight children per group produced a minimum of two obligatory contexts.

Not only the frequency but also the accuracy in the production of personal pronouns in nominative differed significantly across the groups, $F$ (3, 61) = 3.19, $p$ = .030, $\omega^2$ = .09. Gabriel's procedure revealed that the children with SLI produced significantly less correct personal pronouns than the ND-A group, $p$ = .033. The findings appear in Figure 20, representing the proportional scores before arcsine transformation.

**Table 33.** Language Samples: Occurrence of Obligatory Contexts for Personal Pronoun Nominative

| ***Personal pronoun*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative** | **M** | $27.19_a$ | $50.25_a$ | 46.82 | 42.25 |
| | **SD** | 22.13 | 27.31 | 18.61 | 16.17 |
| | **Range** | 2 – 71 | 14 – 120 | 12 – 76 | 4 - 74 |
| | **Total** | 435 | 804 | 796 | 676 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation

**Table 34.** Language Samples: Proportional Scores of Correct Use of Personal Pronoun Nominative

| ***Personal pronoun*** | | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative** | **M** | $0.81_a$ | 0.96 | $0.98_a$ | 0.94 |
| | **SD** | 0.33 | 0.07 | 0.02 | 0.13 |
| | **Range** | 0.0 – 1 | 0.71 – 1 | 0.92 – 1 | 0.50 – 1 |
| | **n** | 16 | 16 | 17 | 16 |

*Note*. Means in a row sharing subscripts are significantly different when arcsine transformations analysed with ANOVA.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group
M = Mean proportional score; SD = Standard deviation; n = Number of participants

Table 53 in Appendix 8 suggests a production accuracy of 95.9% as the most accurate cut-off nominative personal pronouns, resulting in an overall discrimination accuracy of 79%, a sensitivity of 0.625 [0.386 to 0.815], a specificity of 0.941 [0.73 - 0.99], LR+ = 10.625 [1.528 to 73.859] and LR- = 0.398 [0.209 to 0.758].

Percentiles relative to the ND-A group are given in Appendix 10. Percentiles of the performance of the SLI group relative to the ND-L group are presented in Appendix 11.

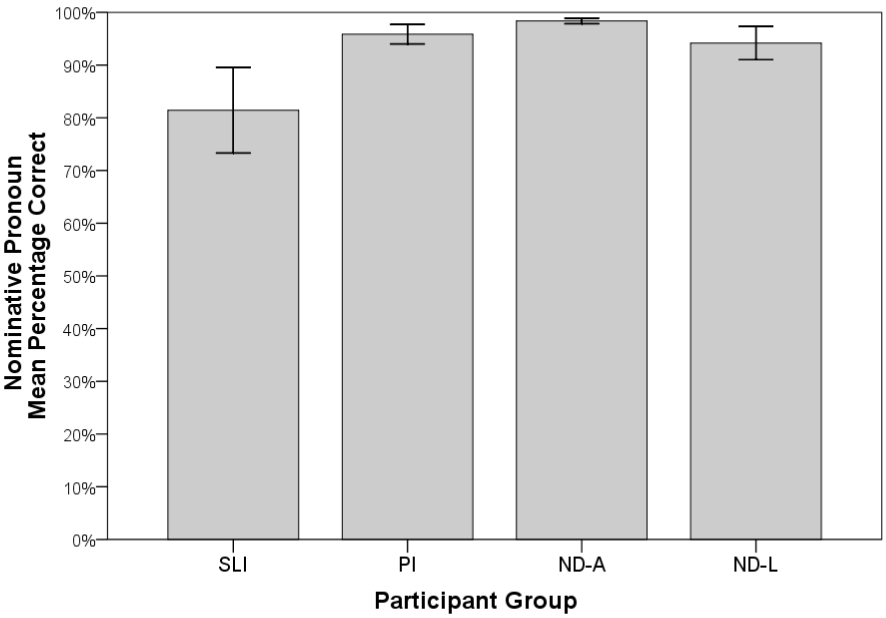



**Figure 20.** Mean percentages of correct use in spontaneous language (± 1 SE) for personal pronouns in nominative.

### 6.6.7 Qualitative analysis

A qualitative analysis was performed for the same variables that met the criterion for statistical analysis. However, definite and indefinite articles were combined in order to focus on case marking rather than the choice of definite versus indefinite articles. Proportions were calculated of the different production patterns, i.e. the proportion of correct responses and the proportion of all error patterns respectively.

For this, the total number of obligatory contexts for nominative, accusative and dative respectively was calculated by summing-up the number of obligatory contexts within each participant group. Per error category and for the category correct productions, the number of productions was then computed per group. This number of productions was divided by the total number of obligatory contexts for each target case in order to calculate the proportion of the respective response pattern. The category *phonological error* does not appear as the children did not make any errors resulting from a speech output disorder.

No statistical between group comparisons were performed regarding production pattern because the available quantity of data per category was too limited.

### *6.6.8 Results of qualitative analysis*

**Production pattern in articles**

The following categories of production patterns in articles were found across cases:

- Correct case marking
- Gender error
- Case error
- Gender and case error
- Filler word
- Omission of the article

Errors regarding definiteness did not occur. Figure 21 presents an overview over the production patterns across the groups for nominative articles. Figure 22 illustrates the production patterns for accusative articles and Figure 23 the production patterns for dative articles. Proportions in percentages are given in Table 35.

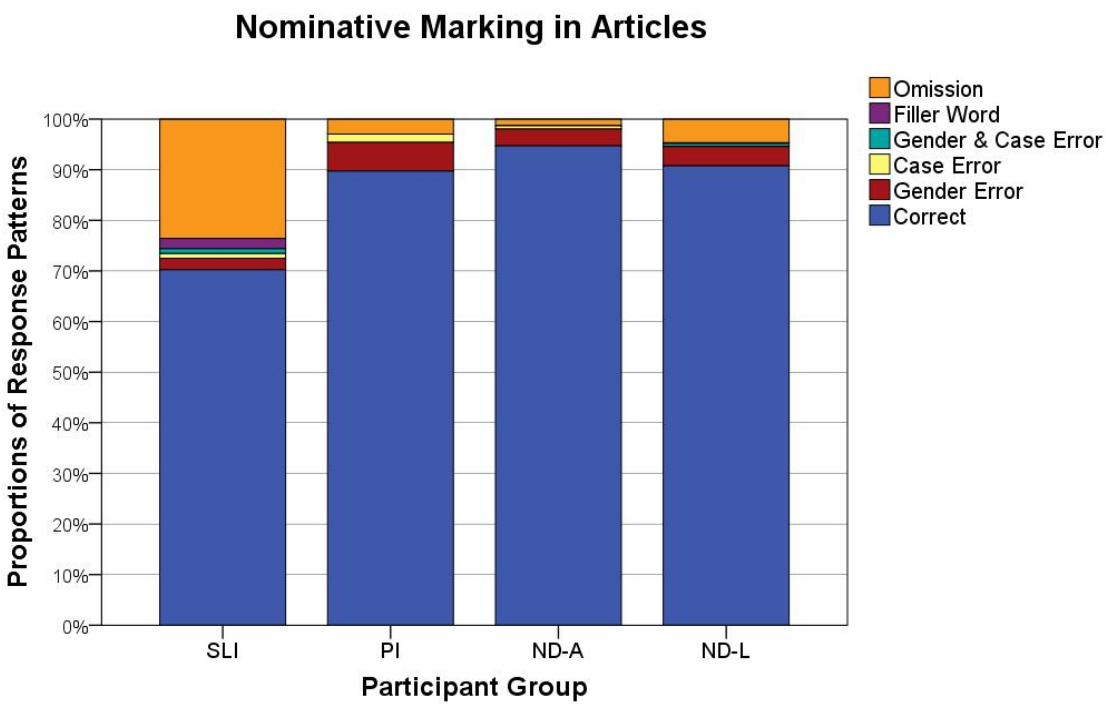



**Figure 21.** Spontaneous production patterns for nominative articles.

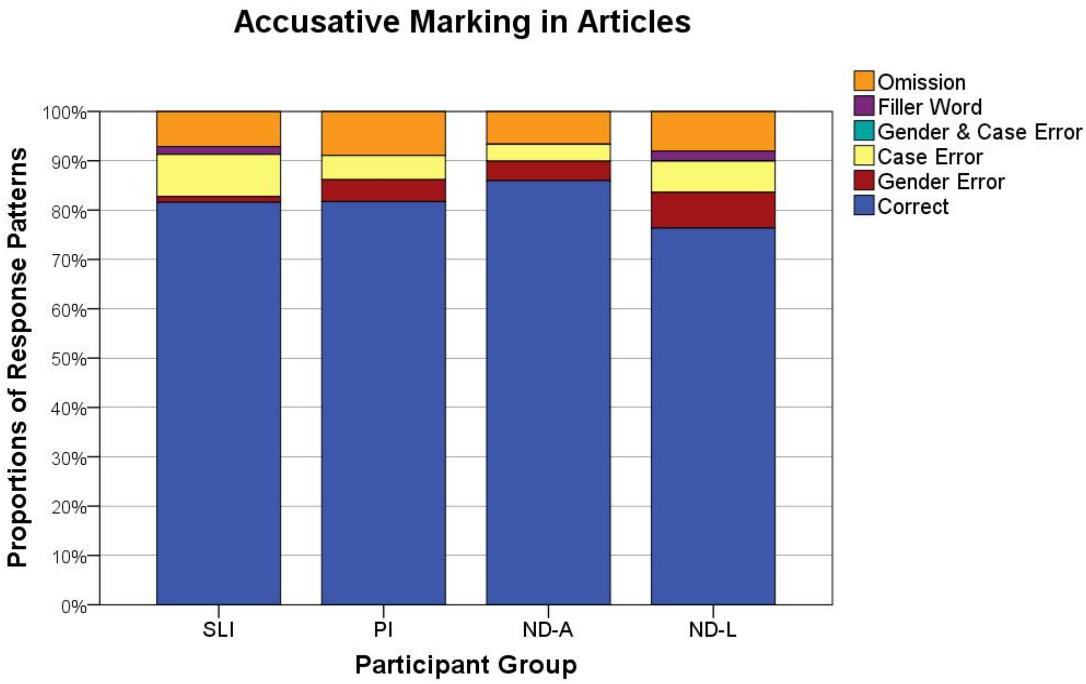



**Figure 22.** Spontaneous production patterns for accusative articles.

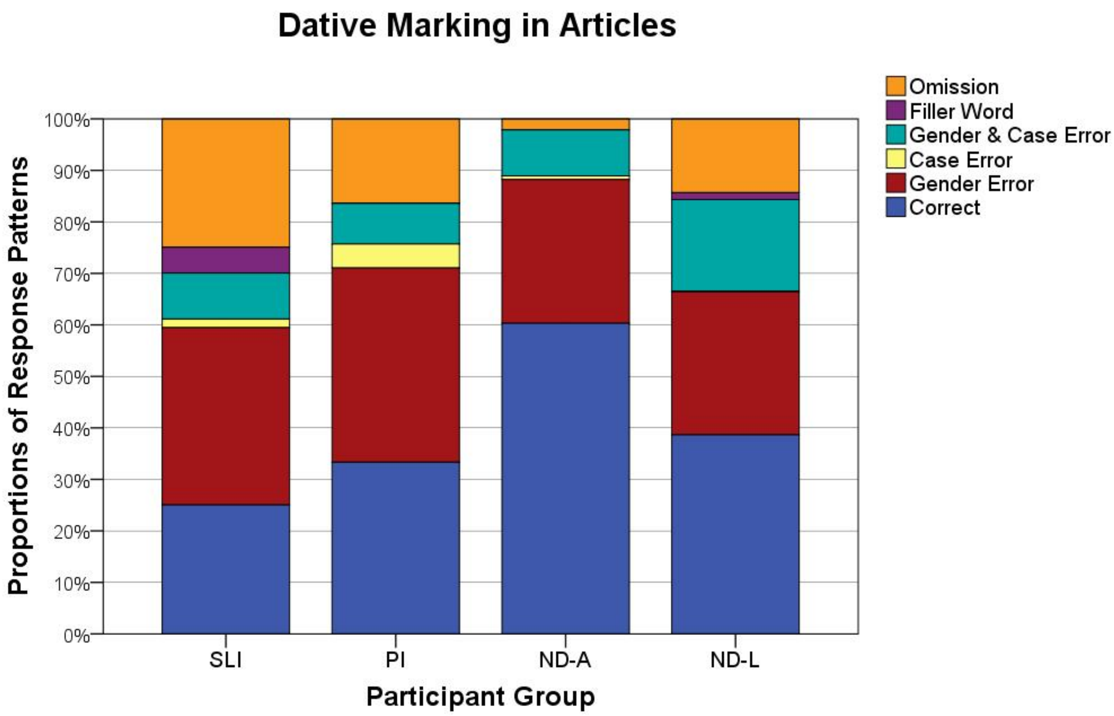



**Figure 23.** Spontaneous production patterns for dative articles.

**Table 35.** Language Samples: Production Patterns for Case Marking in Articles

| | ***Production Pattern*** | ***SLI*** | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| **Nominative** | **Nominative article correct** | 70% | 90% | 95% | 91% |
| | **Gender error** | 2% | 6% | 3% | 4% |
| | **Case error** | 1% | 2% | 1% | 0% |
| | **Gender & case error** | 1% | 0% | 0% | 1% |
| | **Filler word** | 2% | 0% | 0% | 0% |
| | **Omission article** | 24% | 3% | 1% | 5% |
| **Accusative** | **Accusative article correct** | 82% | 82% | 86% | 76% |
| | **Gender error** | 9% | 5% | 3% | 6% |
| | **Case error** | 1% | 4% | 4% | 7% |
| | **Gender & case error** | 0% | 0% | 0% | 0% |
| | **Filler word** | 2% | 0% | 0% | 2% |
| | **Omission article** | 7% | 9% | 7% | 8% |
| **Dative** | **Dative article correct** | 25% | 33% | 60% | 39% |
| | **Gender error** | 2% | 5% | 1% | 0% |
| | **Case error** | 34% | 38% | 28% | 28% |
| | **Gender & case error** | 9% | 8% | 9% | 18% |
| | **Filler word** | 5% | 0% | 0% | 1% |
| | **Omission article** | 25% | 16% | 2% | 14% |

*Note*. Percentages are rounded.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

**Production pattern in nominative personal pronouns**

The production of personal pronouns in nominative was analysed regarding the error types that occurred in the children's spontaneous language. The following categories of production patterns were found across different persons and numbers:

- Correct production
- Omission of the nominative personal pronoun
- Case error
- Substitution

The most common error pattern was the omission of the pronoun. Two children showed substitution errors, both for the first person singular personal pronoun *ich* (Engl.: *I*). A case error occurred only once, in the PI group. Figure 24 illustrates the distribution of the production patterns for each group. Table 36 presents the relative frequency for each error pattern.

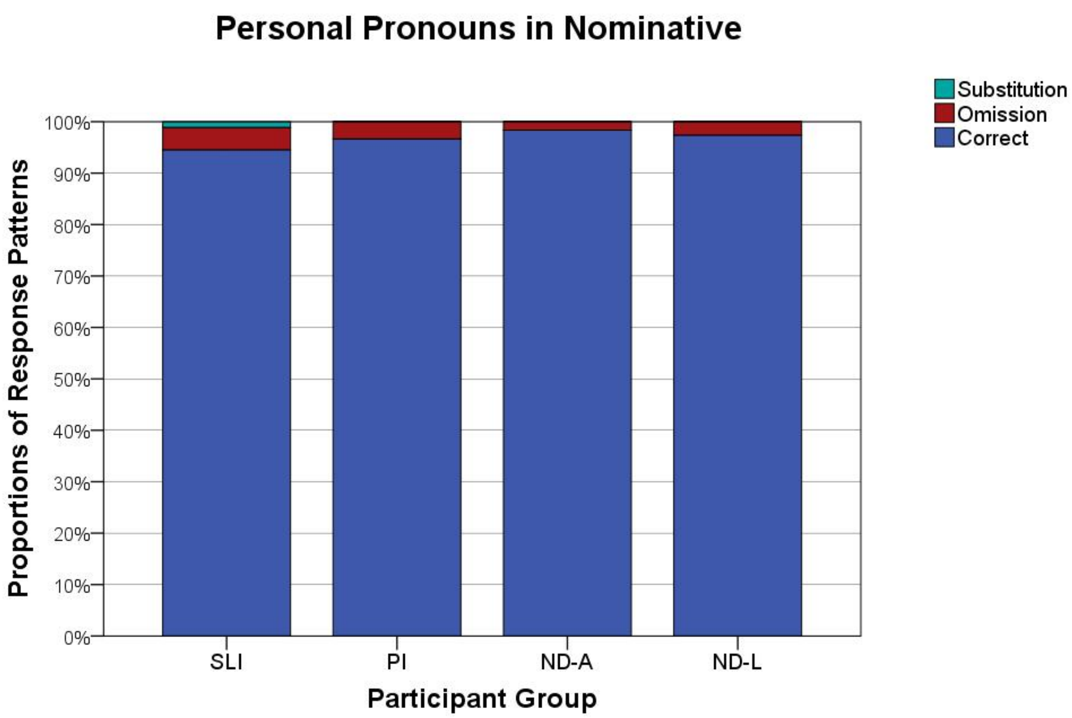



**Figure 24.** Spontaneous production patterns for personal pronouns in nominative.

**Table 36.** Language Sample: Production Patterns for Personal Pronouns in Nominative

| | *Production Pattern* | *SLI* | *PI* | *ND-A* | *ND-L* |
|---|---|---|---|---|---|
| **Nominative** | **Nominative pronoun correct** | 81% | 96% | 98% | 94% |
| | **Omission pronoun** | 6% | 4% | 2% | 6% |
| | **Case error** | 0% | 0% | 0% | 0% |
| | **Substitution** | 13% | 0% | 0% | 0% |

*Note*. Percentages are rounded.
LI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

### 6.6.9 Interpretation

The between group comparisons of spontaneous case marking revealed a delay within delay pattern for the SLI group in the production of nominative articles. The production accuracy of definite and indefinite nominative articles combined may therefore be a suitable candidate for a clinical marker for German SLI. Moreover, the children with SLI lagged significantly behind their age-matched, typically developing peers on dative articles as well as on a case composite derived from spontaneous language. The findings will be discussed below further regarding their relevance for SLI, PI and the findings from the elicitation tasks.

## 6.7 Discussion

### *Case marking as clinical marker for German SLI?*

The primary intention of the presented study was to investigate the potential of case marking as clinical marker for SLI in German. Findings from two measures were presented above: morphology probes involving elicitation tasks for definite articles and spontaneous language samples. The results from between group comparisons across both measures are summarised in Table 19..

**Table 37.** Results from Between Group Comparisons for Case Marking

| | Data source | |
|---|---|---|
| ***Dependant variable*** | ***Elicitation tasks*** | ***Spontaneous language*** |
| Nominative | - | SLI < ND-A; SLI < ND-L |
| Nominative definite articles | n.a. | main effect |
| Nominative indefinite articles | n.a. | - |
| Personal pronoun nominative | n.a. | SLI < ND-A |
| Accusative | SLI < ND-A | - |
| Accusative definite articles | n.a. | - |
| Accusative indefinite articles | n.a. | - |
| Dative | SLI < ND-A; ND-L < ND-A | SLI < ND-A; PI < ND-A; ND-L < ND-A |
| Dative definite articles | n.a. | SLI < ND-A; PI < ND-A; ND-L < ND-A |
| Case composite for articles | SLI < ND-A; PI < ND-A; ND-L < ND-A | SLI < ND-A; ND-L < ND-A |

*Note.* Only statistically significant group differences are presented ($p < 0.05$).
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group

### *6.7.1 Delay within delay*

As mentioned above, one of the two measures revealed a delay within delay pattern for the SLI group on case marking: children with SLI did not only lag behind age matched typically developing children in spontaneous case marking in nominative articles but they also fell behind children matched for language comprehension. On average, children with SLI produced at least 20% more errors on nominative articles in spontaneous language than children with PI, age matched controls and language matched controls. This pattern meets the prediction for a clinical marker. Only the difference between the SLI and PI group does not meet statistical significance. Yet, this seems surprising. The mean production accuracy of the PI group is minimally below that of the ND-L group; standard deviation and range are even smaller than that of the ND-L group. It is probably the arcsine transformation of the data that explains the lack of a significant between group differences. This transformation, on the other hand, was necessary in order to change the binominal distribution of percentage data to a normal distribution. The data from elicited nominative articles, on the other hand, substantiate the doubts in how far weak nominative marking is specific to SLI. The advantage of the PI group over the SLI group almost disappears in elicited data as the children with PI seem to have more difficulties with the morphology probes than with spontaneous production. Nevertheless, the findings from spontaneous language samples suggest that the production of nominative articles may be a suitable diagnostic instrument for the discrimination between typically developing German-speaking children and children with SLI. Yet, the specificity of such a marker remains questionable.

Considering the other measure tool used in the present study, statistical analysis failed to present the same picture for the accuracy of nominative articles. The production of elicited nominative articles did not reveal significant group differences. This raises doubts whether nominative marking poses indeed a particular difficulty for children with SLI. The fact that multiple between group comparisons were carried out bears the risk of a significant difference for the spontaneous data purely due to chance, i.e. a type I error. A look at mean percentages correct instead of *p*-values, though, supports the statistical results for the spontaneous data. As mentioned, the advantage of all control groups over the SLI group is considerable in spontaneous language. Moreover, mean percentages correct for elicited data, in fact, show quite a comparable pattern. The children with SLI had an accuracy rate of 20% less than both control groups of typically developing children. Even the younger control children demonstrated a very high performance for elicited nominative articles. This is also reflected in a relatively small performance range within this group. The lack of a significant group difference between the SLI group and ND-A as well as ND-L group may result from the relatively low number of participants, especially within the SLI group, and from the low number of items per grammatical case. In spontaneous language the impact of both these factors is reduced: Group sizes do not differ between SLI and typically developing control groups and the number of obligatory contexts for nominative articles is more than three times as much as for the elicitation tasks. Moreover, the ND-A group demonstrates a considerably lower performance range in spontaneous than in elicited language. All these factors are likely to have contributed to the results of the statistical analysis. A type I error in the analysis of spontaneous data seems therefore improbable. Instead, it may be worth to investigate the elicited production of nominative articles again with a larger number of participants and more items.

These promising results in nominative articles demand a closer look at the usefulness of this variable as potential clinical marker. The suggestion that children with SLI lag behind their general language development in the production of nominative articles is supported by an acceptable specificity rate of 82%. Only three children of the ND-A group would incorrectly be identified as SLI if a cut-off of 90.5% was applied to the spontaneous language data. A specificity rate of 90% or above and a lower LR- ratio, though, would be even more desirable for an appropriate clinical marker. Furthermore, in order to use this error pattern as clinical marker it should be evident in the language of all children with SLI, i.e. present a satisfying sensitivity rate. Actually, two facts shed doubts regarding the marker's sensitivity for German SLI. Firstly, nominative marking has commonly been reported to be rather unaffected in German SLI (e.g. Eisenbeiss et al., 2006; Schöler et al., 1998a), and secondly, the SLI group of the current study does not present a homogeneous production pattern. The impact of these facts on the clinical marker's potential is discussed below.

Eisenbeiss et al. (2006) as well as Schöler, Fromm and Kany (1998a) report that nominative articles are usually produced correctly by German-speaking children with SLI. Both groups of researchers, though, investigated the language of older children than in the present study (age 7;6-12;0 years and age 5;9-7;11years respectively). A possible explanation for the conflicting findings may therefore be that a weakness in nominative articles is only a temporary phenomenon in young German-speaking children with SLI. There is indeed a small correlation between age and accuracy of nominative articles (both elicited and spontaneous) in the present data. For spontaneous data, this correlation is not statistically significant, though, but represents a trend. Consequently, the present data cannot answer satisfyingly in how far the weakness in nominative articles is restricted to young children with SLI.

Further previous research on German SLI is unfortunately of little help in this respect. Nominative articles were rarely considered explicitly: Clahsen (1989, 1991) focused predominantly on accusative and dative marking. Roberts and Leonard (1997) reported high omission rates for articles in children with SLI aged between 4;8 and 7;2, but they do not present the data for grammatical cases separately. The qualitative analysis of the present data suggests in fact that a large proportion of the omissions reported by Roberts and Leonard may concern the case nominative. It is this grammatical case, for which omissions represent the largest proportion of error patterns in the present study. This is true for both measures. If the children produced obligatory contexts for accusative or dative, case or gender errors presented greater obstacles to the children than the omission of the article. This pattern for accusative and dative is conform with the developmental course as it is described in typical language development (Schrey-Dern, 2006; Szagun, 2006). In the early acquisition of case marking, noun phrases are not marked yet because they miss the article. During the next phase, children start producing articles in nominative which is at that stage considered as case neutral (Christina Kauschke, 2012). The awareness that noun phrases include an article is therefore rather a step in syntactic than morphological development. The following phase of case acquisition is characterised by the differentiation between a subject case (nominative) and an object case in which accusative is often overgeneralised to dative articles. Only later, children make a distinction between accusative and dative (Christina Kauschke, 2012). This sketch of developmental trajectory illustrates that the ‘options’ of possible errors increase for accusative and dative as opposed to nominative. But are omissions of nominative articles a phenomenon characteristic of young children in general acquiring German, young children with SLI only or pre-school children with SLI in general?

One possibility could be that children with SLI remain longer at the first two stages of case development. A look at the children’s individual data reveals that only one child of the present study can be considered to be still at the very first developmental stage: One boy of the SLI-group (SLI2) omitted the determiner in all noun phrases which he produced spontaneously. All remaining 65 children participating in this study produced correct and incorrect noun phrases, also those children who presented omissions of the nominative article. For the SLI group that were nine children. At best, these latter children could be described as being in a transition phase between the first and second stage of case development. The remaining seven children of the SLI group, though, can be considered to have passed these phases and are more advanced as they did not show any omissions of nominative articles. Longitudinal data providing developmental trajectories would be necessary in order to investigate if children with SLI stagnate temporarily on the early stages of case development. Findings in this respect could raise the marker’s sensitivity.

The question of the clinical markers’ sensitivity brings us back to the second fact mentioned above, challenging nominative marking as clinical marker: not all children with SLI demonstrated high omission rates of nominative articles. In spontaneous production, seven out of 17 children with SLI performed above an accuracy level of 90% in nominative articles, of which three children did not make any errors. This is reflected in a rather low sensitivity level of 71% even if the cut-off is set as high as at an accuracy level of 90.5%. In elicited articles, seven out of 13 children produced the correct nominative article for all items.

They clearly did not present any difficulties and would not have been identified by a clinical marker test tapping nominative in articles. The analysis of the results differs between both measures in terms of the exclusion of four particularly weak children, i.e. outliers. Whereas one of these children showed few obligatory contexts for nominative articles in spontaneous language but all of them correct, the other three children pulled the mean percentage correct considerably down for the whole group. Only one further child with SLI produced below 20% of the nominative articles correct while the remaining 13 children produced 75% or more correct. It is consequently likely that this methodological difference between both measures contributed to the significant group difference between the SLI and ND-L group in spontaneous language and the lack of a significant difference in elicited production. In this respect, the difficulties in nominative marking cannot be considered as a consistent weakness in children with SLI. Alternatively to a homogeneous production pattern across the whole population of children with SLI, an age threshold could be used for the identification of children with SLI if this particular error pattern turns out to be present only in young children with SLI. As discussed above, though, the role of the children’s age would need to be investigated in greater detail in order to explore the potential of nominative articles as clinical marker for a certain age group within the SLI population.

In conclusion, children with SLI show a delay within delay pattern in the spontaneous production of nominative articles as it would be predicted for a clinical marker. In comparison with typical language development, children with SLI seem to follow a typical developmental trajectory in their production of nominative articles. But they fall nevertheless even behind the production accuracy expected relative to their language comprehension. A large within group variability and a high level of performance accuracy in a number of children with SLI, though, indicate that this grammatical morpheme is rather unsuitable as clinical marker. Further research might identify a developmental pattern within the SLI group. As a consequence, the accuracy of nominative articles may present potential as clinical marker for a certain age group only. However, results of the current study lack a significant group difference for the SLI and PI group. Difficulties with nominative articles are thus neither characteristic for all pre-school children with SLI nor necessarily specific to SLI. For now, the average age of the ND-L group (3:3 years) should be used as latest cut-off for the decision to carry out further diagnostics if children demonstrate the omission of nominative articles – this should thus involve language testing as well as phonology tasks.

### 6.7.2 Delay

#### 6.7.2.1 Is case marking delayed in German-speaking children with SLI?

Besides nominative marking in articles, several other variables of case marking were investigated, too. Between group comparisons demonstrated that German-speaking children with SLI have difficulties with elicited accusative articles relative to age-matched peers, with the elicited and spontaneous production of dative articles and with personal pronouns in nominative. A case composite score across all three tested grammatical cases also suggested a delay in the development of case marking in articles. Taking these results and the discussion above together, it can be concluded that case marking may be problematic for children with SLI but it does not generally fall behind the children's general language impairment. The fact that the results for the case composite correspond across both measures validates this conclusion. Eisenbeiss, Bartke and Clahsen (2006), too, drew a similar conclusion. The authors' comparison of spontaneous case marking in children with SLI to case marking in younger children matched for MLU did not result in significant differences between both groups.

At a first glance, these results may be surprising. Case marking in general had been identified as the most vulnerable area of the children's morphology in German SLI in earlier research (Kany & Schöler, 1998). This is confirmed by the data of the present study. In elicited case marking, children with SLI succeed on average only in 47 % of the test items. The average production accuracy for articles in spontaneous language does not exceed 62%. The assumption raised in the literature review above, however, that general case marking may be a key characteristic of young children with SLI (s. Clahsen, 1989, 1991; Roberts & Leonard, 1997) could not be confirmed. The inclusion of a language matched control group demonstrated that these difficulties are in fact in line with the children's language development. Case marking is therefore a definite weakness in SLI, but this weakness does not stand out from the children's general language difficulties.

However, Eisenbeiss et al. (2006) go much further than this, claiming that structural case does not even pose a problem to neither children with SLI nor young typically developing children. This claim could not be confirmed in the present study. The data presented here allows indeed some interpretations regarding the distinction of structural versus lexical case marking. The morphology probe of the present study for dative articles elicited lexical case marking due to the use of the verb *gehören* (engl.: *to belong to*). In agreement with Eisenbeiss et al's prediction, the correct production of the dative article was very difficult for all children, especially the SLI group (8%), the PI group (13%) and the ND-L group (14%). The accusative task, on the other hand, involved a direct object and thus structural case marking. Although the children with SLI were much more successful than in dative marking, an error rate of 40% could hardly be called ‚error-free' (Eisenbeiss et al., 2005, p. 35). Unfortunately, the data of the spontaneous language samples does not allow a distinction between structural and lexical case to follow this up further. Nevertheless, a significant delay on the case composite is unlikely to result by chance entirely from difficulties in lexical case marking.

Several methodological differences between the present study and the work of Eisenbeiss and colleagues could account for differing results between the studies. First of all, as intended, the participants with SLI of the present study were approximately 2-3 years younger than the children with SLI in Eisenbeiss et al.. A higher error rate in the present study could therefore be expected. This explanation, though, seems insufficient given the fact that even the age-matched typically developing children (who were slightly older than the MLU matched controls of Eisenbeiss et al.) presented an error rate of 9% in elicited accusative articles. The control children should rather have presented a similar ceiling effect on structural case marking as the children in Eisenbeiss et al..
Several differences with a presumably larger impact on the results can be found between both studies in the approach of analysis. In contrast to the present study, Eisenbeiss excluded case marking on indefinite articles as well as substitution errors that could result from low phonological perceivability, such as *–n* for *–m* substitutions. Eisenbeiss et al. argued that *–m* → *–n* substitutions (in masculine noun phrases evident as accusative instead of dative marking) do

not represent case errors unless feminine and neuter noun phrases are marked in accusative, too. This, however, is the dominant error pattern in dative described for typically developing children (Szagun, 2004, 2006). Feminine and neuter articles are much less often affected. Also nominative instead of accusative marking occurs the most frequently in indefinite rather than in definite articles (especially in masculine articles) in typical language development (Szagun, 2004, 2006). It is therefore likely that Eisenbeiss et al. excluded the most dominant error patterns from their analysis which may account for the ceiling affect in their participant groups. Due to the different intention, these errors were included in the analysis of the current study and only excluded if it would have been evident that the respective children had difficulties with these phonological contrasts. This was not the case and consequently no errors were excluded from analysis of the present study. In this way, the data analysis of the present study seems more representative of the typical language development than the analysis of Eisenbeiss and colleagues.

Looking at the error patterns described in the current study, one further question arises regarding Eisenbeiss et al.'s error analysis. Across all cases, omissions form a considerable proportion of errors in case marking. This applies especially to the SLI and the ND-L group – the two groups also involved in Eisenbeiss et al.'s study. Eisenbeiss and colleagues, though, did not mention this error pattern for their participants. It is however unlikely, that omissions occurred only in the present study but not in the language of Eisenbeiss et al.'s participants. Clahsen (1989, 1991) and Roberts and Leonard (1997) reported very high omission rates, too. Eisenbeiss et al. (2006, p. 21) present one utterance as example. In the present study, this sentence would have been analysed as omission of a dative article marking an indirect object (*zu Ente geben* (English: *to duck give* (*give to duck*))– correct: *der*$_{DAT}$ *Ente geben* (English.: *give (to) the duck*)). Eisenbeiss, however, scored this as an error in lexical case marking: The child inserted inappropriately the preposition *zu* (Engl.: *to*). The required case was still dative, but, normally, prepositions require the structural case accusative. The children need to learn lexically that the grammatical case associated with the preposition *zu* is dative rather than accusative. Because of this, Eisenbeiss et al. interpreted the sentence not longer as error in structural case but in lexical case marking although the preposition was inserted incorrectly. Moreover, the authors interpreted the sentence as realisation of the structural case accusative (substituting the lexical case dative), although no overt case marking was produced (but omitted). Taking all these differences in the approach of analysis together, differing results between the present study and Eisenbeiss et al.'s study do not seem surprising any longer. The high proficiency in case marking reported by Eisenbeiss et al. results rather from the exclusion of a number of error patterns rather than a high accuracy in the production of articles. Taking the present study and earlier research together, it can therefore be concluded that children with SLI experience considerable difficulties in case marking relative to their children's age matched peers.

#### *6.7.2.2 Delay in case marking as clinical marker?*

According to the results discussed above, case marking is vulnerable in German-speaking children with SLI. Even if this vulnerable spot in the children's language development does not represent a delay relative to the children's other language problems, case marking may be a valuable tool for the distinction between children with SLI and their peers of same age. The variables' potential as clinical marker for SLI will be discussed in the following section. The individual grammatical forms will be discussed first, followed by the case composite.

**Dative articles**

Only one individual grammatical form was significantly more difficult for the SLI group than for the group of age-matched controls across both measures: dative marking in articles. Dative articles were particularly difficult not only for children with SLI, who had an average accuracy level of 8 % in the morphology probes, but also for typically developing children. On the same measure, the age-matched controls produced on average only 46% of the definite dative articles correctly. The production of dative articles in spontaneous language, both definite articles only and independent of definiteness, was better (21% (SLI) and 60% (ND-A) independent of definiteness) but the difference between the SLI and the ND-A group reached nevertheless significance. A look at the individual data of the dative elicitation task reveals that in fact 10 out of 13 children with

SLI scored zero. The remaining three children of this group remained below 50% correct production. However, the advantage of the ND-A children is not large. Eight of the control children performed below an accuracy level of 40%, three of these children did not succeed at all in the production of the dative article. Thus, although a cut-off of 0% accuracy would result in an appropriate level of specificity (82%) a marker in dative articles would fail to identify a considerable number of children with SLI. Possibly, dative marking in articles would show more promising results if a larger number of items had been included in the elicitation task. This could result in smaller within-group variations, and a more refined cut-off due to a more pronounced advantage of the ND-A group. Unfortunately, the data from the language samples does not offer help either in this respect. The number of obligatory contexts for dative articles varied between 2 and 10 within the SLI group and between 3 and 19 within the ND-A group with a mean of only almost four and almost seven respectively. Nevertheless, the data from the language samples points into this direction that a higher cut-off could have some potential. The spontaneous data suggests a cut-off of 30% accuracy which results in an excellent specificity (100%). The sensitivity, though, is not satisfying enough possibly reflecting the small amount of data and also the considerable overlap between both groups Thus, although dative marking in articles may have some potential, the overlap between the groups is still substantial and the available data does not provide evidence that dative marking is suitable as clinical marker for SLI.

Another fact, that questions the appropriateness of dative marking in articles as clinical marker is that it would not be specific to SLI. The data from language samples revealed not only significant differences between the SLI and ND-A group but also between the PI and the ND-A group. This latter between group difference did not reach significance for the morphology probes due to a large variation within the PI group. However, Figure 12 illustrates that the PI and ND-A group can by no means be considered as performing at a similar performance level on elicited dative marking in articles either. A clinical marker tapping this linguistic form would consequently identify both children with SLI and a good proportion of children with PI in the absence of language problems.

**Accusative articles**

The second case morpheme that resulted in significant differences between the SLI and the ND-A group on the morphology probes was accusative marking in definite articles. The SLI group reached overall an accuracy level of only 60% whereas the ND-A group produced overall 91% of the accusative articles correctly. However, not all children with SLI experienced difficulties with definite accusative articles. Five children from the SLI group produced 80% or more correctly. The age matched control group, on the other hand, did not all perform above a threshold of 80%. Given this overlap of the groups, it is questionable whether the production of accusative definite articles could be a suitable clinical marker for German SLI. A follow-up investigation with a larger sample of children as well as more probe items may present more promising results, though.

Looking at the results of the language samples, this seems improbable. The mean group production accuracy for accusative articles varied little between the SLI (80%) and ND-A (86%) group. However, it is possible that the exclusion of ambiguous cases of accusative marking may have biased the results towards a higher accuracy. Szagun (2004) reported for typically developing children that accusative marking is more advanced in definite articles than in indefinite articles. It is noteworthy, that this pattern is reversed in the present study. The differences between the production accuracy was not statistically compared for definite versus indefinite articles but the mean percentages correct indicate that all groups were more successful in indefinite than in definite articles. Szagun explains the discrepancy between indefinite and definite accusative articles in favour of the latter with a low perceptual discriminability between nominative and accusative indefinite articles. The masculine indefinite article in accusative *einen* is often pronounced as *ein'n*. The only difference between this form and the nominative article *ein* for masculine and neuter nouns is a prolonged *n.* This difference is very subtle in contrast to the difference between nominative and accusative in definite masculine articles: *der* → *den*. The form of neuter (*das*) and feminine (*die*) definite articles does not change from nominative to accusative.

The present study cannot confirm Szagun's claim. However, it is probable that a higher proportion of indefinite than definite articles was excluded from analysis on the basis of ambiguity. Presumably, the low perceptual discriminability between *einen* und *ein* led frequently to an uncertainty which case had been produced. These instances were marked as ambiguous in the transcript and did not enter statistical analysis. As a consequence, a number of errors in indefinite masculine and possible even neuter articles may not have been recognised as such. A better performance in indefinite rather than definite accusative articles could be the outcome. The fact that the advantage of indefinite articles is consistent across all four groups in this study supports this suggestion. The exclusion of ambiguous cases affected all groups likewise. This explanation seems more plausible than all groups presenting an 'a-typical' weakness in definite accusative articles. However, if this assumption is correct it is possibly the SLI group that benefited the most of the exclusion of ambiguous cases. This again could account for the failure to find the same significant limitations in spontaneous accusative marking as in elicited accusative. Is accusative marking then nevertheless a suitable candidate as clinical marker for German SLI?

The results of the spontaneous language samples provide results for definite and indefinite articles separately. Surprisingly, it is the SLI group that performs the best on accusative indefinite articles. Even if these between group differences do not reach statistical significance, they support the assumption that the SLI group may have benefited unproportionally from the exclusion of ambiguous cases. Otherwise, it would be unexpected that the SLI group presented the lowest mean percentage correct on definite accusative articles in spontaneous as well as elicited language. Coming back to the consideration of accusative marking as clinical marker for SLI, we then should expect that the children with SLI fall significantly behind the ND-A children. This is not the case, although the mean percentages correct do point in that direction. The mean percentages correct suggest even another difference that would be desirable for a clinical marker: in both elicited and spontaneous definite accusative articles, the PI group presented higher levels of accuracy than the SLI group. These differences did not reach significance but may do so if it was investigated with larger groups and more test items. It may therefore indeed be worth while to investigate again the potential of accusative articles as clinical marker in a larger data set. In this case, it seems advisable to target only definite articles due to the facilitated discriminability of correct forms and errors. Nevertheless, two children of the SLI group responded to all accusative tasks correctly in this study, two further children made only one mistake. Presumably, these children would have presented a high level of proficiency on a longer elicitation task, too. Doubts remain therefore in how far an accusative task would provide the sensitivity for SLI that is desirable for a clinical marker.

**Personal pronouns in nominative**

The use of personal pronouns in nominative was not a variable in the morphology screening but met the criterion for analysis[17] for spontaneous data.

The ND-A exceeded the SLI significantly in the production of correct nominative pronouns. However, both groups demonstrated an advanced level of proficiency. The ND-A group performed basically at ceiling level but also the children with SLI had a mean level of accuracy of 81%. In particular, 10 out of the 16 children with SLI who had two or more obligatory contexts for a nominative personal pronoun produced these pronouns in more than 90% correctly. A look at the individual data explains the error patterns. The group as a whole substitutes 13% of personal pronouns in nominative. In this, the group seems distinct from the three control groups. However, only two children of the SLI group showed such substitutions. Child SLI1 and child SLI2 had not yet acquired the first person singular pronoun *ich* (Engl.: *I*) but they produced their own names instead. As they did so in 100% of their obligatory contexts for nominative pronouns they raised the group mean for substitutions considerably, although the remaining children showed production patterns comparable to the other groups. This therefore rejects nominative pronouns as clinical marker.

In how far the use of personal pronouns in accusative or dative may be an alternative cannot be judged. The number of obligatory contexts for these

[17] at least two obligatory contexts in the language sample of a minimum of eight children per participant group

morphemes did not meet the criterion for analysis. This suggests that methodological difficulties make the use of pronouns as clinical marker problematic. For accusative and dative pronouns, large language samples would be necessary in order to obtain an adequate data set. Elicitation tasks would be of little help. Morphology probes eliciting personal pronouns are difficult to design, independent of the grammatical case they are supposed to elicit. With the exception of first person pronouns, children could simply repeat names or noun phrases in order to avoid the pronoun. These methodological considerations taken together with the high proficiency of SLI children on nominative pronouns make personal pronouns inappropriate as clinical marker for SLI.

**Case composite**

Both measures revealed significant differences between the SLI and the ND-A group on a case composite. For this composite, the correct production of nominative, accusative and dative articles was combined into one score. The production of indefinite articles in the morphology probes was accepted as correct although the context would rather prompt a definite article. In agreement with this decision, the composite case derived from spontaneous language data comprises also both definite and indefinite articles. The most suitable cutoff point lies higher for the language samples (95.9% accuracy) than for the morphology probes (59% accuracy). This replicates findings presented in the literature demonstrating that free production facilitates the accuracy in contrast to elicitation tasks (e.g. Schöler et al., 1998). The discussion on accusative articles demonstrated, though, that the exclusion of ambiguous cases may have raised the spontaneous production accuracy for indefinite articles considerably, and thus the case composite derived from spontaneous production, too. This may account for the fact that sensitivity and specificity area higher for the elicited data. However, even for this data source the sensitivity (77%) and LR+ value (4.80) were not promising. This may change, though, if a larger sample of participants is tested with morphology probes comprising a larger number of elicited items for each grammatical case.

A look at the percentile scores for the children with SLI confirms the potential of elicited case marking. If the percentage scores are transformed into t-scores relative to the ND-A group, only one child performed on age-appropriate level. In spontaneous case marking, ten of the seventeen children with SLI fall below the 16$^{th}$ percentile. Five children with SLI, on the other hand, perform well above the 40$^{th}$ percentile and therefore within their age-appropriate level. As discussed above, the elicitation tasks seem to be more sensitive to SLI, presumably to some extent due to methodological differences regarding the definiteness of the articles.

Nevertheless, elicited case marking as clinical marker for SLI faces several challenges. First of all, task complexity is a disadvantage of the morphology probes. The four most impaired children of the SLI group could not perform on the tasks. It is unlikely, that the task complexity was too high in general for the children. None of the remaining participants refused to answer to the morphology probes, although some of the ND-L children were even younger than the four outliers of the SLI group. Some children did not reply to all of the items (especially regarding the dative probe) but the answers to other items of the same task indicated that the task complexity was well manageable for the rest of the children, including the younger ND-L children. It is noteworthy, that the measure tool used for language matching backs this up. The ND-L children were matched on language comprehension to the children with SLI. The claim that the task instructions were too difficult to comprehend for the outliers of the SLI group, but not for the ND-L children, can consequently be rejected. Nevertheless, case marking seems difficult to elicit from children at a very low language level. The diagnostic procedure of such children should therefore be complemented by language samples in order to decide why these children did not complete the tasks.

A second challenge to a case composite as clinical marker is that it may not be sensitive enough as indicated by the low sensitivity and LR+ value. This may possibly improve in a larger data set. Additionally, further research of case marking as clinical marker should take different factors into account that could influence the marker’s sensitivity, such as age or MLU.

However, a third restriction is represented in the fact that case marking in articles could not discriminate between children with SLI and children with PI. Hence, whereas the sensitivity of a case composite is already questionable in the classification of impaired children as such, it is certainly inadequate in the distinction between the two clinical groups employed in this study. This will be discussed further in section 6.8.

In summary, elicited case marking in articles faces several restrictions as potential clinical marker for SLI. A case composite may nevertheless be useful for the distinction between German-speaking children with SLI and their age-matched peers but further research is necessary in order to improve task design, to define accurately the target group and to replicate the findings with a larger sample.

## 6.8 Case Marking in German PI

The dominant purpose of including children with PI in this study was the comparison to the SLI group. A clinical marker for SLI should be specific to this impairment only, but not to another developmental disorder of comparable prevalence[18]. The selection criteria for this group included that the children performed within age expectations on a standardised language test. Given this, we should expect the children with PI showing similar levels of proficiency in case marking as their typically developing age matched peers. However, there are two variables from spontaneous language samples for which this prediction is violated: the children with PI produced dative articles in general and dative definite articles significantly more often incorrectly than the ND-A group. The prediction was further violated for one variable from the morphology probes: the case composite. Also a look at the mean percentages correct for the remaining elicited variables draws a picture of the PI group falling rather between the ND-A and the SLI group than presenting the PI group at a similar level as the ND-A group. The question arises therefore whether the children of the PI group had been identified correctly. We only can expect the children with PI being differentiated by a clinical marker for SLI if the children actually are distinct from the children with SLI. Do the children with PI have minor language deficits, too, which had not been detected by the standardised language test? Or do the children fall behind their typically developing peers on case marking for another reason than the children with SLI?

One of the most obvious explanations would be that the children with PI made their errors for other reasons than the SLI group, i.e. due to their speech impairments. This suggestion, though, can be rejected because such errors would have been excluded from analysis in order to avoid this mismatch of errors. None of the errors observed in the participants of this study could be a direct result from the child's speech problems. However, it is possible that the phonological receptive skills of the children hampered the acquisition of case marking as suggested by Chiat in her phonological theory (2001). Although this

[18] 3-10% of children show a speech output disorder (Fox, 2003; Shriberg et al., 1999)

theory has been proposed to account for the language difficulties in SLI, it should be applicable to children with PI, too. By definition, children with PI have underlying receptive phonological deficits rather than only speech output problems (Dodd, 1995). According to the phonological theory, the errors of the PI group should predominately affect those structures with both a low phonological salience and abstract semantic cues.

The children with PI fell particularly behind age expectations in dative marking. In fact, looking at the mean percentages correct rather than statistic results, the PI group produced even less dative articles correct than the younger typically developing children across both measures. The qualitative analysis of elicited data allows speculations on the developmental stage the children had achieved. As described above, typically developing children are reported to use accusative temporarily as default object case and overgeneralise accusative therefore to dative (Christina Kauschke, 2012). The analysis of error patterns in Table 27 indicates that this overgeneralisation of accusative is indeed the dominant error pattern across all four groups. However, while the SLI and ND-L group produced also nominative instead of accusative articles (i.e. category *case error*), this happened rarely within the PI group. The PI group seems therefore further advanced than the SLI and the ND-L group according to the typical developmental trajectory of case marking. Although the qualitative data analysis of the spontaneous language samples does not include the category *overgeneralisation accusative*[19], this assumption can be confirmed by the spontaneous data. Besides case errors, the children of the ND-L group showed a large proportion of *case and gender errors.* Overgeneralisations of the masculine accusative article *den* errors fell into this category if the children applied *den* to feminine or neuter nouns in dative. The children with PI, on the other hand, presented fewer of these errors, indicating that the children overgeneralised accusative but preserved the gender of the noun phrase in most of the contexts correct. Unfortunately, the error analysis does not allow

[19] While a clear quantitative criterion could be defined for the administered dative probes (at least 3 out of four feminine & neuter articles had to be produced in accusative), this was not applied to spontaneous productions because the latter error analysis did not distinguish between the three genders.

any conclusion on the distribution across the genders of the overgeneralisations of accusative. If low phonological salience causes the difficulties, masculine nouns (*dem*$_{DAT}$ → *den*$_{ACC}$) should be far more affected than neuter (*dem*$_{DAT}$ → *das*$_{ACC}$) or feminine (*der*$_{DAT}$ → *die*$_{ACC}$) nouns. Omissions are a frequent error pattern for the PI group, too. While this error does hardly appear in the ND-A group, it is a typical developmental error pattern as indicated by the data of the ND-L group and reported in the literature (Szagun, 2004). It is likely that this is related to the phonological surface characteristics of definite articles: they are monosyllabic and appear in unstressed pronominal position. Overall, children with PI show error patterns in dative marking that are in line with typical language development. They are more advanced than on average 10 months younger, typically developing children but they lag nevertheless behind age expectations.

The question remains why the PI group seems more delayed in the acquisition of dative than in the acquisition of nominative and accusative articles. The most obvious explanation is that this represents the developmental stage the children reached so far. The children with PI have passed the difficulties with nominative and accusative but still experience more difficulties in dative than their age matched typically developing peers. As mentioned, one alternative possibility could be that the children make unproportionally more errors with masculine than feminine or neuter articles. In this case, the low discriminability of *dem*/*den* could account for these difficulties. This error pattern would be the most expectable error pattern with regards to the children’s receptive phonological problems. Yet, at the same time, we should expect fewer omissions in this case. Szagun (2004) suggested that the late acquisition of dative marking may have its origin in the semantics of dative marking. Dative articles occur most of the time in noun phrases following prepositions. These prepositions convey the essential semantic information. Dative marking in articles succeeding a preposition seems therefore less essential than case marking in accusative noun phrases, i.e. direct objects. From the point of view of the phonological theory, this implies that the semantics in dative articles are more difficult to discover than in accusative articles because misunderstandings are less frequent than for example in accusative-nominative substitutions. This would account for the delay in dative marking rather than nominative or accusative marking.

The data of the spontaneous language samples is in agreement with this proposition. However, as indicated by the significant group difference between the PI and ND-A group on the composite for elicited case marking, the children with PI showed higher error rates in elicited case marking than in their spontaneous production. The significant group difference may result to some extend from the delay in dative marking, but the mean percentages correct tend to be lower for elicited accusative and nominative marking, too. It seems that children with PI are especially prone to the effect of accuracy facilitation in free production (e.g. Schöler et al., 1998a). This could have a number of reasons, for example the children’s awareness of the own limitations, but would need to be explored further. It can be concluded, though, that the children with PI are indeed at a more advanced language level than the children with SLI. The children’s production of articles seems nevertheless to be affected by their receptive phonological difficulties, especially in dative marking.

## 6.9 Error Patterns in Case Marking

Another possible approach to the identification of children with SLI could be the search for deviant error patterns rather than the search for a significant delay or delay within delay. In general, children with SLI of different language backgrounds are reported to make grammatical errors that resemble those made by typically developing children (Leonard, 2000). The error analyses in this study offer the opportunity to reassess this claim and to find indications for underlying deficits that may cause these errors. All conclusions rely simply on descriptive results, though. Differences regarding the children's error patterns were not compared statistically across groups. Conclusions can therefore be of only preliminary nature and would need to be investigated further in future research.

Comparing the children with SLI with typically developing children, thus the ND-A as well as ND-L group, there are five error patterns that seem to occur more frequently in language-impaired than typically developing children: omissions, case errors, filler words, gender errors and the use of indefinite instead of definite articles. These error patterns do not always occur more frequently across both measures or across all different grammatical cases. Below, each error pattern will therefore be described and discussed separately. The number of zero responses did not occur frequently and this error pattern was therefore not considered any further.

**Omission**

Looking at average proportions of error patterns, the difference between children with SLI and typically developing children seems the most striking regarding the frequency of omissions. Across all three grammatical cases, children with SLI appear to omit articles more frequently than the control groups. In nominative, this was the case for both measures, in accusative for the morphology probes and in dative especially for the spontaneous production of articles. Different questions arise from these observations. First, are the children with SLI in this respect indeed distinct from typically developing children? Second, why did the different methods result in different results regarding the omission of articles? Third, are there theoretical approaches to SLI that could account for the large number of omitted articles?

Regarding the first question, Table 38 provides an overview of the average omission rate per child within the SLI group and both groups of typically developing children. At least in comparison to the age-matched control children, the omission rate of children with SLI appears notably higher, with the exception of the spontaneous production of accusative articles. However, the significance of these differences needs to be investigated statistically (and replicated) in order to answer the first question, i.e. in how far omissions are a distinct error pattern of children with SLI.

**Table 38.** Average Omission Rate of Articles

| ***Case*** | | ***Group*** | | |
|---|---|---|---|---|
| | | **SLI** | **ND-A** | **ND-L** |
| **Nominative** | | | | |
| | Elicited | 12% (*n*=2) | 2% (*n*=2) | 1% (*n*=1) |
| | Spontaneous | 24% (*n*=10) | 1% (*n*=1) | 5% (*n*=5) |
| **Accusative** | | | | |
| | Elicited | 10% (*n*=2) | 1% (*n*=1) | 1% (*n*=1) |
| | Spontaneous | 7% (*n*=6) | 7% (*n*=9) | 8% (*n*=6) |
| **Dative** | | | | |
| | Elicited | 15% (*n*=5) | 2% (*n*=2) | 17% (*n*=5) |
| | Spontaneous | 25% (*n*=7) | 2% (*n*=2) | 14% (*n*=5) |

*Note.* Percentages are rounded.
SLI = Specific language impairment group; PI = Phonologically impaired group;
ND-A = Normally developing, age-matched group;
ND-L = Normally developing, language-matched group;
*n*=number of children per group who show one or more omissions

Given that the SLI group falls significantly behind the ND-A group in the general accuracy of spontaneously produced nominative articles and that most errors in spontaneous nominative are omissions, it is likely that the difference regarding the omission rate on this measure would be statistically significant, too. The difference regarding the remaining variables is difficult to judge as the SLI group showed various error patterns in each of them. The omission of articles would be useful as clinical marker only if it can be shown to be an error pattern unexpected for the typical language development but at the same time typical for children with SLI. The latter has been described earlier in the literature. Leonard (Leonard, 2000) reported that children with SLI produce grammatical morphemes rarely inappropriately but rather tend to omit these. Also the literature review above demonstrated that omissions of articles have been described as frequent error pattern in German SLI (Clahsen, 1989; Roberts & Leonard, 1997). Nevertheless, the omission of articles occurs in the typical acquisition of German, too (Szagun, 2006; Weissenborn, 2000). As mentioned above, Szagun reports that noun phrases miss the article in early language acquisition. Weissenborn (2000) calls this phenomenon “das Prinzip der minimalen Struktur” (p. 159), i.e. ‚the principle of minimal structure’. According to this principle, children tend to avoid those functional language structures that are intellectually still too complex for them. The considerably higher omission rate of the ND-L group in dative than of the ND-A group supports this claim. This suggests that the omission of articles is indeed a prominent error pattern as long as the children did not fully grasp the grammatical rule. The two further stages of case development described by different authors (Christina Kauschke, 2012; Schrey-Dern, 2006; Szagun, 2006), the overgeneralisation of nominative and accusative, refer to accusative and dative contexts only. Consequently, in nominative contexts, omissions are the error pattern to expect. It is therefore little surprising that children with SLI showed this error pattern, too. Thus, even if the difference between the groups turns out to be statistically significant in nominative, the rate of omissions does not appear suitable to differentiate reliably between children with SLI and typically developing children. The high omission rate of articles of the children with SLI is simply a delay in language development but no sign of a deviant development.

This claim is further supported by a closer look at the group data. Table 38 provides the number of children per group who show one or more instances of omissions for each variable. This demonstrates that not all children with SLI can be characterised by the omission of articles. This overview demonstrates further that the number of children within the SLI group exceeds that of the younger typically developing children only considerably for nominative. However, as shown above, it is especially in nominative that young typically developing children show omissions of articles rather than other error patterns. Nevertheless, since the children with SLI fall here even behind typically developing children matched for language comprehension, longitudinal data might reveal a developmental stagnation that is characteristic for young children with SLI (see discussion in section 7.4). Thus, the omission of articles, especially nominative articles, might temporarily be distinctive for children with SLI. Yet, an age threshold would need to be identified for diagnostic purposes.

The second question arose regarding the differing results across both measures. A look at Table 38 illustrates that more children tend to omit articles in spontaneous language than in elicited language. This seems more or less the case across the groups. However, the average omission rate per variable is more inconsistent than the number of affected children across the measures. For the SLI group, though, the trend that spontaneous production is more prone to omissions than elicited production seems obvious in nominative and dative. But also in accusative, more children omit articles in spontaneous language than in the accusative probes. What could be the underlying reason for these methodological differences? Comparing the productions, one major difference seems to be the utterance length. In general, the elicitation task required the children to fill in only two words, i.e. the noun phrase containing the determiner and the noun. In spontaneous language, the utterance length was far more variable. This could have increased the processing load the children had to deal with resulting in a larger number of omissions. Given that especially children with SLI are said to have processing limitations (e.g.Kail, 1994; Leonard, 1989) this could explain why it is mainly the SLI group that shows more omissions in spontaneous language than in the elicitation tasks.

The third question raised above regards theoretical underpinnings of SLI accounting for the large number of omissions of articles. As concluded above, the large number of omissions seems to be a phenomenon characteristic for typical language development, too. However, the present study demonstrated that the children with SLI seem to fall even behind younger typically developing children in their degree of article omissions. What could be the reasons for this? One of the explanations could be that the children with SLI are simply delayed in their language development and omit the articles generally for the same reasons as typically developing children. The above mentioned ‘principle of minimal structure’ by Weissenborn (2000) would be one approach. However, there are also a number of theories specific to SLI that could account for the large number of article omissions. One possibility, for example, is the implicit rule deficit account by Gopnik and colleagues (Gopnik, 1994, 1997; Gopnik & Crago, 1991) claiming that the children are not able to construct certain grammatical rules. In this case, though, we should expect omissions not only to be the dominant error pattern but also the exclusive one. This is not the case as Table 35 illustrates – especially in accusative and dative marking.

Another explanation suggested by Gerken (1994) is the influence of stress patterns within utterances. Gerken proposed that unstressed syllables occurring in prestress position are more vulnerable than unstressed syllables that immediately follow a stressed syllable. In German, articles are generally unstressed whereas the following noun begins most of the time with a stressed syllable[20]. Articles occur thus in German most of the time in prestress position. Regarding the present data collection, the stress pattern was not considered in particular. However, the elicitation tasks allow a distinction. The nominative task involved only the noun phrase, representing the pattern weak-strong syllable. The accusative task, too, involved this stress pattern. The demand *Gib mir…* *(*English: *Give me…)* preceding the noun phrase carries its stress on the first word *gib* whereas *mir* is unstressed. In dative, though, the noun phrase is directly preceded by a stressed syllable (*Der Kopf gehört …*: weak-strong-weak-strong). According to Gerken, it is thus the dative context that should stimulate the production of the article more than in accusative (and possibly nominative, although the latter does not involve any preceding syllable). However, the present results show the opposite pattern. Also children with SLI demonstrated much less omission of accusative articles than dative articles. It is therefore unlikely that the stress pattern had a major impact.

[20] The trochee is the most dominant stress pattern in German bisyllabics, i.e. the stress is placed on the first syllable.

However, as discussed for the PI group, theories based on weaknesses in the children's phonological processing, such as the surface account (Leonard, 1989, 1998; Leonard et al., 1997; Leonard, McGregor et al., 1992) or the phonological theory (Chiat, 2001) are more general and could indeed account for the large number of omissions observed in the production of articles. Hsu and Bishop (2011) suggested recently that a deficiency in statistical learning hampers the acquisition of grammatical rules in children with SLI. This account, too, seems feasible given that the children with SLI in general show a developmental pattern by omitting articles, although it would be necessary to consider input rates in order to assess this suggestion further. These input rates are unfortunately not available for the present data. The question of possible theoretical underpinnings of the grammatical symptoms in German SLI will be followed up in greater detail in the general discussion in Chapter 7 in combination with error patterns from other variables.

**Case errors**

The choice of a wrong case occurred more frequently in children with SLI than in typically developing children in nominative and dative. In nominative, case errors occurred across the groups almost exclusively in the elicited production. The children with SLI showed on average three times more case errors (9%) on the nominative task than the typically developing children. Interestingly, the PI group fell in between the control groups and the SLI group. In both groups three children showed case errors in nominative but only one child of all typically developing children. Case errors in nominative articles have not been described in the literature on typical development of German (Christina Kauschke, 2012; Szagun, 2004). However, although the current data points into the same direction, the conclusion seems premature that any child showing deviant case errors in nominative articles should be classified as SLI or PI. Also one child of the ND-A group showed case errors in this context, although this child met neither the criterion SLI nor PI.

All case errors in nominative resulted in the production of the masculine accusative article *den*. This suggests that the children showing this kind of error were at a more advanced stage of development. It is possible that the children had originally acquired nominative articles correctly but that they overgeneralised the accusative during accusative acquisition. Both feminine and neuter articles have the same surface form in nominative and accusative which explains why only masculine articles are involved in this overgeneralisation. Children may have drawn the conclusion that the masculine article has the same surface form across both cases, too. However, three children overapplied the masculine accusative article also to neuter nouns in nominative. Interestingly, though, this overapplication did not affect feminine nouns. A possible explanation could be the phonological alikeness of the substitute *den* to the neuter dative article *dem*. Although the children disregard the difference *n / m* they preserve the genus neuter but overapply dative rather than accusative to nominative. The lack of differentiation between the wordfinal *n* and *m* may also account for the slightly higher frequency of case errors in the PI group. However, these suggestions lack substance due to the very limited data. Longitudinal data would be necessary in order to investigate trajectories in nominative marking across the different populations to allow more substantial conclusions. For now, case marking errors in nominative marking should result in more detailed language testing in order to detect any possible language or speech impairment.

In dative articles, different case errors were categorised accordingly: the overgeneralisation of accusative to dative, the overgeneralisation of the masculine accusative article *den*, case & gender errors, and remaining case errors (i.e. the overgeneralisation of nominative). For spontaneous language, the distinction was made only between case error and case & gender error. Overall, the error analysis showed a rather similar picture for the SLI and ND-L group. Especially in elicited dative marking, only a slightly higher rate could be observed in the production of nominative articles for the SLI group (SLI: 24%; ND-L: 20%). In comparison to the ND-A group, though, this difference is very large because the older control children hardly produced any nominative in the context of dative. This indicates a developmental delay within the SLI group but no developmental deviance. In spontaneous language, the SLI children produced on average a fifth more of case errors, although this could be due to statistical chance. The frequency of case & gender errors resembled that of the

ND-A children and was only half of that of the ND-L group. Consequently, the production patterns in dative marking do generally not point to distinct case errors in dative marking but towards delayed developmental patterns.

In accusative marking, typically developing children showed slightly higher average frequencies of case errors than the children with SLI. This highlights the caution that should be applied when interpreting minor group differences without employing statistics. The results discussed above could simply be due to chance - probably with the exception of the high proportion of nominative overgeneralisation in elicited dative contexts relative to the ND-A group. The data presented here does therefore indicate a developmental delay regarding the frequency of case errors. There is furthermore the possibility that children with SLI tend to overapply accusative and possibly dative more frequently to nominative articles than typically developing children. Especially the overgeneralisation of dative to nominative may be triggered by an underlying phonological deficit. However, more and longitudinal data is necessary in order to draw conclusions on the children’s developmental trajectories and differences between these, regarding the different groups.

**Filler words**

If children produced a substitute such as [*də*] across different genders and/or cases these were categorised as *filler words.* In nominative articles, these filler words hardly occurred in any of the participant groups. In accusative and dative marking, on the other hand, the children with SLI produced this sort of error more frequently than typically developing children. This seems especially true for elicited data. A look at the children’s individual data, however, illustrates that these are single cases. In the probe data for accusative marking, it is only one child with SLI that used a filler word instead of a correct article. As this boy did this consequently, though, his productions resulted in a proportion of 8% of all productions that were realised as filler words. In elicited dative marking, three children used filler words. They did not use them consequently which explains a lower rate of filler words in dative (6%) than in accusative. Nevertheless, this illustrates that filler words occur occasionally in some children with SLI but cannot be described as typical error pattern for German-speaking children with SLI.

**Gender errors**

The children with SLI showed gender errors of noteworthy frequency only in accusative marking: on average 17% in elicited data and 9% in spontaneous language. These frequencies exceeded those observed for the ND-A group considerably (8% and 3%) but were comparable to those observed for the ND-L group. It is therefore likely that the gender errors represent a developmental delay in accusative marking rather than a distinct production pattern.

**Indefinite articles**

The use of indefinite articles in the context of definite articles was only analysed for the probe data. The tasks were identical across all groups and designed to elicit definite rather than indefinite articles as the noun had been introduced at the beginning of the task and each item occurred only once in the task. Comparing the four participant groups, it is especially in nominative that the typically developing children picked up more easily the pragmatic cues that stimulated the definiteness than both the children with SLI and with PI. Both clinical groups produced indefinite articles in about one fifth to one fourth of the test items. In accusative marking, this was a problem only for the PI group. Although the substitution of definite articles was documented for the probe data this was not counted as error in order to maintain the comparability between the two measures elicitation tasks and spontaneous language samples. While the context within the elicitation tasks obliges clearly definite articles it is hard to determine unambiguously in spontaneous language whether the context requires definite or indefinite articles.

Facing this problem when analysing spontaneous data, most researchers combined definite and indefinite articles in their analyses. One exception is a recent study by Polite and Leonard (2011) who employed elicitation tasks. The authors investigated the use of definite and indefinite articles in English-speaking children with SLI aged 4;5-7;0 and in two control groups matched for chronological age and MLU respectively. For the elicitation of definite and indefinite articles, the children were read short stories and asked questions by the examiner. In the definite articles condition, the target had already been mentioned in the story. The results showed that the children with SLI were less successful in using definite articles than both control groups whereas no significant difficulties were apparent in the use of indefinite articles. The majority of errors were substitutions, thus the children produced the indefinite instead of the definite article. However, also about one third of the errors were omissions. This latter error pattern cannot be compared to the data of the present data. Nevertheless, the high proportion of substitution errors reported by Polite and Leonard support the results of the present study: Children with SLI seem to have difficulties in the appropriate distinction between definite and indefinite articles.

It is important, though, to keep in mind that the two studies investigated different languages. Since case marking does not appear at the surface of English articles, Polite and Leonard do not consider the grammatical cases of the test items as important. The experimental task items are provided in Appendix by Polite and Leonard. Two of the items in the definite article condition refer to nominative and six items to accusative. In the indefinite article condition, none of the test items refers to nominative but three to accusative and five to dative. Regarding the results of the present study, these methodological differences could have had an influence on the results. In the present study, nominative is the only grammatical case for which the children with SLI demonstrated problems with definite articles. Accusative was further difficult for children with PI, but none of the children used indefinite rather than definite articles in the dative task of the present study. Given the low number of nominative items in the study of Polite and Leonard, this could account for only about one third of the errors observed. Nevertheless, the high proportion of dative items could have influenced the results of the indefinite article condition positively. In this respect, it would be interesting to know how often the children included the preposition *with*[21] in their response. It may be possible that the production of a preposition preceding the noun phrase enhanced the correct production of the article. As only the indefinite article condition included dative phrases this would have facilitated only the production of indefinite articles. However, this explanation cannot account for the lack of substitution errors in the dative task of the present study. The grammatical case dative was prompted by a verb triggering the lexical case dative but not by the inclusion of a prepositional phrase. It seems therefore necessary to consider the grammatical as well as phrase structure of test items for interpretation.

Differences in both the target languages and in task design do not allow a direct comparison between the results of Polite and Leonard and the present study. However, Polite and Leonard propose different possible explanations for their results which may also be applicable for the results presented here. One of the potential aspects discussed by Polite and Leonard are possible memory factors. In contrast to the context for indefinite articles, speakers have to be aware of the fact that a referent is already known to the listener in order to choose the definite article. Memory factors may impact on the retention of an earlier reference. Polite and Leonard, though, came to the conclusion that these memory factors cannot be the sole explanation of the problem with definite articles. The participants in their study often used the previously mentioned target nouns but made errors in the selection of the article. In the present study, a major impact of memory factors seems even less likely. The tasks of the present study did not only involve that the referents were mentioned in the introduction of the task but secondly, the focus by speaker and listener on the referents was further facilitated by the use of pictures and the reference to these. These pictures should enhance the retention and the choice of a definite rather than indefinite article. Nonetheless, one finding points indeed in the direction of a retention deficit. The children with SLI showed difficulties with the distinction between definite and indefinite articles only in nominative. The order

[21] All experimental test items in dative by Polite and Leonard (2011) included the preposition *with* in the question eliciting the children’s response (e.g. ‘Sue wants to write a letter. What can she write the letter with?’)

of the tasks could have been of influence here. While the same target items were used for both tasks, it was the nominative tasks that preceded the accusative tasks. When the children performed the accusative task, the referents had thus been mentioned at least twice before, but only once when the children performed the nominative task. The retention was consequently additionally stimulated for accusative. This again may explain why the children did not experience difficulties with the selection of the definite article in accusative.

Another possible explanation, though, for the difference between nominative and accusative in the present study could be the pragmatic context. In the nominative task, the children had to answer to the question: '*What goes here?*' with the examiner pointing to a picture. The children had to select and name the correct small picture that had to be placed on the larger picture board. In the accusative task, on the other hand, the children had to address a puppet and give him an order: '*Give me the…!*' The fact, that the children addressed the listener directly may have stimulated the choice of the definite article more than the rather passive reply to a question. Schafer and de Villiers (2000) demonstrated that children as well as adults apply definite articles to differing degrees depending on the pragmatic context. However, there is no literature known investigating the choice of German articles in similar pragmatic contexts. Schafer and Villiers suggested furthermore that, in early language acquisition, the use of indefinite articles does not represent a determiner phrase but rather a number phrase. This would be applicable to German, too, as the indefinite article corresponds to the quantifier *one.* The definite article, though, does require a determiner phrase according to Schafer and Villiers and involves further that the children take the point of view of the listener in order to be aware of his knowledge. The authors suggested that both features are difficult for young children. Regarding the pragmatic differences between the nominative and accusative task of the present study, it is likely that addressing the puppet may have stimulated more explicitly to take the point of view of the listener. This would account for lack of difficulties with definite articles in accusative.

Nevertheless, both explanations cannot account for the fact that the PI children presented difficulties with indefinite articles in both grammatical cases. However, as this study is concerned with SLI rather than PI this issue is not followed up here any further.

None of the participant groups struggled particularly with the choice of definite articles in dative marking. A large proportion of the children, though, showed grammatical errors in dative. It seems therefore probable, that the children had more severe difficulties with dative marking than the distinction between definite and indefinite articles. Moreover, those children who got dative marking right, did not present difficulties with this distinction either. It seems likely that this is a developmental course. Once the children have acquired dative marking they have acquired the distinction of definite and indefinite articles, too.

To sum up, German speaking children with SLI show difficulties with the correct selection of definite rather than indefinite articles, especially in nominative. Retention deficits as well differences in task design may have contributed to these results. Further research is necessary, though, in order to replicate differences between the grammatical cases in this respect and to investigate the influence of retention and the degree of theory of mind.

## 6.10 Conclusion

The data presented here allows a number of conclusions. In agreement with previous research, German-speaking children with SLI presented considerable difficulties in case marking, falling significantly behind their age-expectations in nominative, accusative and dative as well as in a case composite. The spontaneous production of nominative articles is even significantly lower than that of language-matched control children. However, longitudinal research would be necessary in order to identify a potential time frame during which nominative articles could be used as clinical marker providing not only a satisfying specificity but also sensitivity. This could possibly also improve the discriminative power of such a marker using morphology probes although further research involving larger groups and a larger number of test items would be needed. So far, it is especially the discrimination between the SLI and the PI group that remains questionable using elicitation tasks for the production of nominative articles.

Regarding production patterns in case marking, there are a number of errors that should raise the alertness of any clinician. Omission was the most dominant error pattern in the production of articles by children with SLI. Although omissions occur frequently in typical language development, too, this error pattern was particularly rare in nominative articles even in typically developing children with an average age of 3;3. Also case errors in nominative and the use of filler words instead of articles should result in a detailed diagnostic procedure. Case errors in accusative and dative as well as gender errors in accusative, on the other hand, represent a developmental delay as they occur in both typical and impaired language development.

# 7 A Clinical Marker for SLI in German?

The primary intention of the presented work was the identification of a grammatical clinical marker for SLI in German. Two measures were employed in order to identify those grammatical structures that have the greatest potential for a clinical marker: morphology probes involving elicitation tasks for verb morphology and articles as well as spontaneous language samples. In addition to two groups of typically developing children, a group of children with PI was included in the current study. This inclusion allowed the investigation of the overlap between clinical groups as well as their distinction and the possible impact of speech difficulties on the children's morphology.

## 7.1 Is There a Clinical Marker for SLI in German?

In summary, the data from the present study does not reveal a useful clinical marker for SLI in German-speaking children. The overlap between the SLI group and the ND-A group is, for most variables, considerable and results in insufficient categorisation accuracy. The only variables that resulted in satisfactory levels of sensitivity, specificity, LR- and LR+ values are two composites referring to verb inflections. However, both markers are based on very high cut-offs in proficiency and they are not robust to minor changes in the analysis procedure. Also, the analysis of error patterns failed to reveal a valid marker for SLI. The majority of errors made by German-speaking children with SLI mirrors typical language development. Some exceptions did occur but they were not characteristic for all children with SLI. Thus, according to the presented data, there is no grammatical deficit in the language of German-speaking children with SLI that stands out as a potential marker for SLI.

Furthermore, also the comparison between the SLI and PI group revealed few significant group differences. At first glance, both groups do not seem distinct of each other. A closer look, however, indicated that although children with PI produced grammatical errors that could not be explained by their speech difficulties alone they seem more advanced in grammatical development than the children with SLI.

The chapter below discusses three questions putting the presented findings in a wider perspective:
First, what are the methodological factors that could have impacted on the presented results? Second, how can the findings of the present study further our understanding of the relation between SLI and PI? Third, how shall the category SLI be considered in future and what are the implications of this for future research of primarily language impaired children?

## 7.2 Methodological Factors

At first sight, these results may be surprising and they raise questions as to whether a suitable methodology was employed for the study. The procedure followed in this study may have its flaws and a number of decisions had to be made which are disputable. Nevertheless, it is crucial to evaluate whether these methodological issues can account for the outcome of the study and whether a different study design would have resulted in the identification of a clinical marker in the children's grammar for SLI in German.

### *7.2.1 Measurement tools*

Two different measures were employed in order to analyse the children's use of grammatical morphemes: the morphology probes involving elicitation tasks and spontaneous language samples. These two methods were intended to be complementary; an intention that was confirmed from the presented data. In general, elicitation tasks are often perceived as an unnatural way of language data collection but they also have been proven to be more reliable to unearth data on a linguistic structure of interest than spontaneous language sampling (e.g. Dromi et al., 2003; Schöler et al., 1998). Indeed, with respect to the use of 2$^{nd}$ person singular verb inflection, for example, no conclusions could have been drawn from the present study if only the spontaneous data was available. The elicitation task provides an obligatory context for a certain morpheme to *every* child and does not leave this up to 'chance' or to the child's avoidance strategy. On the other hand, children produced most morphemes at a higher accuracy level in spontaneous language than in the probe data. Most evident is this for the three cases of SLI with the weakest language profile. Whereas most of the elicitation tasks were far beyond their language ability and/or were pragmatically or developmentally inaccessible to them, the children did produce obligatory contexts for some of these morphemes in spontaneous language. This indicates that the data of the language sample is not redundant as it was found to be by Schöler et al. (1998) with older children. Instead, both measures complement one another and this approach can be considered as strength of the present study.

**Comparison of measurement tools**

Overall, the results were generally consistent across the measures although some differences were apparent in specific conditions. Third person singular inflection as well as the production of accusative articles was significantly more difficult for the SLI group than the ND-A group in the elicitation tasks but not in spontaneous production. These results are not surprising because elicitation tasks have previously been shown to be more difficult than spontaneous production (e.g. Rice & Wexler, 1996; Schöler et al., 1998), a difference which is less pronounced in age-matched, typically developing children than in language impaired children (Rice & Wexler, 1996). The fact that the SLI group is, but the ND-A group is not affected by the choice of measure in the production of 3rd person singular inflection and accusative articles explains why a significant difference between both groups is only evident in the probe data.

The production of dative articles is the only variable for which the ND-A group, too, shows a weakness in elicited production relative to spontaneous production. Both the SLI and the ND-L group perform significantly below the ND-A group on the production of dative articles irrespective of the measure tool. However, the difference between the PI and the ND-A group reaches statistical significance only for the language sample. The difference between group means is larger in the probe data, though. This appears to be explained by the larger within group variance for the ND-A group in the probe data than in spontaneous data. The lack of a significant difference between the ND-A and the PI group in the probe data on dative articles is consequently likely to be due to this larger within group variance of the ND-A group. A similar reason accounts for the lack of a significant difference between the SLI and ND-A group in elicited nominative articles. Although the mean percentages correct are for both groups very much alike across measures, one measure results in a significant difference whereas the other does not. A larger variance within the ND-A group may account for this discrepancy.

In spontaneous language, a large number of items were excluded from analysis if they could not unambiguously be identified. As discussed in chapter 6, this may have had an impact on the results for accusative marking in particular.

However, it is difficult to justify any other approach for data analysis if the data does not provide sufficient information on the target structure intended by the child. Similar considerations can be taken into account for missing data in the probe data, especially the gaps in the responses of the most severe cases with SLI. The tasks were excluded from analysis if the children did not attempt the completion of the task or, in the case of verb inflections, if the response did not involve a verb. This approach, though, had the potential to raise the child's estimate of the performance accuracy. For example, if all responses of a child involved a zero or ambiguous response they were regarded as missing data. If a child, on the other hand, produced for example two zero responses in addition to two responses with correct verb inflection, the child scored still 100% correctly for this respective inflection. The first two responses were excluded from analysis as it was unclear whether the child actually intended the production of a verb in present tense, whereas the remaining responses *with* the target structure were then considered as the total of attempts. It is therefore possible that the results of the morphology probes present a more advanced picture than the reality. An alternative approach could have been to treat all these missing responses as errors. Such an approach would certainly have pulled down the performance of the SLI group considerably and would possibly even have resulted in the identification of a clinical marker. Nevertheless, the design of the present study does not justify this approach. A non-attempt can occur because of all sorts of reasons besides a lack of linguistic proficiency. For example, the child could be tired, distracted or bored, or have difficulties with non-linguistic task demands. A much larger amount of data, preferably even longitudinal, is consequently required in order to answer the question whether a certain structure is present or not in a child's language production.

In summary, although the two different measures result for some variables in different findings, these discrepancies can be accounted for by task nature and a loss of statistical power due to large within group variations. Both measures provide valuable and complementary information and are overall consistent.

**Quantity of data**

One aspect that could certainly have been improved in this study is the quantity of data. The more data is collected the more representative the picture that is intended to be drawn from the real world. However, a line needs to be drawn somewhere to remain the balance between costs and resources.

The aim of this study was to obtain the most meaningful results on the basis of the available resources. With respect to the investigator, these resources involved time and knowledge for participant recruitment, data collection, transcription of language samples, data analysis and data interpretation. With respect to the participants, resources involved the time for testing (including the absence from nursery programme or other activities), the children's attention span and cooperation, the parents' willingness to give their consent and partly their initiative to accompany the child to the clinic for speech and language therapy as well as the cooperation of the institutions such as the nurseries or the clinics for speech and language therapy to provide assistance and rooms. It is thought, that the present project provides the best possible compromise between resources and the intention to answer the research questions. Nevertheless, some of the decisions taken are discussed in the following.

With regards to the language sample, the most important factor that could be controlled was a time limit which needed to be set. A sample of 20 minutes has been recommended in the literature (e.g. Dunn, Flax, Sliwinski, & Aram, 1996; Gavin, Klee, & Membrino, 1993) and was applied in the present study. On average, a 20 minute sample corresponded to 202 complete and intelligible utterances per child. Similar guidelines regarding duration length or number of utterances were used in several studies (e.g. Hansson & Nettelbladt, 1995; Klee, Gavin, & Stokes, 2007; Rice & Wexler, 1996). However, whereas some studies relied on even shorter language samples of 100 utterances only (e.g. Beverly & Williams, 2004) many researchers try to use much larger samples as data source (e.g. Leonard et al., 2004; Miller & Leonard, 1998; Paradis & Crago, 2001). Although long language samples are certainly preferable, it needs to be pointed out that these large data sets stem frequently from corpora.

This option was not chosen for the present study because a large number of studies on German SLI draw on data corpora with data often collected decades ago (implying for example less valid selection procedures) (e.g. Eisenbeiss et al., 2005; Hamann et al., 1998; Rice et al., 1997; Roberts & Leonard, 1997; Rothweiler et al., 2012).

One advantage of the analysis procedure employed in the present study is that the current project did not intend to provide any information about the course of language acquisition. Consequently, no conclusions have been drawn from the information whether a certain linguistic structure was absent or present in the language transcript. In this case, the length (and frequency) of language recording would be far more relevant. In the present study, however, the lack of obligatory contexts for a certain morpheme resulted in missing data and no conclusion could be drawn from that. As a result, shorter language samples provide information on a lower number of variables. No information could be obtained, though, about the children's general ability or inability of using certain linguistic forms as long as the child did not create a minimum of two obligatory contexts respectively. However, although structures were dismissed in analysis if obligatory contexts occurred only once in the language sample, proportional scores are nevertheless influenced by the length of the transcript. The weight of each obligatory context decreases with the increase of the total of obligatory contexts and it is obviously more likely to find a larger number of obligatory contexts the longer the language sample is. Proportional scores entail therefore that the weight of individual productions vary across transcripts depending on the number of obligatory contexts. Whereas one child achieves a proportional score of 50% correct with one correct and one incorrect production, another child may score at 90% correct with a similar number or errors but nine correct productions. Consequently, the length of the transcripts is a very important factor in order to influence and minimise this measurement error. Looking at the mean of occurrences of obligatory contexts across the different morphemes, it would have been preferable to have longer language transcripts. It is very likely that this would have reduced the within group variation and therefore increased the statistical power of the data.

With regards to the morphology probes, two factors could be controlled: the number of variables tested and the number of items for each of these variables. The variables were chosen on empirical and theoretical grounds. Furthermore, the results of the language samples did not suggest any morphemes which should have been additionally included in the elicitation tasks. However, the number of items per variable (4-6 depending on the task) is rather limited. A larger number of items would probably have resulted in overall higher proportional scores because errors would have been proportionally less impacting. More importantly, though, within group variations would have been smaller with a larger number of items and this would probably have increased the statistical power of the data. Nevertheless, the data showed clearly, that there was always a good proportion of the SLI group performing at a high level of accuracy. These scores would have been as high or even higher if the number of items were increased. It is therefore evident that it is not the limited number of items that prevented the identification of a clinical marker for German SLI but the lack of one distinct weakness in the children's morphology.

### *7.2.2 Participants*

**Sample size**

Another approach to increase the quantity of data could be to increase the number of participants. A replication of the study with larger samples would probably reveal more statistically significant differences between groups than the current results. The reason is similar to the effect of larger quantities of data per child: the within group variation would most likely decrease and the statistical power therefore increase. More importantly, the larger the sample size is the more representative is this sample for the population. Is it therefore necessary to replicate the current study with more children in each group before we can draw conclusions about a clinical marker in German SLI?

This depends on what we expect from such a clinical marker. The expression 'clinical marker' implies that it should be applicable in clinical context. However, a clinician is usually confronted with diagnosing individuals rather than whole groups. Thus, a reliable clinical marker should mark nearly *every child*, not only the *'average child'* with SLI. Even if the current SLI group is only a small sample of the population of German-speaking children with SLI, each of the participants is an individual representative of this population. The outcome that this small sample is too heterogeneous for the identification of a clinical marker is notable in itself. It implies that a clinical marker that is based on the mean performance of a large sample would be of little use in clinical context. Applied to the 17 children with SLI in the current study, a clinical marker would not have identified all of these children. Possibly, a number of the children of the current SLI group were unusually proficient in their morphology. In this case, a larger sample would pull down the mean percentages correct for some morphemes. It would be difficult, on the other hand, to call the weaknesses of a large sample as *characteristic* for SLI if the mean accuracy is above 80% in the current sample. From this point of view, one can postulate that a study with a larger sample would be more likely to identify a clinical marker but at the same time it would bear the risk to mask the population's heterogeneity. It is therefore essential to focus on smaller samples and single cases as well as on large, representative studies. In this sense, the present study can be considered as pilot study. If the outcome had revealed a promising candidate for a clinical marker the replication with a larger sample would have been the necessary following step. The fact, though, that children of the present study were proficient in different areas of their morphology indicates that the continued search for a clinical marker is unlikely to reveal a morphological deficit as being characteristic for German SLI.

**Language deficit**

Nevertheless, before accepting this conclusion it is necessary to re-evaluate the participant selection. The conclusion that German-speaking children with SLI cannot be characterised by a clinical marker is only valid if the groups of children are representative for their respective population. Most importantly, do the children in the SLI group have SLI? As the literature review on the identification of SLI demonstrated, there is no straightforward answer to this question. Inclusion and exclusion criteria were employed in the participant recruitment. Given the fact that the SLI group of the present study performed at a rather high level of proficiency it seems more relevant to question the inclusion rather than the exclusion criteria. Hence, the question arises whether the language deficit is severe enough to call it SLI.

The criterion for the language discrepancy was the performance at a minimum of 1.1 standard deviations below the age-equivalent score in at least one of the three subtests of the SETK 3-5. At the outset of the project, it was intended to use a cut-off of -1.25 SD as recommended by Tomblin et al. (1996). The beginning of the participant recruitment showed quickly, though, that it was very difficult to find these children. The children had to be impaired severely enough in order to meet this criterion, besides the prerequisites not to present any of the limitations listed in the exclusion criteria and not to be enrolled yet in therapy. Three approaches were possible to deal with this problem: a) abandon the aim that the children were not yet enrolled in intervention, b) continue searching and c) lower the inclusionary criterion.

The inclusion of children who are enrolled in an intervention programme would have resulted in the risk that the therapy may already have improved the children's language – which exactly is the intention of every therapy programme. Although the aim to include only children who did not receive language intervention yet is relatively uncommon in research on SLI, it is not clear why intervention should not influence the outcome. Such an assumption entails that the symptoms must be therapy resistant. In other words, a clinical marker study that includes children irrespective of their intervention status does not search for a key symptom but for a therapy-resistant key symptom. Of course, a marker should ideally mark also resolved cases. This, however, is only feasible if the search does not focus on a linguistic marker because by definition, linguistic symptoms should no longer be evident in resolved cases. Furthermore, the inclusion of children enrolled in intervention seems to contradict the ultimate goal of research on SLI: a better understanding of the impairment in order to offer best intervention to the children. If we assume that the linguistic symptoms are therapy-resistant, though, it seems redundant to follow this path. Thus, in order to abide by the rationale of this project, it is not possible to abandon the criterion regarding the children's intervention status. A modification of the criterion, e.g. controlling for therapy content or method, seemed difficult to realise. A number of different clinics for speech and language therapy were involved in the study with an even larger number of therapists. However, even if it could have been ensured that intervention had not focussed on grammar a generalising transfer across language domains could not have been ruled out. The proposition of a phonological deficit at the roots of SLI (e.g. Chiat, 2000), for example, demonstrates that such an effect across language domains is a reasonable possibility. It becomes therefore clear that any kind of language therapy should be ruled out. In the present study, though, a compromise was made regarding the number of therapy sessions. It was assumed that a maximum number of five therapy sessions could not have had a major impact yet on the children's language since the first session involve case history, diagnostic procedures and the establishment of a good contact between client and therapist. Any child that had received more than five sessions was excluded from the study. In the light of this criterion, the results of the present study seem even more powerful. It can clearly be rejected that the absence of a clinical marker is related to intervention effects.

The possibility to continue searching for children who met the original inclusion and exclusion criteria would have had implications for the clinical applicability. The present project was aimed to contribute to our understanding of a developmental impairment, to an easier identification of affected children and eventually to their therapy. If it becomes more and more difficult, though, to find those individuals who meet the criteria *and* are causing concern to parents and professionals and who would consequently benefit most from intervention the project loses in its significance. It was therefore decided to loosen the inclusion criteria in order to strengthen the relevance of the present project. The criterion for language discrepancy was set at a cut-off of -1.1 rather than -1.25 SD.

Given the findings that no clinical marker could be identified for SLI, this modified inclusionary criterion might seem too generous. It could be argued, that the present sample of children with SLI may represent in fact only a ‘mild’ form of SLI, especially after the exclusion of the 4 most affected children from the morphology probes. Information about the diagnostic accuracy of the SETK 3-5 (Grimm, 2001) and cut-off criteria for language impairment applied by the author could shed some light on this issue. Unfortunately, though, no such information is provided in the handbook. However, the selection by means of the SETK 3-5 was supported by a clinical diagnosis. All children within the SLI group did not only perform at maximum at a percentile of 13.57 on one subtest of the SETK 3-5 and met the exclusion criteria stated in the Methods section, but they were also diagnosed by professionals as SLI. This clinical back-up is especially useful as several authors have raised questions regarding the diagnostic congruence between clinical and research contexts of SLI (Aram et al., 1993; Merrell & Plante, 1997; Plante & Vance, 1994). Moreover, the fact that the children were language impaired was additionally confirmed through the standard measures from the language samples. Even if the four weakest children were excluded from the SLI group, the comparison between this and the ND-A group reached statistical significance for MLU in words as well as MLU in morphemes and the number of different words within the first 100 utterances (see Appendix 9). Thus, the SLI group clearly differed from typically developing children in terms of their language abilities, even if only the ‘milder’ cases were considered. If a clinical marker is meant to be of practical use in clinical work it should be relevant to the target group of clinicians – and not only relevant to a construct created by scientists. The language impairment of the current SLI sample may be less severe than in other studies and this may have prevented the discovery of a clinical marker for SLI. Nevertheless, a clinical marker would be of value especially for these children because of their need to receive language intervention and because their language impairment may be less obvious to non-professionals than a much more severe form. However, although the inclusion criteria meet the intention of the investigator it is important to note that a clinical marker may exist for German SLI of greater severity.

**Age range**

Another aspect of participant recruitment that has influenced the outcome is the age of the children. A view at the individual data shows that especially the older children tend to perform better than the younger children within the SLI group. It is consequently worth asking whether the age range is too large in relation to the sample size. In comparison to other research projects, the age range of the present study seems conform to customary guidelines (e.g. Bedore & Leonard, 1998; Bortolini et al., 2006; Charest & Leonard, 2004). Nevertheless, the exclusion of children receiving intervention pulled the age inevitably down. Stokes et al. (2003) point out that the classifictory power of a clinical marker decreases the younger the children are because it is more difficult to identify SLI in younger than in older children. Yet, it is this reason that justifies the inclusion of children as young as 3 years. A clinical marker would be of special value if it could already be applied to young children.

Three of the 4 children whose language was so severely delayed that they could not yet complete most of the elicitation tasks were 3-year-olds. This may be an argument that the test design was not suitable for this age. However, the fourth child was 4 years old. Two other children of the SLI group, on the other hand, were 3;0 and 3;1 years respectively but could complete the elicitation tasks. It is thus probably more a matter of the language impairment's severity that resulted in the children's inability to respond to the elicitation tasks. However, an age range of 2;1 years seems large within a sample of 17 children, or for many variables eventually only 13 children or even fewer depending on the number of children who produced obligatory contexts. The fact that missing data was to be expected should have been taken into account prior to subject recruitment. The alternative in the present study, though, would have been a smaller sample due to the difficulties in recruiting children with SLI. Nevertheless, the large age range contributed to the large within group variations and restricts the conclusions that can be drawn from the present study.

### *7.2.3 Language matching procedure*

One remarkable outcome of the present study is the lack of a statistically significant difference between the SLI and the ND-L group on almost all variables. If grammar was an area of particular difficulty for the children with SLI they should perform in this area below their general language proficiency. In contrast to children with SLI, typically developing children are considered to show an even language profile across language domains. It was consequently expected that the children with SLI make more grammatical errors than the ND-L children. However, this was not confirmed by the results of the present study. This project differs in terms of language matching from many other studies in the field of SLI. Whereas most researchers select their language-matched participants on the basis of the children's MLU, the children of the present study were selected if their sentence comprehension (SC) score corresponded to the respective SETK score of a child with SLI. This procedure was chosen because the matching criterion was intended to be unrelated to morphology. From this point of view, the findings of the present study are especially remarkable. If morphology is a particular weak area relative to other language areas, MLU-matched control children would be even younger than children matched on language comprehension. Hence, significant differences should be easier to obtain in comparison to children matched on language comprehension than in comparison to MLU-matched children. The fact that these differences were not evident underlines the conclusion that morphology cannot be considered as an area of particular weakness in German-speaking children with SLI.

Two further characteristics of the ND-L group should have provoked a significant difference between the two groups rather than prevented it. First of all, the age of the children: seven of the children with SLI were below an age of 4 years. Table 7 in the Methods section illustrates that the age gap between SLI and ND-L group was on average only 8 months. The age difference between the SLI group and language matched children lies in many studies noticeably above this and reaches often 1.5-2 years (e.g. Bedore & Leonard, 2001; Rice & Wexler, 1996; Stokes et al., 2006). This implies that it should have been more likely for the SLI group in the present study to fall below the performance level of the ND-L group than in other studies with larger age differences. Furthermore, the SC subtest of the SETK was one of the subtests that could determine the inclusion in the SLI group. Only seven of the children with SLI, though, performed 1.1 SD below the age equivalent score and met the definition SLI on grounds of their SC score. The remaining children performed within 1 SD from the age score. This relatively high performance level contributed consequently further to a higher age of the ND-L group than in other studies, although none of the children with SLI scored above their age equivalent score. However, all these factors should rather have facilitated a significant group difference. The absence of such is therefore even more noteworthy.

### *7.2.4 Conclusion on methodological considerations*

In summary, several methodical aspects have been identified which could be improved in a replication study. As the discussion above demonstrated, though, some of these limitations even strengthen the conclusion that no clinical marker exists for German SLI. Nevertheless, a larger quantity of data and much stricter inclusion criteria could possibly reveal group differences, however whether it could be used as a clinical marker remains moot. The children with SLI in the current study represent a sample of the population treated in German clinics for speech and language intervention. At the same time, they evidently did not have a consistent weakness on a particular morphological structure or word order.

## 7.3 The Relation Between Specific Language Impairment and Phonological Impairment

The second research question of the present study concerns the second clinical group: the group of children with an isolated phonological impairment (PI). This group was included in order to investigate whether a potential clinical marker for SLI in German can discriminate also between children with SLI and children with PI.

### *7.3.1 General predictions*

All children with PI are age and gender matched to the children with SLI and consequently also to the age and gender matched control group, ND-A. Given the fact, that the children with PI were selected as experiencing no language difficulties other than their phonological speech output problems, they expectedly should not have marked difficulties in morphology. This prediction is especially based on the fact that the data analysis of the present study took the individual speech output errors of the children into account. All errors with grammatical morphemes that could result from the children's speech problems were excluded prior analysis. Therefore, the following pattern could be expected from group comparisons regarding the production accuracy of grammatical morphemes:

1. On average, the PI group should perform on a similar level as the ND-A group.
2. The PI group should demonstrate a larger production accuracy in grammatical morphemes than the younger control group ND-L.
3. The PI group should use grammatical morphemes more accurately than children with SLI. Especially, if the children with SLI demonstrate a special weakness in particular morphemes (their production accuracy is significantly lower than that of the ND-A and possibly also than that of the ND-L group), the PI group should show a significantly larger production accuracy on these morphemes than the SLI group.

The overview of the results in chapters 5 and 6 show that these predictions are not all fulfilled. Formally, prediction 1 seems confirmed. On most of the variables, statistical group comparisons did not present a statistical difference between the PI and the ND-A group. However, this prediction is not held for the spontaneous production accuracy of dative articles, the case composite derived from the probe data and a composite for present tense inflections derived from the probe data. Secondly, a glance at the raw data and the figures in the studies illustrates that the production accuracy is lower on many morphemes in the PI group than the ND-A group, although these differences failed to reach statistical significance. Thus, the group differences apparent in the raw scores could theoretically be fortuitous. But this seems unlikely given the fact that the children with PI perform consistently lower than the ND-A on many of the morphemes across both measures. A weakness in grammatical morphemes seems rather a stable tendency in the PI group.

The second prediction, claiming that the PI group should produce grammatical morphemes correctly more often than the younger control group ND-L, cannot be supported by the data. Even a look at the basic measures from spontaneous language, MLU and NDW, confirms this trend: the PI group does not have a significant longer MLU than the ND-L group, neither a larger NDW. These observations raise doubts in the participant selection and the definition of the group. Does the PI group indeed represent children with an isolated phonological impairment, i.e. an isolated speech output disorder? Or was the participant selection so imprecise that the children have concomitant language problems? One fact that argues against this is the significant difference between the SLI and the PI group on the basic measures MLU and NDW. The children with PI have a significant longer MLU measured in both morphemes and words as well as they use significantly more different words than the children with SLI. In comparison to the typically developing control children, MLU and NDW values of the PI group fall in between those of the ND-A and ND-L group. The finding, that the values do not reach those of the ND-A group, may be explained by a compensation strategy of the children with PI. The children are most likely less intelligible due to their phonological impairment than typically developing children. It seems therefore a sensible compensation strategy to restrict their output to shorter utterances and words of high frequency in order to enhance the own intelligibility. This again results in lower MLU values and a lower NDW. Transcription errors may be an alternative explanation for the lower MLU of the PI group as function words may have been less intelligible for this group. However, this is unlikely to be the case as all ambiguous cases were transcribed as such but nevertheless as separate morpheme. Thus, ambiguous forms were excluded from the analysis of the respective morphemes but still counted in the total of morphemes. Most important, though, is the finding that statistical analyses did not reveal significant differences between both control groups, ND-A and ND-L. The production accuracy of dative articles and a case composite derived from spontaneous language are the only exceptions from this finding. It is therefore little surprising that the performance of the PI group was not statistically better from that of the ND-L group either.

The third prediction relates to the findings from different morphemes individually. The SLI group performed significantly below both groups of typically developing children on two of the variables: nominative articles and a present tense composite including the copula, both derived from spontaneous production. The PI group, however, performed only on the latter, the present tense composite, significantly better than the SLI group. Furthermore, this was overall the only variable for which significant differences were obtained between both clinical groups. The third prediction is therefore violated to a large extent. This is especially remarkable since the data analyses of the present studies accounted for speech errors. As a consequence, the low performance of the children with PI cannot be explained by their speech output problems. Moreover, the grammatical errors observed within the PI group cannot be assigned to a few outliers only within the group who may have been misidentified as having an age appropriate language development. The present study indicates instead that the children with PI experienced indeed minor problems in verb and case morphology concomitant to their phonological problems.

### *7.3.2 Are the SLI and PI group distinct groups?*

Given the grammatical deficits observed in the PI group, the question needs to be asked whether the SLI and PI group represent indeed distinct populations. The criteria for participant selection allowed an overlap between the SLI and PI group regarding the children's speech output. Phonological deficits were the explicit inclusion criterion for the PI group, but no exclusion criterion for the SLI group. An overlap regarding the children's grammatical skills, though, was unexpected. First of all, the children with PI performed within age expectations on the standardised language test while the children with SLI did not. One of the problems might have been that the inclusion criterion for the SLI group was relaxed from 1.25 SD to 1.1 SD below age norm. The language deficit of the SLI group may consequently not be pronounced enough in order to result in statistical differences between the PI and SLI group on the morphemes tested in the present study. However, as mentioned above, the statistical difference on the basic measures MLU and NDW does indicate a distinct difference between both groups regarding their language skills. Moreover, the occurrence of grammatical errors within the PI group can hardly be explained by a too proficient SLI group. Secondly, the analyses of production patterns indicated that the children with PI produced developmental grammatical errors but at the same time they were often a developmental stage ahead not only of the SLI group but also ahead of the ND-L group. Hence the children with PI show unexpected difficulties in case and verb morphology but they are less pronounced than those of the SLI group.

The analyses of production patterns highlight further the impact of speech errors on grammatical morphemes. This was especially relevant in the present study regarding the production of verb inflections. Within the PI group, about one fifth of the responses resulted in errors that could be accounted for by the children's phonological processes. The SLI group was less affected by phonological errors but the rate in spontaneous production of 3$^{rd}$ person singular reached almost a similar level as for the PI group. All phonological errors were excluded prior to statistical between group comparisons. The high rates of phonological errors underline the importance of this methodological approach. If phonological errors are not recognised as such, phonological errors can be misclassified as grammatical errors. In the present study, qualitative differences between the developmental stages of the PI and SLI group would probably have been more difficult to identify. A comparison of the studies by Rice, Noll and Grimm (1997) on the one hand side and Roberts and Leonard (1997) on the other hand side illustrates the importance of the identification of speech errors, too. Roberts and Leonard reported verb stems as the most frequent production pattern in German SLI. Rice and colleagues, however, excluded phonological errors and found only a small proportion of verb stems in the language of children with SLI. Nevertheless, the distinction between phonological and grammatical errors has often been ignored in data analysis (e.g. Archibald & Joanisse, 2009; Conti-Ramsden, 2003; Conti-Ramsden et al., 2001; Rothweiler et al., 2012). This neglect, however, can have a major impact on research results. A population of children with language and / or speech problems may seem more homogeneous on the basis of such analysis procedures than the present findings suggest. In order to disentangle such an inappropriate overlap between difficulties in two different language domains it is therefore strongly recommended to differentiate between speech and grammatical errors.

Another reason for the careful differentiation between phonological and grammatical errors is the choice of an appropriate intervention method. A single case study by Seef-Gabriel, Chiat and Pring (2012), for example, illustrated that grammatical deficits can be resolved by targeting the relevant phonology for the missing inflections if the child has both language difficulties and co-occurring phonological difficulties. Most importantly, however, this study demonstrates that the language domains phonology and morphology are not independent of each other. The results from the PI group of the present study point to such interplay between both language domains, too. Although the children’s primary difficulties were in the phonological domain they demonstrated additional difficulties in case and verb morphology. As these morphological difficulties were less pronounced it is likely that they may be secondary to the phonological deficits although they did not directly result from the children’s individual speech output errors. For example, the children with PI showed a slowed acquisition of dative marking although none of the children presented relevant speech errors. The error analysis, though, suggested that the children had particular difficulties with the masculine dative articles *dem* and that they overgeneralised instead the accusative article *den*. The contrast of word final *n*/*m* may pose a particular problem to the children with PI especially if this contrast is semantically not very relevant (see discussion chapter 6 for more details). It is therefore likely that the grammatical errors observed in the PI group result from the interplay between phonology and morphology and, accordingly, that they are secondary to the phonological impairment of the children.

There is a vast amount of literature that suggests an effect of phonology on morphology. For example, both the ‘surface account’ by Leonard and colleagues (Leonard & Eyer, 1996; Leonard et al., 1997) and the ‘phonological theory’ by Chiat (2001) suggest that deficits in phonological processing are at the root of the morphological errors observed in children with SLI. Haskill and Tyler (2007) compared the morphology of children with isolated language impairment and children with concomitant language and speech deficits. The children with concomitant speech and language impairment had poorer morphosyntactic skills than the children with isolated language deficits. Phonological processes evident in the children’s speech could not account alone for the group differences. More recently, there is also growing evidence for an influence of phonotactic frequency on the production of grammatical morphemes (Leonard, Davis, & Deevy, 2007; Marshall & van der Lely, 2006; Ott & Höhle, 2013). Most of the research investigating the relation between phonology and morphology, though, take children with morphological difficulties as starting point. Research questions address the impact of phonology on the language of children with SLI. Few researchers take children with PI as starting point and investigate the influence of the children’s deficit on other language domains. One exception is the work by Seeff-Gabriel, Chiat and Dodd (2010). The authors analysed the error patterns in a sentence repetition task across two different groups of children with speech output deficits, children with a consistent phonological disorder and children with an inconsistent phonological disorder, and compared their error patterns with those of typically developing children and those of children with SLI. The results showed that especially the group of children with an inconsistent phonological disorder presented errors that were due to morphosyntactic difficulties but not due to the children's speech difficulties. Hence, previous research supports strongly that there is an interaction between the domains of phonology and morphology. It is consequently likely that the grammatical errors observed in the PI group of the present study are a secondary result of the children's phonological impairment.

As discussed in chapter 5, an impact on the children's morphological performance may not be the only secondary problem of the children with PI. The children with PI produced an unexpected rate of verb infinitives in the 3rd person singular task. This may have resulted from pragmatic difficulties the children experienced during the assessment. Such an interplay between phonology and pragmatics remains still to be investigated. However, the results from the PI group taken on the whole highlight that the interplay of language domains should not be underestimated. Two conclusions are possible from this. First, phonological difficulties are likely to result in grammatical difficulties and PI should therefore be considered as subgroup of SLI. As a consequence, theories of SLI based on processing deficits appear more probable than linguistic theories referring to one language domain only because the problems of the children are likely to generalise across language domains. A second approach, on the other hand, could be that PI and SLI are two distinct problems but that they are sometimes difficult to disentangle due to the interplay between the different language domains. Each of the impairments, though, would probably have its own origin although both may have a longitudinal effect on other language domains. In this case, linguistic approaches to SLI seem more promising than processing accounts. Presumably, though, linguistic accounts may be restricted to particular deficits evident in SLI while further difficulties may develop secondary to the original impact.

In order to investigate both possibilities further, we need research taking children with speech output problems as starting point rather than children with

SLI. PI is an excellent example because this disorder is well researched and stable subgroups of PI could be identified (Dodd, 1995, 2005). Moreover, the subgroups could be identified not only in English-speaking children but for example also in German-speaking children (Fox, 2003), Cantonese (So & Dodd, 1994), Spanish (Goldstein, 1996) and Putonghua (Zhu Hua & Dodd, 2000). Furthermore, the different subgroups can be related to breakdowns on different levels of psycholinguistic models as Dodd (1995, 2005) and also Fox illustrate. Based on these assumptions regarding the children's breakdown, effective intervention programmes have been developed (Dodd, 1995, 2005; Fox & Teutsch, 2005). This clinical evidence, in turn, supports strongly the validity of the suggested classification system. Longitudinal research investigating the impact of the different speech disorders on other language domains, such as morphology and syntax, could provide an important indication whether children with PI should be considered as part of the SLI population or as a distinct group. Different outcomes for the different subgroups would point into the direction of distinct populations whereas similar outcomes may suggest that PI can be considered as SLI. Research taking children with PI as starting point is therefore crucial and highly recommended. If grammatical and phonological errors will be differentiated, such research would offer a good opportunity to investigate the interplay between phonology and other language domains and to increase our understanding of both PI and SLI.

## 7.4 Specific Language Impairment - A Clinical Category or a Theoretical Construct?

The findings of the present study highlight the heterogeneity of the population of children with SLI. Although morphology and syntax had been reported as weakest language domains in German-speaking children with SLI (Grimm, 1993), no clinical marker could be identified in the children's morphology or verb placement that would discriminate successfully between children with SLI and typically developing children. The group of children with SLI participating in the current study presented language profiles that were too heterogeneous to be 'marked' by one or a composite of grammatical deficits. Considering these findings in a wider perspective they question in fact the validity of the diagnostic term *specific language impairment*. Is this indeed a valid clinical category or is it rather a theoretical construct?

### *7.4.1 Critique in the category of SLI*

This question is not a new one. A number of authors cast doubts on the validity of the category SLI earlier. Aram Morris and Hall, for example, called the construct SLI 'an hypothesis in need of testing and validation' (1993, p. 582). The authors found a considerable mismatch between populations selected by means of clinical identification procedures and by means of identification criteria used in research. Aram (1991) stressed further the need to identify valid subgroups of children with SLI in order to obtain more homogeneous groups. Leonard (1991) and Dollaghan (2004b, 2011) both claimed that the label SLI does not apply to a diagnostic category but rather to the lower end of continuous distribution of language skills. Moreover, Norbury and Sparks (2012) illustrated that the diagnosis of SLI is highly dependent on a number of cultural factors such as the socio-economic background of the child, clinical and educational resources available (e.g. assessment material and resources for intervention) and cultural values. The authors pointed out that prevalence rates of developmental disorders vary often drastically across countries and cultural backgrounds. Silveira (2011), on the other hand, criticised that the definition of

SLI varies enormously even within scientific context. Silveira demonstrated that not only the identification procedures of participants change considerably across studies but also that different language tests used are based on very different theoretical approaches. This is supported by findings of Plante and Vance (1994) who compared the discrimination accuracy across a large number of standardised language assessments. The authors failed to find a satisfying diagnostic congruence across the different assessments. According to Silveira, this lack of congruence results from different views on the category SLI. These differences, in turn, are often not acknowledged by the authors but instead results are frequently compared across studies neglecting the different approaches to the category SLI.

### *7.4.2 Heterogeneity of the SLI population across studies*

These examples from the literature show that doubts on the category SLI have been expressed already for more than two decades. In the meanwhile, the field on SLI has moved on. A shift towards some more consistency regarding inclusion criteria for SLI could be observed. A review of research on English-speaking children with SLI published in 2003 and 2004 shows that most of the work in this period had been based on standardised scores for inclusion (Heilmann, 2004 cited in Miller & Fletcher, 2005; Spaulding, Plante, & Farinella, 2006). Previously, and in other languages than English even beyond this time, the recruitment from clinical pools had been very common without insuring that the selected participants would form a homogeneous group by setting clear additional criteria, such as employing norm-referenced tests (e.g. Clahsen, 1989, 1991; Clahsen et al., 1997; Conti-Ramsden et al., 2001; C. Dollaghan & Campbell, 1998; Eisenbeiss et al., 2005; Grimm, 1993). Another milestone for the identification of children with SLI is the study by Tomblin et al. (1996). The authors tested different cutoff criteria in order to obtain the best sensitivity and specificity for the identification of children with SLI and suggested a cutoff of –1,25 SD below age norms. The review by Heilmann indicated that the cutoffs used in research fell usually between -1 and -1.5SD. Apparently, the suggestion by Tomblin et al. had some impact on the decisions made for inclusion criteria although some variation continues to exist between studies.

Nevertheless, the population of children with SLI has often been described as heterogeneous (e.g. Leonard, 1998; Schwartz, 2009). This is the case both across different studies as well as outcome of individual studies. Across studies, the variation in inclusion criteria is probably one of the main reasons. This variation does not only apply to clinical versus assessment based approaches, different cutoff criteria for standardised language tests or different age-ranges, but also to the definition of SLI in general. As illustrated in chapter 2, the discussion whether children with a phonological disorder should be classified as SLI is one example and researchers answered this question differently. Leonard and colleagues, for instance, use consistently language measures referring to different language areas such as vocabulary and language comprehension for participant selection. Furthermore, they also exclude children from their studies who show speech output difficulties that could impact the production of the tested morphemes (e.g.Charest & Leonard, 2004; Dispaldro, Leonard, & Deevy, 2013; Leonard et al., 2007; Leonard et al., 1997; Leonard, Miller, & Gerber, 1999). For Ott and Höhle (2013), on the other hand, an isolated PI is an explicit inclusion criterion. This is certainly also an issue that needs to be considered when standardised tests are used that are based on or include nonword repetition tasks. Although sensitivity and specificity rates regarding SLI may be high for nonword repetition tests, children with PI are likely to fail them, too. Another factor that contributes to the heterogeneity of the population of children with SLI is that different standardised language tests use different 'golden standards' in order to set their differentiation criterion between typical and impaired language development. Letts, Edwards, Schäfer and Sinka (2013) illustrate the difference between a broad or a narrow definition of language impairment and the impact it may have on sensitivity and specificity rates of the test. High rates can be accomplished by setting strict criteria identifying only children with severe language deficits in more than one language domain as impaired. However, authors of other assessment tools, such as the New Reynell Developmental Language Scales (S. Edwards, Letts, & Sinka, 2011), classify children as language impaired if they were clinically identified as having a primary language impairment, irrespective of the language domain. It is obvious that the latter group of language impaired children is likely to be more heterogeneous, although also clinically more relevant than the first group. Another factor that influences specificity and sensitivity rates of an assessment tool is the question whether language impaired children were included in the population tested for standardisation or not. Children with a language deficit are more likely to be identified as language impaired if the norm data is based on typically developing children only (Pena, Spaulding, & Plante, 2006b). The development of a standardised language test requires thus a number of crucial decisions. As a result, groups of children with SLI can differ considerably because authors tend to take these decisions differently.

Given these various factors that contribute to the heterogeneity of the SLI population, consistent findings across studies seem rather remarkable. For example, a large number of studies reported a particular weakness of English-speaking children with SLI in the production of tense and agreement morphemes (see Leonard, 1998 for review). Less consistency seems to exist in languages other than English. The present study is a good example for German, but also for SLI in French conflicting findings were reported whether morphology is particularly vulnerable (Jakubowicz, Nash, Rigaut, & Gérard, 1998; Paradis & Crago, 2001) or not (Thordardottir & Namazi, 2007). For English SLI, though, this language area has widely been accepted as clinical marker (e.g. Conti-Ramsden, 2003; Leonard, 1998; Rice, 2003). As discussed above, consistency regarding the inclusion criteria cannot explain this consistency in findings. It is more probable that linguistic features of English can account for it. In comparison with other languages, English is a language with few grammatical inflections. This provides on the one hand side few possibilities for morphological errors and results consequently in homogeneous error patterns if errors occur. Leonard (2000), on the other hand, proposes an explanation for the particularly high error rate in English morphology as opposed to other languages. At first sight, it seems paradoxical that children with SLI who are acquiring a language with a rich morphology seem to have fewer difficulties to produce grammatical inflections correctly than children with SLI who are acquiring a language with a sparse inflectional morphology such as English. According to Leonard's morphological richness account, however, it is exactly the extent of a language's inflectional morphology that stimulates the children's awareness for grammatical morphemes. The fewer rules there are, the less important they appear to the children. Both factors are hence likely to contribute considerably to the consistency in findings across different studies on English-speaking children with SLI. This is an important reason to carry out cross-linguistic research. While the population of English-speaking children with SLI seems more homogeneous it is research from languages other than English that questions this homogeneity, and thus the clinical category SLI. The present study based on German-speaking children with SLI is consequently a valuable contribution to the discussion of the category SLI.

### *7.4.3 Heterogeneity of the SLI population within studies*

The number of differences across studies listed above may suggest that at least studies considered individually should present homogeneous results. However, even individual studies on children with SLI reported frequently a substantial within group variation (e.g. Leonard et al., 1999; Leonard et al., 2002; C. F. Norbury et al., 2001; Rice & Wexler, 1996). According to Silveira (2011), this is often due to the broad definition of SLI. While some authors restrict the inclusion criteria to children with difficulties in one particular language area (e.g. Grammatical SLI (van der Lely, Rosen, & McClelland, 1998)), most authors include children with impairments in grammar and / or vocabulary irrespective whether receptive and / or expressive language is involved. This approach was suggested amongst others by Tomblin et al. (1996) in order to account for the heterogeneity of the SLI population. Another reason for this broad approach is that the identification of valid subgroups of SLI has a long standing history but has not been particularly successful so far. Although a number of authors could identify subgroups (Aram & Nation, 1975; Conti-Ramsden et al., 1997; Fletcher, 1992; Rapin & Allen, 1987), the findings lacked agreement across the studies. Furthermore, longitudinal data presented by Conti-Ramsden and Botting (1999) indicated that identified subgroups are not stable over time. Children with SLI can move from one to another subgroup over time because the children's individual language profiles change often with continuing language development.

The broad inclusion criteria are very likely to have contributed to the heterogeneity found in the present SLI group, too. In order to be included in the present study, the children with SLI had to perform 1.1 SD below their age norm on at least one of three subtests of the SETK 3-5. As a result, the children could have been included on the basis of very different language difficulties – difficulties in language comprehension, in morphology, in sentence repetition (children 4;0 years and older) or picture description (children aged 3;0-3;11). In other words, the children performed all at the low end of the developmental curve, but in different language areas. It seems thus little surprising that the children did not present congruent language profiles in their morphosyntax. The question arises therefore whether this is indeed the appropriate approach to investigate SLI as long as no valid subgroups of SLI can be identified.

One of the arguments that can back up this approach is that the different language areas are thought to interact with each other. As illustrated above, a vast body of research indicates a relation between phonology and grammar. However, phonological factors and the children's acquisition of vocabulary have been shown to be interrelated, too, (see Chiat, 2001 for overview; McKean, Letts, & Howard, 2013b; Munson, Kurtz et al., 2005) and also the grammatical development is assumed to be dependent on semantic cues (e.g. Moyle, Weismer, Evans, & Lindstrom, 2007; Szagun, 2006) and the growth of the vocabulary (Bates & Goodman, 2001; Marchman & Bates, 1994; Moyle et al., 2007) as well as pragmatic skills (Schaeffer & Matthewson, 2005). The interplay between the different language domains may justify considering children with any primary language deficit as one group because the affected language areas are likely to affect one another. The fact that children with SLI move from one subgroup of SLI to another over time supports this assumption. However, the results of the present study failed to provide further support. The children's individual linguistic profiles varied so that no common characteristics could be identified. Furthermore, considering SLI as one valid category may bear a lot of risks - especially with regards to both conclusions drawn from individual studies as well as conclusions drawn across studies.

### *7.4.4 SLI as clinical category – Impact of this assumption*

The decision hierarchy presented in chapter 2 illustrates many different levels that are involved in the decision making process on SLI. The ultimate goal of research on SLI is twofold: First of all, we want to improve our understanding of this primary language impairment. Why do the children have a language impairment? Questions around this issue go in different directions. One direction concerns the aetiology of SLI. A number of researchers, for example, assumed one common underlying deficit and investigated possible genetic underpinnings of SLI (Gopnik, 1997; O'Brien, Zhang, Nishimura, Tomblin, & Murray, 2003; SLI-Consortium, 2002, 2004). Another direction is the development of theories that can account for the particular linguistic deficits the children show. Several authors see the origin of SLI in linguistics deficits (e.g. Clahsen, 1989; Rice et al., 1997; van der Lely, 2005) whereas other authors suggest processing difficulties of different nature at the root of SLI (e.g. Chiat, 2001; Gathercole & Baddeley, 1990; Leonard et al., 1997; Ullman & Pierpont, 2005). The second ultimate goal is the development of effective intervention programmes. Obviously, a better understanding of the aetiology of SLI would facilitate the development of intervention methods enormously. However, it is almost needless to say that the concept of SLI underlying the respective research projects, (e.g. choice of assessment, inclusion and exclusion criteria, differentiation from other developmental disorders, age range, error analysis, etc.) is likely to have a major impact on the conclusions we can draw from this research.

The review of research history on SLI in chapter 2 illustrates that this process of decision making results in varying conclusions. It seems that the further this process, the less likely it is that researchers will question earlier decisions. Only when returning to earlier levels of the decision hierarchy, in order to test the developed theory, questions arose that resulted eventually in the modification of investigated variables and theories. However, the starting point of the research – the definition of SLI – remained the same for a surprisingly long time. Instead, the incongruence of SLI populations led to two different movements within the field of SLI research. First, a rather narrow approach to SLI and secondly, a broad approach to SLI.

Initially, research on SLI was driven by a modular view of language acquisition. A nativist approach to SLI seemed confirmed by the identification of a marked difficulty in the children's grammar. The development of linguistic theories on SLI evolved from this, assuming a deficit in the linguistic knowledge module at the root of the language impairment. Clinical marker research became temporarily very popular. However, while early identified markers referred to particular linguistic characteristics, such as difficulties in tense and agreement marking, markers tended to become more general. Errors in sentence repetition, for example, difficulties in nonword repetition or a reduced MLU can represent various language difficulties. These findings stimulated a broader view of the phenomenon SLI. The origin of these difficulties was no longer interpreted within a nativist framework but rather in a constructivist view of language acquisition with the acknowledgement of the important interaction between cognitive factors and language skills (s. Tomasello, 2000; Tomasello, 2003). As chapter 2 demonstrates, this was carried further by even more recent approaches to language acquisition. However, the tendency remained to view SLI as a rather broad term.

Both approaches have been based on empirical evidence from the field of SLI. Nevertheless, both approaches struggle with the heterogeneity of the population SLI. For example, the narrow approach to SLI may be clinically less relevant than it is desirable. One way to get over this may be to have large population samples, but these should not been considered as a guarantee that the outcome is indeed representative for the whole population. In contrast, large samples may encourage the reader to lose individual cases out of sight. A clinical marker may be valid for the majority of the population. Those children, however, who are not 'marked' by the clinical marker experience rather disadvantages than advantages from such research. Individual cases may not be identified as needing intervention and, at the same time, explanatory accounts as well as intervention programmes would not necessarily be applicable to these individual cases. Large samples bear therefore the risk of pretending a clinically distinct group while masking individual cases. Another methodological approach to overcome the population's heterogeneity has been the application of stricter inclusion criteria, such as lowering the cut-off point on standardised tests. As discussed above, this approach may result in a larger homogeneity but raises questions, too, about the clinical relevance of these findings. One positive example of narrowing down the approach to SLI, though, has been the work by van der Lely and colleagues (van der Lely, 2003, 2005). Van der Lely identified a sub-group of children with SLI and applied strict inclusion criteria regarding the children's linguistic symptoms. Consequently, the findings of her research refer only to this subgroup which seems far more homogeneous than the usually considered population. However, a modular view of the disorder is evident in this approach, too. The authors tend to include rather old children, beyond pre-school age. From an emergent point of view, though, the children's age may have a large impact on their performance. The focus on subgroups may consequently have even greater potential for the field of SLI if the age limit as well as age range are lowered.

As mentioned, one obstacle in this approach has been the difficulty of identifying stable subgroups of SLI (s. chapter 2 for details). The lack of valid subgroups may therefore suggest that a broad identification procedure is still the best option. This however bears the risk that the pictures will always stay blurred due to the population's heterogeneity. If we acknowledge the individual variability between children of SLI, linguistic markers should rather be considered as 'vulnerable points' within the language development of children with SLI than as clinical identification tool. Furthermore, underlying factors of SLI are difficult to identify if the participant samples are heterogeneous and this hampers the development of effective intervention programmes. It is obvious that this has been more and more acknowledged in the field of SLI. More recent literature does not necessarily refer to SLI any longer but to language impairment (LI) or primary language impairment (PLI) (Law, 2004; McKean et al., 2013a; J. B. Tomblin, Zhang, Buckwalter, & O'Brien, 2003a). This trend is evident in the most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-V)* (American Psychiatric Association, 2013), too. While the previous edition *DSM-IV* (American Psychiatric Association, 2000) included the category SLI it is no longer listed as clinical category in the DSM-V. These changes in terminology may mirror a cultural problem first identified by Bishop (2009): SLI is not well known in public. It is likely that the heterogeneity of the population contributes considerably to this lack of public attention. The phenomenon SLI is difficult to grasp as its appearance can be very varied. However, these acknowledgements of the heterogeneity of SLI within the broader perspective of SLI did not stop researchers from taking the whole population of children with SLI / PLI as starting point for their investigations. One reason besides the incongruent findings regarding subgroups is that researchers hope nevertheless to identify eventually common underlying processes for the various language difficulties and the development of a valid model of SLI. Within this broad view of SLI, the origin of the disorder is no longer considered to lie within the language domain but result from a deficit in cognitive processing that allows the children to acquire language. Given the large number of cognitive processes proposed to be relevant for language acquisition, however, this underlines even more the potential of subgroups. If we could succeed in the identification of stable subgroups of SLI, this could facilitate identification procedures, our understanding of underlying factors, the development of effective intervention programmes as well as raising the public’s recognition of different language problems.

### *7.4.5 Future research on ‘specific language impairment’*

The changes in the view of SLI reflect changes of language acquisition models. Most recent theories of SLI described in chapter 2 are compatible with a neuroconstructivist (Thomas, 2005; Thomas & Karmiloff-Smith, 2005) or on an emergent approach to language acquisition (Evans, 2001). These accounts no longer assume the existence of an intrinsic linguistic system but emphasise the gradual development of language modules. Cognitive processes are triggered by external language input and result in the emergence of new capacities in language skills. Both approaches to language development are characterised by a very dynamic interplay between language input, processing skills and language abilities.

Returning to the original goals of SLI research, the identification of underlying causes of SLI and the development of effective therapy programmes, such a dynamic model of language acquisition presents several challenges to SLI research. First of all, the dynamic interplay of acquisition processes implies longitudinal effects across domain-general as well as domain-specific processes. Profiles of children with SLI are therefore not static but likely to change – regarding both, nonverbal skills as well as language skills . Secondly, genetic as well as environmental factors are thought to play an important role in language acquisition. These, however, are factors that are often difficult to control for in research. Thirdly, neuroconstructive, emergent language models are still relatively new. Even regarding typical language development, gaps remain in our understanding of involved cognitive processes and their interaction with each other. SLI research can provide valuable contribution to a better understanding of typical language development, but at the same time has to deal with these uncertainties. Future research on SLI faces therefore the question how to deal with these challenges in order to gain the best insight for both the understanding of typical as well as impaired language development.

This suggests that the main aim of research should focus on the identification of longitudinal effects. Only the description of developmental trajectories can help to improve our understanding of the interplay between domain-general as well as domain-specific processes and to discriminate between typical and impaired language development. Hence, methodological procedures should therefore involve predominantly longitudinal or, where not possible, cross-sectional research. Additionally, analysis procedures should allow a statistical comparison of changes over time between typical and impaired children and also the investigation of possible interactions between the variables.

Regarding participant selection, research should start as early as possible. With respect to language impairment, this would imply that the population of late talkers seems the best group to start with. It is unlikely that this allows a start at the very beginning of the language impairment if domain-general cognitive processes are at the root of the language impairment. However, to date, the symptoms around an age of 24 months allow the earliest diagnosis of a recognised language delay. Given the fact that a number of the participants would outgrow their language difficulties, such a participant selection would be even broader than the common inclusion and exclusion criteria for SLI. Longitudinal research with late talkers would have nevertheless several advantages. First of all, the longitudinal approach allows a subsequent discrimination between children with a persistent language impairment and so-called late bloomers. Secondly, an early start of observing developmental trajectories can help to identify parallel developmental patterns which might be less dependent on the children's age than on passing a series of developmental stages. This is especially relevant from an emergent perspective on language development as skills and processes are thought to emerge during language development and may change over time again once they have stimulated the following developmental phase. Thirdly, potential differences in developmental stages might allow the discovery of subgroups of language-impaired children. Fourthly, if both language-specific and domain-general processes are investigated, even the inclusion of late bloomers can contribute considerably to improve our understanding of the interplay between language processes and cognitive skills. Fifth, environmental factors are easier to control for the younger the participants are. This latter advantage will obviously lose its benefit with progress of the longitudinal data collection. At least at the outset of the research project, though, the group of participants is likely to be more homogeneous with respects to environmental factors if the children are 2 years old rather than 5 years. Language therapy, for example, will probably not have taken place yet. The focus on very young children seems therefore a very suitable starting point for research. Once developmental trajectories are better known for different populations other stages may be more appropriate.

An alternative to the selection of late talkers could be the focus on clinically relevant subgroups. This would involve that the children's main aspect of difficulties determines the classification or, as suggested by Leonard (2009), the children's response to different types of intervention. The pursuit of the children's developmental changes could indicate whether such clinical subgroups are indeed distinct of each other or whether other factors may account for the linguistic differences. For example, it may be possible that children with SLI pass through a typical order of developmental stages each characterised by particular difficulties and strengths. According to the severity of differences in underlying causes, though, children may pass these stages in their individual pace. To date, developmental trajectories of children with SLI have been described as reflecting those of typically developing children, although delayed (Law, Tomblin, & Zhang, 2008). Two studies investigated the potential of qualitative differences in developmental trajectories regarding different subgroups of SLI (Conti-Ramsden, Clair, Pickles, & Durkin, 2012; Law et al., 2008). Both studies found very little differences regarding the linguistic skills. Subgroups, however, could be identified on the basis of different trajectories in non-verbal skills (Conti-Ramsden et al., 2012). Such research offers a very valuable contribution in the sense of the above mentioned goals. However, both studies followed children from an age of 7 years onwards. Research following a similar outset but focussing on younger children would therefore be an appropriate next step.

The question of a broad versus narrow approach refers also to the choice of research measures. As long as the aim to identify developmental trajectories concerns particular language domains, studies could theoretically focus on these language areas only. However, this would ignore the potential interplay between different domains and processes. According to a neuroconstructive view, the modularity of language is a product of language development but not present yet at its onset. The focus on single language domains seems therefore inappropriate. For that reason, data collection should involve a number of language domains and relevant processing skills. Language samples complemented by tasks targeting particular processing skills seem therefore a good choice. However, the aim to detect developmental changes requires a dense data collection. Tomasello and Stahl (2004) illustrated that the probability of capturing particular target structures is highly dependent on their frequency and that conclusions are almost impossible on low frequent structures. Additional tasks tapping target structures may therefore be necessary in order to keep sampling, transcription and analysis time in a realistic balance to the gain of knowledge. Our increasing knowledge about typical and impaired language acquisition, on the other hand, will probably permit concentration more and more on particular skills. Furthermore, if subgroups of language impairment can successfully be identified and mapped onto language models it would be possible to narrow down the methodological approach accordingly. Thus, the more we know the more precisely measures as well as participants can be selected.

The outcome of such research bears a number of potentials which again would have a positive impact on a number of levels presented in the decision hierarchy on SLI in chapter 2. The identification of different developmental trajectories could result in the description of different subgroups each pointing to different factors causing the children's difficulties. Alternatively, developmental trajectories could indicate that the children follow a certain pathway suggesting a common origin of the language difficulties. Results of regression analyses could provide additional evidence in order to identify the factors that determine the appearance of the individual language profiles. As suggested above, these might point into the direction of the severity of an underlying deficit, to environmental factors or to deficits regarding different cognitive processes or a combination of these. This information would eventually lead to a revision of models of language development which can account for both typical and impaired language acquisition. Intervention programmes that are based on these theoretical accounts would allow the validation of the applied model or if necessary indicate where this needs to be revised. Possibly, categories of language impairment can be divided even further into subgroups, as illustrated with the example of speech output disorders above. Another advantage of valid subgroups is that the awareness of different language deficits would probably increase. Language impairment would be more approachable for clinicians as well as the public because symptoms would be easier to detect and to classify. However, the children with language impairment would benefit most of these advances. The chance that the children's difficulties would be diagnosed and treated effectively would increase enormously and long-term effects of these difficulties could be minimised. The children could benefit in respect of their psychological, social and professional prospects.

## 7.5 Conclusion

The present study supports current trends in SLI research suggesting that SLI is not suitable as clinical category. The findings underline the importance of differential diagnoses between different subtypes of language impairment. However, the approach to SLI at present is too broad to provide a valid classification system. The investigation of developmental trajectories of young children would therefore serve as excellent starting point in order to improve our knowledge of possible subcategories of SLI. Such a classification system would allow a narrow approach to language impairment refining eventually models of language acquisition as well as the support for affected children.

## References


Aguilar-Mediavilla, E. M., Sanz-Torrent, M., & Serra-Raventos, M. (2002). A Comparative Study of the Phonology of Pre-School Children with Specific Language Impairment (SLI), Language Delay (LD) and Normal Acquisition. *Clinical Linguistics & Phonetics, 16*(8), 573-596.

Ahmmed, A., Parker, D., Adams, C., & Newton, V. (2006). Auditory Temporal Resolution in Children with Specific Language Impairment. *Journal of Medical Speech-Language Pathology, 14*(2), 79-96.

Alt, M., Plante, E., & Creusere, M. (2004). Semantic Features in Fast-Mapping: Performance of Preschoolers with Specific Language Impairment Versus Preschoolers with Normal Language. *Journal of Speech, Language, and Hearing Research, 47*(2), 407.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (Revised 4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Angermaier, M. (1974). *Psycholinguistischer Entwicklungstest - PET*. Weinheim: Beltz.

Aram, D. M. (1991). Comments on Specific Language Impairment as a Clinical Construct. *Language, Speech and Hearing Services at Schools, 22*, 84-87.

Aram, D. M., Morris, R., & Hall, N. E. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech & Hearing Research, 36*(3), 580-591.

Aram, D. M., & Nation, J. E. (1975). Patterns of language behavior in children with developmental language disorders. *Journal of Speech & Hearing Research, 18*(2), 229-241.

Archibald, L. M. D., & Gathercole, S. E. (2006a). Nonword Repetition: A Comparison of Tests. *Journal of Speech, Language and Hearing Research, 49*(5), 970-983.

Archibald, L. M. D., & Gathercole, S. E. (2006b). Prevalence of SLI in Language Resource Units. *Journal of Research in Special Educational Needs, 6*(1), 3-10.

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52*(4), 899-914.

Baddeley, A. D. (2000). The Episodic Buffer: A New Component of Working Memory? *Trends in Cognitive Sciences, 11*(44), 417-423.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In E. D. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Bates, E., Dale, P. S., & Thal, D. J. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language*. Oxford, UK: Blackwell.

Bates, E., Elman, J., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1999). Innateness and emergentism. In G. Graham & W. Bechtel (Eds.), *A companion to cognitive science* (pp. 590-601). Oxford: Blackwell.
Bates, E., & Goodman, J. (2001). On the inseparability of grammar and the lexicon: Evidence from acquisition. In M. Tomasello & E. Bates (Eds.), *Language development: The essential readings* (pp. 134-162). Oxford, England: Blackwell.
Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing.* (pp. 3-73). Cambridge, England: Cambridge University Press.
Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech Language and Hearing Research, 41*(5), 1185-1192.
Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech Language and Hearing Research, 44*(4), 905-924.
Bedore, L. M., & Leonard, L. B. (2005). Verb Inflections and Noun Phrase Morphology in the Spontaneous Speech of Spanish-Speaking Children with Specific Language Impairment. *Applied Psycholinguistics, 26*(2), 195-225.
Beery, K. E., & Buktenica, N. A. (1989). *The VMI: developmental test of visual-motor integration* (3rd revision (3R) ed.). Cleveland, Ohio: Modern Curriculum Press.
Beitchman, J. H., Brownlie, E. B., Inglis, A., Wild, J., & et al. (1996). Seven-year follow-up of speech/language impaired and control children: Psychiatric outcome. *Journal of Child Psychology and Psychiatry, 37*(8), 961-970.
Beitchman, J. H., Nair, R., Clegg, M., & Patel, P. G. (1986). Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. *Journal of Speech & Hearing Disorders, 51*(2), 98-110.
Bishop, D. V., Bishop, S. J., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *Journal of speech, language, and hearing research : JSLHR, 42*(1), 155-168.
Bishop, D. V. M. (1994). Grammatical Errors in Specific Language Impairment - Competence or Performance Limitations. *Applied Psycholinguistics, 15*(4), 507-550.
Bishop, D. V. M. (2000). Pragmatic language impairment: A correlate of SLI, a distinct subgroup, or part of the autistic continuum? In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 99-113). New York, NY, US: Psychology Press.
Bishop, D. V. M. (2009). Specific Language Impairment as a Language Learning Disability. *Child Language Teaching and Therapy, 25*(2), 163-165.
Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for

causing language impairment in children. *Journal of Speech Language and Hearing Research, 42*(6), 1295-1310.
Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *journal of Child Psychology and Psychiatry, 37*(4), 391-403.
Bittner, D. (2003). The emergence of verb inflection in two German-speaking children. In D. Bittner, W. U. Dressler & M. Kilani-Schoch (Eds.), *Development of Verb Inflection in First Language Acquisition: A Cross-linguistic Perspective* (21 ed., pp. 53-88). Berlin: Walter de Gruyter.
Black, M., & Chiat, S. (2003). Noun-verb dissociations: A multi-faceted phenomenon. *Journal of Neurolinguistics, 16*(2-3), 231-250.
Bortolini, U., Arfé, B., Caselli, M. C., Degasperi, L., Deevy, P., & Leonard, L. B. (2006). Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders, 41*(6), 695-712.
Bortolini, U., Caselli, M. C., Deevy, P., & Leonard, L. B. (2002). Specific language impairment in Italian: the first steps in the search for a clinical marker. *International Journal of Language & Communication Disorders, 37*(2), 77-93.
Bortolini, U., Caselli, M. C., & Leonard, L. B. (1997). Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech Language and Hearing Research, 40*(4), 809-820.
Bortolini, U., & Leonard, L. B. (2000). Phonology and children with specific language impairment: Status of structural constraints in two languages. *Journal of Communication Disorders, 33*(2), 131-150.
Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry, 46*(3), 317-326.
Botting, N., & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *International Journal of Language & Communication Disorders, 36*(4), 421-432.
Botting, N., & Conti-Ramsden, G. (2003). Autism, primary pragmatic difficulties, and specific language impairment: can we distinguish them using psycholinguistic markers? The omission of subject arguments in children with Specific Language Impairment. *Developmental Medicine & Child Neurology, 45*(8), 515-524.
Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological Processing, Language, and Literacy: A Comparison of Children with Mild-to-moderate Sensorineural Hearing Loss and Those with Specific Language Impairment. *Journal of Child Psychology and Psychiatry, 42*(3), 329-340.
Broomfield, J., & Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Language Teaching & Therapy, 20*(2), 135-151.
Calderon, J. (2004). Working Memory in Spanish-English Bilinguals with Language Impairment. *Dissertation Abstracts International, B: Sciences and Engineering, 64*(7), 3233-B.
Centre for Evidence-Based Medicine. (2000, 2013). CEBM Statistics Calculator. Retrieved 23 January 2013, from http://ktclearinghouse.ca/cebm/practise/ca/calculators/statscalc
Charest, M. J., & Leonard, L. B. (2004). Predicting tense: finite verb morphology and subject pronouns in the speech of typically-developing children and

children with specific language impairment. *Journal of Child Language, 31*(1), 231-246.
Chiat, S. (2000). *Understanding Children with Language Problems*. Cambridge, UK: Cambridge University Press.
Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes. Special Issue: Language and Cognitive Processes in Developmental Disorders, 16*(2-3), 113-142.
Chomsky, N. (1959). A Review of B. F. Skinner's Verbal Behavior. *Language, 35*, 26-58.
Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
Clahsen, H. (1986). *Die Profilanalyse: Ein linguistisches Verfahren für die Sprachdiagnose im Vorschulalter*. Berlin: Marhold.
Clahsen, H. (1989). The Grammatical Characterization of Developmental Dysphasia. *Linguistics, 5*(303), 897-920.
Clahsen, H. (1991). *Child language and developmental dysphasia*. Amsterdam: John Bejamins.
Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics, 10*(2-3), 151-171.
Clahsen, H., & Rothweiler, M. (1992). Inflectional rules in children's grammars: Evidence from German participles. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1992* (pp. 1-34). Dordrecht, The Netherlands: Kluwer.
Coady, J. A., & Evans, J. L. (2008). Uses and Interpretations of Non-Word Repetition Tasks in Children with and without Specific Language Impairments (SLI). *International Journal of Language & Communication Disorders, 43*(1), 1-40.
Cole, K. N., Schwartz, I. S., Notari, A. R., Dale, P. S., & Mills, P. E. (1995). Examination of the Stability of 2 Methods of Defining Specific Language Impairment. *Applied Psycholinguistics, 16*(1), 103-123.
Connell, P. J., & Myles-Zitzer, C. (1982). An analysis of elicited imitations as a language evaluation procedure. *Journal of Speech & Hearing Disorders, 47*(4), 390-396.
Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 46*(5), 1029-1037.
Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42*(5), 1195-1204.
Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(6), 741-748.
Conti-Ramsden, G., Clair, M. C. S., Pickles, A., & Durkin, K. (2012). Developmental Trajectories of Verbal and Nonverbal Skills in Individuals With a History of Specific Language Impairment: From Childhood to Adolescence. *Journal of Speech, Language, and Hearing Research, 55*(1716-1735).

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research, 40*(4), 765-777.
Conti-Ramsden, G., & Durkin, K. (2007). Phonological short-term memory, language and literacy: Developmental relationships in early adolescence in young people with SLI. *Journal of Child Psychology and Psychiatry, 48*(2), 147-156.
Conti-Ramsden, G., & Hesketh, A. (2003). Risk markers for SLI: A study of young language-learning children. *International Journal of Language & Communication Disorders, 38*(3), 251-263.
Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: A new look at an old hypothesis. *Journal of Speech, Language, and Hearing Research, 50*, 647-666.
Crago, M., & Paradis, J. (2003). Two of a Kind? The Importance of Commonalities and Variation Across Languages and Learners. In Y. Levy & J. C. Schaeffer (Eds.), *Language competence across populations : toward a definition of specific language impairment* (pp. 97-110). Mahwah, N.J.: L. Erlbaum Associates.
Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*(4), 482-486.
De Jong, J. (1999). *Specific language impairment in Dutch : inflectional morphology and argument structure*. Groningen, Netherlands: Groningen Dissertations in Linguistics.
De Jong, J. (2004). Grammatical impairment: An overview and a sketch of Dutch. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications.* (pp. 261-281). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
de Villiers, J. G. (2003). Defining SLI: A linguistic perspective. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment.* (pp. 425-447). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-Word and Nonword Repetition in ItalianSpeaking Children With Specific Language Impairment: A Study of Diagnostic Accuracy. *Journal of Speech, Language, and Hearing Research, 56*, 323–336
Dodd, B. (1995). *The differential diagnosis and treatment of children with speech disorder*. San Diego, Calif.: Singular Pub. Group.
Dodd, B. (Ed.). (2005). *Differential diagnosis and treatment of children with speech disorder* (2nd ed.). London: Whurr.
Dollaghan, C., & Campbell, T. F. (1998). Nonword Repetition and Child Language Impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136-1146.
Dollaghan, C. A. (2004a). Evidence-Based Practice in Communication Disorders: What Do We Know, and When Do We Know It? *Journal of Communication Disorders, 37*(5), 391-400.
Dollaghan, C. A. (2004b). Taxometric Analyses of Specific Language Impairment in 3- and 4-Year-Old Children. *Journal of Speech, Language, and Hearing Research, 47*(2), 464-475.

Dollaghan, C. A. (2011). Taxometric analyses of specific language impairment in 6-year-old children. *Journal of speech, Language, and Hearing Research, 54*, 1361-1371.

Dromi, E., Leonard, L. B., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech Language and Hearing Research, 42*(6), 1414-1431.

Dromi, E., Leonard, L. B., & Blass, A. (2003). Different methodologies yield incongruous results: A study of the spontaneous use of verb forms in Hebrew. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 273-289). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (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 and Hearing Research, 39*, 643-654.

Durkin, K., & Conti-Ramsden, G. (2010). Young People with Specific Language Impairment: A Review of Social and Emotional Functioning in Adolescence. *Child Language Teaching and Therapy, 26*(2), 105-121.

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19*(2), 279-309.

Edwards, S., Letts, C., & Sinka, I. (2011). *The New Reynell Developmental Language Scales*. London: GL Assessment.

Eisenbeiss, S., Bartke, S., & Clahsen, H. (2005). Structural and Lexical Case in Child German: Evidence from Language-Impaired and Typically Developing Children. *Essex research reports in linguistics, 48*, 1-42.

Eisenbeiss, S., Bartke, S., & Clahsen, H. (2006). Structural and Lexical Case in Child German: Evidence From Language-Impaired and Typically Developing Children. *Language Acquisition, 13*(1), 3-32.

Ellis Weismer, S. (1996). Capacity limitations in working memory: The impact on lexical and morphological learning by children with language impairment. *Topics in Language Disorders, 17*(1), 33-44.

Ellis Weismer, S., Tomblin, J. B., Zhang, X. Y., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech Language and Hearing Research, 43*(4), 865-878.

Evans, J. L. (1996). SLI Subgroups: Interaction Between Discourse Constraints and Morphosyntactic Deficits. *Journal of Speech and Hearing Research, 39*, 655-660.

Evans, J. L. (2001). An emergent account of language impairments in children with SLI: Implications for assessment and intervention. *Journal of Communication Disorders. Special Issue: Clinics issue: Peering into the future, 34*(1-2), 39-54.

Fee, E. J. (1995). The Phonological System of a Specifically Language-Impaired Population. *Clinical Linguistics & Phonetics, 9*(3), 189-209.

Fey, M. E., Long, S. H., & Cleave, P. L. (1994). Reconsideration of IQ criteria in the definition of specific language impairment. In R. V. Watkins & M. L. Rice (Eds.), *Specific language impairments in children* (Vol. 4, pp. 161-178). Baltimore, MD, US: Paul H Brookes Publishing.

Fletcher, P. (1992). Subgroups in school-age language-impaired children. In P. Fletcher & D. Hall (Eds.), *Specific speech and language disorders in children : correlates, characteristics, and outcomes*. London: Whurr Publishers.
Fletcher, P., Leonard, L. B., Stokes, S. F., & Wong, A. M.-Y. (2009). Morphosyntactic Deficits in Cantonese-speaking Children with Specific Language Impairment. In S.-P. Law, B. S. Weekes & A. M.-Y. Wong (Eds.), *Language Disorders in Speakers of Chinese*. Bristol: Cromwell Press Ltd.
Fletcher, P., & Peters, J. (1984). Characterising language impairment in children: an exploratory study. *Language Testing, 1*, 33-49.
Fox, A. V. (2002). *PLAKSS - Psycholinguistische Analyse kindlicher Sprechstörungen*. Frankfurt: Swets-Test Services.
Fox, A. V. (2003). *Kindliche Aussprachestörungen: phonologischer Erwerb, Differenzialdiagnostik, Therapie*. Idstein: Schulz-Kirchner Verlag.
Fox, A. V. (2006). *TROG-D: Test zur Überprüfung des Grammatikverständnisses* (1 ed.). Idstein: Schulz-Kirchner Verlag.
Fox, A. V., & Teutsch, A. (2005). Therapie bei Kindern mit Aussprachestörungen im deutschsprachigen Raum: was wirkt wann, wie und wann nicht? *Die Sprachheilarbeit, 50*(6), 273-279.
Fujiki, M., & Brinton, B. (1987). Elicited imitation revisited: A comparison with spontaneous language production. *Language, Speech, and Hearing Services in Schools, 18*(4), 301-311.
Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49*(1), 87-115.
Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*(4), 513-543.
Gathercole, S. E., & Baddeley, A. (1996). *Children's Test of Nonword Repetition*. London: Harcourt Assessment.
Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the Role of Phonological STM in the Development of Vocabulary in Children: A Longitudinal Study. *Journal of Memory and Language, 28*(2), 200-213.
Gathercole, S. E., & Baddeley, A. D. (1990). Phonological Memory Deficits in Language Disordered Children: Is There a Causal Connection? *Journal of Memory and Language, 29*(3), 336-360.
Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hillsdale, NJ: Lawrence Erlbaum.
Gavin, W. J., Klee, T., & Membrino, I. (1993). Differentiating Specific Language Impairment from Normal Language-Development Using Grammatical Analysis. *Clinical Linguistics & Phonetics, 7*(3), 191-206.
Gerken, L. (1994). A metrical template account of children’s weak syllable omissions from multisyllabic words. *Journal of Child Language, 21*, 567-584.
Gillam, R. B., & Hoffman, L. M. (2004). Information processing in children with specific language impairment. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications.* (pp. 137-157). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
Girbau, D., & Schwartz, R. G. (2007). Non-Word Repetition in Spanish-Speaking Children with Specific Language Impairment (SLI).

*International Journal of Language & Communication Disorders, 42*(1), 59-75.
Gladfelter, A., & Leonard, L. B. (2013). Alternative Tense and Agreement Morpheme Measures for Assessing Grammatical Deficits During the Preschool Period. *Journal of Speech Language and Hearing Research, 56*(2), 542-552.
Goldstein, B. (1996). Error groups in Spanish-speaking children with phonological disorders. In T. W. Powell (Ed.), *Pathologies of Speech and L anguage: Contributions of Clinical Phonetics and Linguistics* (pp. 171-177). New Orleans: ICPLA.
Goorhuis-Brouwer, S. M., & Knijff, W. A. (2001). Efficacy of speech therapy in children with language disorders: Specific language impairment compared with language impairment in co-morbidity with cognitive delay. *International Journal of Pediatric Otorhinolaryngology, 63*, 129-136.
Gopnik, M. (1994). Impairments of tense in a familial language disorder. *Journal of Neurolinguistics, 8*(2), 109-133.
Gopnik, M. (1997). Language deficits and genetic factors. *Trends in Cognitive Sciences, 1*(1), 5-9.
Gopnik, M., & Crago, M. B. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39*(1), 1-50.
Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in nonword repetition performance of children with and without specific language impairment: a meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*(1), 177-195.
Gray, S. (2003a). Diagnostic Accuracy and Test-Retest Reliability of Nonword Repetition and Digit Span Tasks Administered to Preschool Children with Specific Language Impairment. *Journal of Communication Disorders, 36*(2), 129-151.
Gray, S. (2003b). Word-Learning by Preschoolers with Specific Language Impairment: What Predicts Success? *Journal of Speech, Language, and Hearing Research, 46*(1), 56-67.
Gray, S. (2004). Word Learning by Preschoolers with Specific Language Impairment: Predictors and Poor Learners. *Journal of Speech, Language, and Hearing Research, 47*(5), 1117-1132.
Gray, S. (2005). Word Learning by Preschoolers with Specific Language Impairment: Effect of Phonological or Semantic Cues. *Journal of Speech, Language, and Hearing Research, 48*(6), 1452-1467.
Gray, S. (2006). The Relationship between Phonological Memory, Receptive Vocabulary, and Fast Mapping in Young Children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 49*(5), 955-969.
Grela, B. G. (2000). The Influence of Argument Structure and Sentence Length on Morphosyntax in Children with Specific Language Impairment. *Dissertation Abstracts International, B: Sciences and Engineering, 60*(11), 5473-B.
Grela, B. G., & Leonard, L. B. (2000). The influence of argument-structure complexity on the use of auxiliary verbs by children with SLI. *Journal of Speech Language and Hearing Research, 43*(5), 1115-1125.
Grimm, H. (1993). Syntax and morphological difficulties in German-speaking children with specific language impairment: Implications for diagnosis

and intervention. In H. Grimm & H. Skowronek (Eds.), *Language acquisition problems and reading disorders: Aspects of diagnosis and intervention* (pp. 25-63). Berlin: Walter de Gruyter.
Grimm, H. (2001). *SETK 3-5: Sprachentwicklungstest für drei- bis fünfjährige Kinder*. Göttingen: Hogrefe.
Grimm, H., & Doil, H. (2000). *ELFRA-2: Elternfragebogen für zweijährige Kinder*. Göttingen: Hogrefe.
Grimm, H., & Schöler, H. (1991). *Heidelberger Sprachentwicklungstest - HSET* (2. verbesserte Auflage ed.). Göttingen: Hogrefe.
Grimm, H., & Weinert, S. (1990). Is the Syntax Development of Dyspahsic Children Deviant and Why? New Findings to an Old Question. *Journal of Speech and Hearing Research, 33*, 220-228.
Grinstead, J., Baron, A., Vega-Mendoza, M., Mora, J. D. I., Cantú-Sánchez, M., & Flores, B. (2013). Tense Marking and Spontaneous Speech Measures in Spanish Specific Language Impairment: A Discriminant Function Analysis. *Journal of Speech, Language, and Hearing Research, 56*, 352-363
Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA, US: The MIT Press.
Håkansson, G. (2001). Tense Morphology and Verb-Second in Swedish L1 Children, L2 Children and Children with SLI. *Bilingualism: Language and Cognition, 4*(1), 85-99.
Hamann, C., Penner, Z., & Lindner, K. (1998). German Impaired Grammar: The Clause Structure Revisited. *Language Acquisition, 7*(2), 193-245.
Hansson, K. (1997). Patterns of verb usage in Swedish children with SLI: An application of recent theories. *First Language, 2*(50), 195-217.
Hansson, K., Forsberg, J., Löfqvist, A., Mäki-Torkko, E., & Sahlén, B. (2004). Working memory and novel word learning in children with hearing impairment and children with specific language impairment. *International Journal of Language & Communication Disorders, 39*(3), 401-422.
Hansson, K., Nettelbladt, U., & Leonard, L. B. (2000). Specific language impairment in Swedish: The status of verb morphology and word order. *Journal of Speech, Language and Hearing Research, 43*(4), 848-864.
Haskill, A. M., & Tyler, A. A. (2007). A comparison of linguistic profiles in subgroups of children with specific language impairment. *American Journal of Speech-Language Pathology, 16*(3), 209-221.
Häuser, D., Kasielke, E., & Scheidereiter, U. (1994). *Kindersprachtest für das Vorschulalter - KISTE*. Weinheim: Beltz.
Hsu, H. J., & Bishop, D. V. M. (2011). Grammatical Difficulties in Children with Specific Language Impairment: Is Learning Deficient? *Human Development, 53*(5), 264–277.
Ingram, D. (2008). Cross-Linguistic Phonological Acquisition. In M. Ball, J., M. Perkins, R., N. Müller & S. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 626-640). Oxford, UK: Blackwell Publishing.
Jakubowicz, C., Nash, L., Rigaut, C., & Gérard, C.-L. (1998). Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition: A Journal of Developmental Linguistics, 7*(2-4), 113-160.
Jarrold, C., Baddeley, A. D., & Hewes, A. K. (2000). Verbal short-term memory deficits in Down Syndrome: a consequence of problems in rehearsal?

*Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(2), 233-244.

Johnston, J. R., & Weismer, S. E. (1983). Mental Rotation Abilities in Language-Disordered Children. *Journal of Speech and Hearing Research, 26*(3), 397-403.

Kail, R. (1994). A Method for Studying the Generalized Slowing Hypothesis in Children with Specific Language Impairment. *Journal of Speech and Hearing Research, 37*(2), 418-421.

Kany, W., & Schöler, H. (1998). "Ich hab bestimmt schon Fehler gemacht" - Die sprachlichen Leistungen der SSES-Kinder. In H. Schöler, W. Fromm & W. Kany (Eds.), *Spezifische Sprachentwicklungsstörung und Sprachlernen: Erscheinungsformen, Verlauf, Folgerungen für Diagnostik und Therapie* (pp. 111-176). Heidelberg: Edition Schindele.

Kauschke, C. (2012). *Kindlicher Sprachwerwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. Göttingen: de Gruyter.

Kauschke, C., & Siegmüller, J. (2002). *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen* (1 ed.). München: Elsevier.

Kauschke, C., & Siegmüller, J. (2009). *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen* (2 ed.). München: Urban & Fischer Verlag / Elsevier GmbH.

Keller-Bell, Y. D. (2001). Linguistic Processing in Children with Down Syndrome. *Dissertation Abstracts International, B: Sciences and Engineering, 61*(11), 5836-B.

Klampfer, S. (2003). Emergence of verb paradigms in one Austrian child. In D. Bittner, W. U. Dressler & M. Kilani-Schoch (Eds.), *Development of Verb Inflection in First Language Acquisition: A Cross-linguistic Perspective* (pp. 297-322). Berlin: Walter de Gruyter.

Klee, T., Gavin, W. J., & Stokes, S. F. (2007). Utterance length and lexical diversity in American- and British-English speaking children: What is the evidence for a clinical marker of SLI? In R. Paul (Ed.), *Language disorders from a developmental perspective: Essays in honor of Robin S. Chapman. New directions in communication disorders research: Integrative approaches.* (pp. 103-140). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Klee, T., Stokes, S. F., Wong, A. M. Y., Fletcher, P., & Gavin, W. J. (2004). Utterance Length and Lexical Diversity in Cantonese-Speaking Children With and Without Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 47*(6), 1396-1410.

Köhler, K., & Bruyère, S. (1996). Finiteness and verb placement in the L1 acquisition of German. *Wiener Linguistische Gazette*, 53-54.

Korkman, M., & Häkkinen-Rihu, P. (1994). A new classification of developmental language disorders (DLD). *Brain and Language, 47*(1), 96-116.

Krassowski, E., & Plante, E. (1997). IQ variability in children with SLI: Implications for use of cognitive referencing in determining SLI. *Journal of Communication Disorders, 30*(1), 1-9.

Langen-Müller, U. d., Kauschke, C., Kiesel-Himmel, C., Neumann, K., & Noterdaeme, M. (2011, 01/06/2013). Diagnostik von Sprachentwicklungsstörungen (SES), unter Berücksichtigung umschriebener Sprachentwicklungsstörungen (USES): Interdisziplinäre

S2k-Leitlinie. AWMF-Registernummer 049/006. Retrieved June 14th, 2013, 2013, from http://www.awmf.org/uploads/tx_szleitlinien/049-006l_S2k_Sprachentwicklungsstoerungen_Diagnostik_2013-06_01.pdf

Law, J. (2004). The close association between classification and intervention for children with primary language impairments. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications.* (pp. 401-418). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*(2), 165-188.

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*(3), 739-749.

Leonard, L. B. (1989). Language Learnability and Specific Language Impairment in Children. *Applied Psycholinguistics, 10*(2), 179-202.

Leonard, L. B. (1991). Specific language impairment as a clinical category. *language, Speech, and Hearing Services in Schools, 22*, 66-68.

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

Leonard, L. B. (2000). Specific language impairment across languages. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome.* (pp. 115-129). New York, NY, US: Psychology Press.

Leonard, L. B. (2003). Specific language impairment: Characterizing the deficits. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 209-231). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Leonard, L. B. (2009). Some Reflections on the Study of Children with Specific Language Impairment. *Child Language Teaching and Therapy, 25*(2), 169-171.

Leonard, L. B., Bortolini, U., Caselli, M. C., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition, 2*, 151-179.

Leonard, L. B., Davis, J., & Deevy, P. (2007). Phonotactic probability and past tense use by children with specific language impairment and their typically developing peers. *Clinical Linguistics & Phonetics, 21*(10), 747-758.

Leonard, L. B., & Deevy, P. (2006). Cognitive and linguistic issues in the study of children with specific language impairment. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 1143-1171). London: Academic Press.

Leonard, L. B., Dromi, E., Adam, G., & Zadunaisky-Ehrlich, S. (2000). Tense and finiteness in the speech of children with specific language impairment acquiring Hebrew. *International Journal of Language and Communication Disorders, 35*(3), 319-335.

Leonard, L. B., & Eyer, J. A. (1996). Surface properties of grammatical morphology and morphological deficits in children with specific language impairment. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp. 233-247). Hillsdale, NJ: Lawrence Erlbaum.

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research, 40*(4), 741-753.

Leonard, L. B., McGregor, K. K., & Allen, G. D. (1992). Grammatical Morphology and Speech-Perception in Children with Specific Language Impairment. *Journal of Speech and Hearing Research, 35*(5), 1076-1085.

Leonard, L. B., Miller, C., & Gerber, E. (1999). Grammatical morphology and the lexicon in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 42*(3), 678-689.

Leonard, L. B., Miller, C. A., Deevy, P., Rauf, L., Gerber, E., & Charest, M. (2002). Production operations and the use of nonfinite verbs by children with specific language impairment. *Journal of Speech, Language and Hearing Research, 45*(4), 744-758.

Leonard, L. B., Miller, C. A., Grela, B. G., Holland, A. L., Gerber, E., & Petucci, M. (2000). Production operations contribute to the grammatical morpheme limitations of children with specific language impairment. *Journal of Memory and Language, 43*(2), 362-378.

Leonard, L. B., Sabbadini, L., Leonard, J. S., & Volterra, V. (1987). Specific Language Impairment in Children - a Cross-Linguistic Study. *Brain and Language, 32*(2), 233-252.

Letts, C., Edwards, S., Schaefer, B., & Sinka, I. (2013). The New Reynell Developmental Language Scales: Descriptive account and illustrative case study. *Child Language Teaching & Therapy, onlinefirst version* 1-14.

Levelt, W. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Stein, C. M., Shriberg, L. D., Iyengar, S. K., et al. (2006). Dimensions of early speech sound disorders: A factor analytic study. *Journal of Communication Disorders, 39*(2), 139-157.

Liiva, C. A., & Cleave, P. L. (2005). Roles of Initiation and Responsiveness in Access and Participation for Children With Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 48*(4), 868-883.

Lindner, K., & Johnston, J. R. (1992). Grammatical morphology in language–impaired children acquiring English or German as their first language: A functional perspective. *Applied Psycholinguistics, 13*, 115-129.

Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language, 31*(6), 713-733.

Maillart, C., & Parisse, C. (2006). Phonological deficits in French speaking children with SLI. *International Journal of Language and Communication Disorders, 41*(3), 253-274.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*(2), 339-366.

Marshall, C. R., & van der Lely, H. K. J. K. J. (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition, 100*(2), 302-320.

Marshall, C. R., & van der Lely, H. K. J. K. J. (2007). The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Advances in Speech Language Pathology, 9*(3), 191-203.

Marton, K., & Schwartz, R. G. (2003). Working Memory Capacity and Language Processes in Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 46*(5), 1138-1153.

McArthur, G. M., Ellis, D., Atkinson, C. M., & Coltheart, M. (2008). Auditory Processing Deficits in Children with Reading and Language Impairments: Can They (and Should They) Be Treated? *Cognition, 107*(3), 946-977.

McKean, C., Letts, C., & Howard, D. (2013a). Developmental Change is Key to Understanding Primary Language Impairment: the Case of Phonotactic Probability and Nonword Repetition. *Journal of Speech, Language and Hearing Research, 56*(5), 1579-1594.

McKean, C., Letts, C., & Howard, D. (2013b). Functional reorganization in the developing lexicon: Separable and changing influences of lexical and phonological variables on children's fast-mapping. *Journal of Child Language, 40*(2), 307-335.

Merrell, A. W., & Plante, E. (1997). Norm-referenced test interpretation in the diagnostic process. *Language, Speech, and Hearing Services in Schools, 28*(1), 50-58.

Miller, J. F., & Fletcher, P. (2005). Developmental theory and language disorders: background issues. In P. Fletcher & J. F. Miller (Eds.), *Developmental theory and language disorders* (pp. 1-10). Amsterdam: John Benjamins Publishing Company.

Moyle, M. J., Weismer, S. E., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal Relationships Between Lexical and Grammatical Development in Typical and Late-Talking Children. *Journal of Speech, Language, and Hearing Research, 50*(2), 508-528.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships Between Nonword Repetition Accuracy and Other Measures of Linguistic Development in Children With Phonological Disorders. *Journal of Speech, Language, and Hearing Research, 48*(1), 61-78.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The Influence of Vocabulary Size, Phonotactic Probability, and Wordlikeness on Nonword Repetitions of Children With and Without Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 48*(5), 1033-1047.

Nakamura, M., Plante, E., & Swisher, L. (1990). Predictors of Novel Inflection Learning by Preschool-Aged Boys. *Journal of Speech and Hearing Research, 33*(4), 747-754.

Norbury, C. F., Bishop, D. V. M., & Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *Journal of Speech Language and Hearing Research, 44*(1), 165-178.

Norbury, C. F., & Sparks, A. (2012). Difference or Disorder? Cultural Issues in Understanding Neurodevelopmental Disorders. *Developmental Psychology, 49*(1), 45-58.

O'Brien, E. K., Zhang, X., Nishimura, C., Tomblin, J. B., & Murray, J. C. (2003). Association of specific language impairment (SLI) to the region of 7q31. *American journal of human genetics, 72*(6), 1536-1543.

Oetting, J. B., & Cleveland, L. H. (2006). The clinical utility of nonword repetition for children living in the rural south of the US. *Clinical Linguistics & Phonetics, 20*, 7-8.

Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech and Hearing Research, 38*(2), 434-435.

Orgassa, A., & Weerman, F. (2008). Dutch gender in specific language impairment and second language acquisition. *Second Language Research, 24*(3), 333-364.

Ott, S., & Höhle, B. (2013). Verb inflection in German-learning children with typical and atypical language acquisition: the impact of subsyllabic frequencies. *Journal of Child Language, 40*(1), 169-192.

Owen, A. J., & Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children: With specific language impairment: Application of D. *Journal of Speech Language and Hearing Research, 45*(5), 927-937.

Paradis, J., & Crago, M. (2001). The Morphosyntax of Specific Language Impairment in French: An Extended Optional Default Account. *Language Acquisition, 9*(4), 269-300.

Pena, E. D., Spaulding, T. J., & Plante, E. (2006a). The composition of normative groups and diagnostic decision making: shooting ourselves in the foot. *American Journal of Speech-Language Pathology, 15*(3), 247-254.

Pena, E. D., Spaulding, T. J., & Plante, E. (2006b). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech Language Pathology, 15*(3), 247-254.

Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *The Annual Review of Psychology, 60*, 283-306.

Pihko, E., Mickos, A., Kujala, T., Pihlgren, A., Westman, M., Alku, P., et al. (2007). Group intervention changes brain activity in bilingual language-impaired children. *Cerebral Cortex, 17*(4), 849-858.

Pizzioli, F., & Schelstraete, M.-A. (2008). The argument-structure complexity effect in children with specific language impairment: Evidence from the use of grammatical morphemes in French. *Journal of Speech, Language, and Hearing Research, 51*(3), 706-721.

Plante, E., & Vance, R. (1994). Selection of preschool language tests: a data-based approach *Language, Speech, and Hearing Services in Schools, 25*, 15-24.

Polite, E. J., & Leonard, L. B. (2011). The use of definite and indefinite articles by children with specific language impairment. *International Journal of Speech-Language Pathology, 13*(4), 291-300.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language, 38*(3), 265-282.

Rapin, I., & Allen, D. A. (1987). Developmental dysphasia and autism in preschool children: characteristics and subtypes. In *Proceedings of the first International Symposium on Specific Speech and Language Disorders in Children* (pp. 20-35). Bentford, Eng.: Association for All Speech-Impaired Children.

Redmond, S. M. (2005). Differentiating SLI from ADHD Using Children's Sentence Recall and Production of Past Tense Morphology. *Clinical Linguistics and Phonetics, 19*(2), 109-127.

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, N.J.: L. Erlbaum Associates.

Rice, M. L., & Bode, J. V. (1993). GAPS in the Verb Lexicons of Children with Specific Language Impairment. *First Language, 1*(37), 113-131.

Rice, M. L., Hoffman, L., & Wexler, K. (2009). Judgments of Omitted BE and DO in Questions as Extended Finiteness Clinical Markers of Specific Language Impairment (SLI) to 15 Years: A Study of Growth and Asymptote. *Journal of Speech, Language, and Hearing Research, 52*, 1417-1433.

Rice, M. L., Noll, K. R., & Grimm, H. (1997). An extended optional infinitive stage in German-speaking children with specific language impairment. *Language Acquisition: A Journal of Developmental Linguistics, 6*(4), 255-295.

Rice, M. L., Oetting, J. B., Marquis, J., & Bode, J. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research, 37*(1), 106-122.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39*(6), 1239-1257.

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*(4), 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*(6), 1412-1431.

Roberts, S. S., & Leonard, L. B. (1997). Grammatical deficits in German and English: A crosslinguistic study of children with specific language impairment. *First Language, 17*(50, Pt 2), 131-150.

Rodekohr, R. K., & Haynes, W. O. (2001). Differentiating Dialect from Disorder: A Comparison of Two Processing Tasks and a Standardized Language Test. *Journal of Communication Disorders, 34*(3), 255-272.

Roodenrys, S., & Stokes, J. (2001). Serial recall and nonword repetition in reading disabled children. *Reading and Writing, 14*(5-6), 379-394.

Rothweiler, M., Chilla, S., & Clahsen, H. (2012). Subject-verb agreement in Specific Language Impairment: A study of monolingual and bilingual German-speaking children. *Bilingualism: Language and Cognition, 15*(1), 39-57.

Saccuman, M. C. (2006). Language comprehension and production in normally developing children and children with language impairment: An fMRI study. *Dissertation Abstracts International, B: Sciences and Engineering, 66*(10), 5713.

Sahlén, B., Reuterskiold-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Non-word repetition in children with language impairment - pitfalls and possibilities. *International Journal of Language & Communication Disorders, 34*(3), 337-352.

Schaeffer, J., & Matthewson, L. (2005). Grammar and Pragmatics in the Acquisition of Article Systems. *Natural Language & Linguistic Theory, 23*, 53-101.

Schafer, R. J., & De Villiers, J. (2000). Imagining articles: What a and the can tell us about the emergence of DP. In S. C. Howell, S. A. Fish & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development* (Vol. 2, pp. 609–620). Somerville, MA: Cascadilla Press.

Schöler, H., Fromm, W., & Kany, W. (1998a). Die Spezifische Sprachentwicklungsstörung - eine sprachspezifische Störung? In H. Schöler, W. Fromm & W. Kany (Eds.), *Spezifische Sprachentwicklungsstörung und Sprachlernen - Erscheinungsformen, Verlauf, Folgerungen für Diagnostik und Therapie* (pp. 275-294). Heidelberg: Universitätsverlag C. Winter.

Schöler, H., Fromm, W., & Kany, W. (Eds.). (1998b). *Spezifische Sprachentwicklungsstörung und Sprachlernen : Erscheinungsformen, Verlauf, Folgerungen für Diagnostik und Therapie*. Heidelberg: Programm Ed. Schindele im Universitätsverlag C. Winter.

Schrey-Dern, D. (2006). *Sprachentwicklungsstörungen: Logopädische Diagnostik und Therapieplanung*. Stuttgart: Thieme.

Schul, R., Stiles, J., Wulfeck, B., & Townsend, J. (2004). How 'Generalized' Is the 'Slowed Processing' in SLI? The Case of Visuospatial Attentional Orienting. *Neuropsychologia, 42*(5), 661-671.

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

Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language & Communication Disorders, 45*(6), 691-702.

Seeff-Gabriel, B., Chiat, S., & Pring, T. (2012). Intervention for co-occurring speech and language difficulties: Addressing the relationship between the two domains. *Child Language Teaching and Therapy, 28*(1), 123-135.

Serratrice, L., Joseph, K. L., & Conti-Ramsden, G. (2003). The acquisition of past tense in preschool children with specific language impairment and unaffected controls: regular and irregular forms. *Linguistics 41, no. 2*.

Shriberg, L. D. (1994). Five subtypes of developmental phonological disorders. *Clinics in Communication Disorders, 4*(1), 38-53.

Shriberg, L. D., & Austin, D. (1998). Comorbidity of speech-language disorder: Implications for a phenotype marker for speech delay. In R. Paul (Ed.),

*Exploring the speech-language connection* (pp. 73-117). Baltimore: Paul H. Brookes Publishing Co.
Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders: I. A clinical profile. *Journal of Speech & Hearing Research, 37*(5), 1100-1126.
Shriberg, L. D., Lewis, B. A., Tomblin, J. B., McSweeny, J. L., Karlsson, H. B., & Scheer, A. R. (2005). Toward diagnostic and phenotype markers for genetically transmitted speech delay. *Journal of Speech, Language, and Hearing Research, 48*(4), 834-852.
Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research, 42*(6), 1461-1481.
Sick, B. (2004). *Der Dativ ist dem Genitiv sein Tod*. Köln: Kiepenheuer & Witsch.
Silveira, M. (2011). *Specific Language Impairment (SLI) revisited: evidence from a psycholinguistic investigation of grammatical gender abilities in Brazilian Portuguese-speaking children.* University College London, London.
Simkin, Z., & Conti-Ramsden, G. (2006). Evidence of reading difficulty in subgroups of children with specific language impairment. *Child Language Teaching & Therapy, 22*(3), 315-331.
SLI-Consortium. (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *American journal of human genetics, 70*(2), 384-398.
SLI-Consortium. (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *American Journal of Human Genetics, 74*(6), 1225-1238.
Snowling, M., Chiat, S., & Hulme, C. (1991). Words, Nonwords, Phonological Processes: Some Comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12*(3), 369-373.
So, L., & Dodd, B. (1994). Phonologically disordered Cantonese-speaking children. *Clinical Linguistics & Phonetics, 8*, 235-255.
Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility criteria for language impairment: is the low end of normal always appropriate? *Language, Speech & Hearing Services in the Schools, 37*(1), 61-72.
Stackhouse, J., & Wells, B. (1997). *Children's Speech and Literacy Difficulties: a Psycholinguistic Framework*. London: Whurr.
Stark, R. E., & Tallal, P. (1981). Selection of Children with Specific Language Deficits. *Journal of Speech and Hearing Disorders, 46*(2), 114-122.
Stokes, S. F., Wong, A. M. Y., Fletcher, P., & Leonard, L. B. (2006). Nonword Repetition and Sentence Repetition as Clinical Markers of Specific Language Impairment: The Case of Cantonese. *J Speech Lang Hear Res, 49*(2), 219-236.
Swisher, L., Plante, E., & Lowell, S. (1994). Nonlinguistic deficits of children with language disorders complicate the interpretation of their nonverbal IQ scores. *Language, Speech, and Hearing Services in Schools, 25*(4), 235-240.
Szagun, G. (2004). Learning by ear: On the acquisition of case and gender marking by German-speaking children with normal hearing and with cochlear implants. *Journal of Child Language, 31*(1), 1-30.

Szagun, G. (2006). *Sprachentwicklung beim Kind. Ein Lehrbuch.* Weinheim: Beltz.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech Language & Hearing Research, 42*(5), 1275-1278.

Tallal, P., Hirsch, L. S., Realpe-Bonilla, T., Miller, S., & Brzustowicz, L. M. (2001). Familial aggregation in specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(5), 1172-1182.

Tallal, P., & Piercy, M. (1973a). Defects of Non-Verbal Auditory Perception in Children with Developmental Aphasia. *Nature, 241*, 468 - 469.

Tallal, P., & Piercy, M. (1973b). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia, 11*(4), 389-398.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia, 12*(1), 83-93.

Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia, 13*(1), 69-74.

Taylor Kulp, M. (1999). Relationship between Visual Motor Integration Skill and Academic Performance in Kindergarten through Third Grade. *Optometry and Vision Science, 76*(3), 159-163.

Thal, D. J., Miller, S., Carlson, J., & Vega, M. M. (2005). Nonword Repetition and Language Development in 4-Year-Old Children With and Without a History of Early Language Delay. *Journal of Speech, Language, and Hearing Research, 48*(6), 1481-1495.

Thomas, M. S. C. (2005). Constraints on language development: Insights from developmental disorders. In P. Fletcher & J. F. Miller (Eds.), *Developmental theory and language disorders* (4 ed.). Amsterdam: Jon Benjamins Publishing Company.

Thomas, M. S. C., & Karmiloff-Smith, A. (2005). Can developmental disorders reveal the component parts of the human language faculty? *Language Learning and Development, 1*(1), 65-92.

Thordardottir, E. T., & Namazi, M. (2007). Specific Language Impairment in French-Speaking Children: Beyond Grammatical Morphology. *Journal of Speech, Language, and Hearing Research, 50*(3), 698-715.

Tomasello, M. (2000). Acquiring syntax is not what you think. In D. M. V. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome.* (pp. 1-15). New York, NY: Psychology Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA, US: Harvard University Press.

Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: how much is enough? . *Journal of Child Language, 31*(1), 101-121.

Tomblin, B. J. (1996). Genetic and environmental contributions to the risk for specific language impairment,. In M. L. Rice (Ed.), *Toward a genetics op language* (pp. 191-210). Mahwah: Lawrence Erlbaum Associates.

Tomblin, J. B., Freese, P. R., & Records, N. L. (1992). Diagnosing specific language impairment in adults for the purpose of pedigree analysis. *Journal of Speech and Hearing Research, 35*(4), 832-843.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech Language & Hearing Research, 40*(6), 1245-1260.

Tomblin, J. B., Records, N. L., & Zhang, X. Y. (1996). A System for the Diagnosis of Specific Language Impairment. *Journal of Speech and Hearing Research, 39*, 1284-1294.

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 361-382). Cambridge, MA: MIT Press.

Tomblin, J. B., & Zhang, X. (2006). The Dimensionality of Language Ability in School-Age Children. *Journal of Speech, Language, and Hearing Research, 49*(6), 1193-1208.

Tomblin, J. B., Zhang, X., Buckwalter, P., & O'Brien, M. (2003a). The Stability of Primary Language Disorder: Four Years after Kindergarten Diagnosis. *Journal of Speech, Language, and Hearing Research, 46*(6), 1283-1296.

Tomblin, J. B., Zhang, X., Buckwalter, P., & O'Brien, M. (2003b). The Stability of Primary Language Disorder: Four Years After Kindergarten Diagnosis. *Journal of Speech, Language and Hearing Research, 46*, 1283-1296.

Trauner, D., Wulfeck, B., Tallal, P., & Hesselink, J. (2000). Neurological and MRI profiles of children with developmental language impairment. *Developmental Medicine & Child Neurology, 42*(7), 470-475.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The Procedural Deficit Hypothesis. *Cortex. Special Issue: The Neurobiology of Developmental Disorders, 41*(3), 399-433.

Uwer, R., Albrecht, R., & von Suchodoletz, W. (2002). Automatic processing of tones and speech stimuli in children with specific language impairment. *Developmental Medicine and Child Neurology, 44*(8), 527-532.

van der Lely, H. K. J. K. J. (1997). Language and cognitive development in a grammatical SLI boy: Modularity and innateness. *Journal of Neurolinguistics, 10*(2-3), 75-107.

van der Lely, H. K. J. K. J. (2003). Do heterogeneous deficits require heterogeneous theories? SLI subgroups and the RDDR hypothesis. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment.* (pp. 97-133). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

van der Lely, H. K. J. K. J. (2005). Domain-specific cognitive systems: Insight from Grammatical-SLI. *Trends in Cognitive Sciences, 9*(2), 53-59.

van der Lely, H. K. J. K. J., & Battell, J. (2003). WH-Movement in Children with Grammatical SLI: A Test of the RDDR Hypothesis. *Language, 79*(1), 153-181.

van der Lely, H. K. J. K. J., & Howard, D. (1993). Children with specific language impairment: Linguistic impairment or short-term memory deficit? *Journal of Speech & Hearing Research, 36*(6), 1193-1207.

van der Lely, H. K. J. K. J., Rosen, S., & McClelland, A. (1998). Evidence for a grammar-specific deficit in children. *Current Biology, 8*(23), 1253-1258.

van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry, 2*(47), 176-189.

Verhoeven, L., & van Balkom, H. (2004). Developmental language disorders: Classification, assessment, and intervention. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 3-20). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Watkins, R. V., Rice, M. L., & Moltz, C. C. (1993). Verb Use by Language-Impaired and Normally Developing Children. *First Language, 1*(37), 133-143.

Weismer, S. E. (2007). Typical talkers, late talkers, and children with specific language impairment: A language endowment spectrum? In R. Paul (Ed.), *Language Disorders From a Developmental Perspective: Essays in Honour of Robin S. Chapman* (pp. 83-101). Mahwah, NJ: Lawrence Erlbaum Associates.

Weissenborn, J. (2000). Der Erwerb von Morphologie und Syntax. In H. Grimm (Ed.), *Sprachentwicklung* (pp. 141-169). Göttingen: Hogrefe.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*(3), 558-584.

Wilson, B. C., & Risucci, D. A. (1986). A model for clinical-quantitative classification: Generation I: Application to language-disordered preschool children. *Brain and Language, 27*(2), 281-309.

Wolfus, B., Moscovitch, M., & Kinsbourne, M. (1980). Subgroups of developmental language impairment. *Brain and Language, 10*(1), 152-171.

Wong, A. M.-Y., Klee, T., Stokes, S. F., Fletcher, P., & Leonard, L. B. (2010). Differentiating Cantonese-Speaking Preschool Children With and Without SLI Using MLU and Lexical Diversity (D). *Journal of Speech, Language, and Hearing Research, 53*, 794-799.

World Health Organisation. (2010). International Statistical Classification of Diseases and Related Health Problems 10. Retrieved 06/12/2011, 2011

Zhang, X., & Tomblin, J. B. (2000). The association of intervention receipt with speech-language profiles and social-demographic variables. *American Journal of Speech-Language Pathology, 9*(4), 345-357.

Zhu Hua, & Dodd, B. (2000). Putonghua (modern standard Chinese)-speaking children with speech disorder. *clinical linguistics & phonetics, 14*(3), 165191.

# Appendix 1 Phonological Diagnosis for Clinical Groups

SLI-Group:

**Table 39.** Active Phonological Processes of Participants with SLI

| Code | Age | Phonological processes |
|---|---|---|
| SLI1 | 3;0 | fronting of [R], assimilations, deletion of unstressed syllables |
| SLI2 | 3;0 | - (not enough vocabulary) |
| SLI3 | 3;0 | - (not enough vocabulary) |
| SLI4 | 3;0 | reduction of CC (including [st]); deletion of word-final [l] in [əl]; fronting of [k] , [g], [ŋ]; inconsistent stopping of [f]; inconsistent deletion of word-final consonants |
| SLI5 | 3;1 | inconsistent deletion of unstressed syllables; fronting of [ʃ]; inconsistent deletion of word-final consonants (including [st]) |
| SLI6 | 3;6 | word-initial sound preference [h]; fronting of [ʃ]; reduction of CC |
| SLI7 | 3;7 | sound preference [h]: word-initial [s, f, z, v, j]→ [h]; reduction of CC ([st] correct); fronting of [k], [g], [ŋ], [ʃ], [ç]; inconsistent deletion of unstressed syllables |
| SLI8 | 4;0 | Assimilations; fronting of [ʃ]; [R]→ [h] |
| SLI9 | 4;1 | - (not enough vocabulary) |
| SLI10 | 4;1 | reduction of CC ([st] correct); voicing of [f]; fronting of [k] , [g], [ŋ], [ʃ], [ç] + [χ] |
| SLI11 | 4;1 | reduction of CC (including [st]), fronting [k] , [g], [ŋ] and [ʃ] |
| SLI12 | 4;3 | Assimilations; inconsistent reduction of CC ([st] correct), inconstant fronting of [k], [g], [ŋ] |
| SLI13 | 4;8 | deletion of word-final [l] in [əl] or [l] added, inconsistent reduction of CC, fronting of [k] / [g] in CC |
| SLI14 | 4,9 | fronting of [k], [g], [ŋ], [r]; reduction of CC ([st] correct); deletion of unstressed syllables; assimilations |
| SLI15 | 5;0 | fronting of [ʃ], reduction of CC with [R] |
| SLI16 | 5;0 | - |
| SLI17 | 5;0 | - |

## PI-Group:

**Table 40.** Active Phonological Processes of Participants with PI

| Code | Age | Phonological Processes |
|---|---|---|
| PI1 | 3;0 | fronting [k], [g], [ŋ]; stopping / deletion of fricatives (including [st]); reduction of CC |
| PI2 | 3;4 | backing of labials; inconsistent stopping of fricatives ([s/z] sometimes correct); fronting of [ʃ]; inconsistent reduction of CC (including [st]) |
| PI3 | 3;1 | fronting [k], [g], [ŋ] and inconsistent of [ʃ]; reduction of CC ([st] correct) |
| PI4 | 3;3 | backing of [t] , [d] and inconsistent of [s] and [f]; reduction of CC (including [st]); assimilations; inconsistent deletion of word-final consonants; deafrication; s → sometimes h |
| PI5 | 4;1 | word-initial [R]→ [h]; reduction CC ([st] correct); inconsistent word-initial stopping of [f], [v], [z] and [ʃ]; inconsistent fronting of [k], [g] and [ʃ] |
| PI6 | 3;7 | stopping of fricatives or substitution with [h] in word-initial position; inconsistent reduction of CC; fronting of [k], [g], [ŋ] |
| PI7 | 3;8 | fronting [ʃ]; inconsistent backing of [f], [v]; fronting of [k], [g], [ŋ] |
| PI8 | 4;5 | assimilations |
| PI9 | 3;11 | sound preference [t]; fronting of [k], [g], [ŋ]; fronting of [ʃ]; [R]→ [ʕ]; reduction of CC (including [st]) |
| PI10 | 3;10 | fronting of [ʃ]; deletion of word-final [l] in [əl]; reduction of CC with [b], [f], [k] + [g]; deafrication; deletion of final [t] in [st]-clusters; inconsistent fronting of [ç] + [χ] |
| PI11 | 4;7 | fronting of [k], [g], [ŋ] |
| PI12 | 4;11 | fronting of [ʃ]; inconsistent fronting of [f] |
| PI13 | 4;6 | fronting of [ʃ]; deletion of word-final [l] in [əl]; inconsistent stopping of [f] and [ʃ] |
| PI14 | 5;2 | fronting of [k], [g], [ŋ]; [R]→ [h]; reduction of CC (including [st]); deafrication |
| PI15 | 5;0 | fronting of [ʃ]; reduction of CC with [k] / [g]; inconsistent fronting of [k] |
| PI16 | 5,1 | inconsistent deletion of word-final consonants ([st] correct); devoicing of [b], [g] + [d] in CC |

# Appendix 2 Morphologie Untersuchung (German Version)

**Name:** **Geb.:**

## 1. Verbflektionen

**3. Person Singular:**

**Material:** Bildkarten - Verben

**Prüfer:** **Was macht das Mädchen denn da? Das Mädchen ...**

| Zielstruktur | Reaktion |
|---|---|
| Testitem: Das Mädchen trinkt. | |
| 1. Das Mädchen schaukelt. | |
| 2. Der Junge spielt Fußball. | |
| 3. Das Mädchen füttert die Katze. | |
| 4. Der Junge kauft ein Eis. | |
| 5. Der Junge baut einen Turm. | |

**2. Person Singular:**

Der Prüfer stellt Verben pantomimisch dar und das Kind soll raten.

**Prüfer:** **Schau mal, was mache ich?** (Stellt Verb pantomimisch dar.)
**Jetzt kannst du sagen: „Du ...“** (Prüfer richtet Finger auf sich selbst.)

Wenn Verb nur im Infinitiv produziert wird, dann kann der Prüfer es in Form von korrektivem Feedback wiederholen. Genau: „Du trinkst.“ (Zeigt dabei auf sich selbst.)

| Zielstruktur | Reaktion |
|---|---|
| Testitem: Du trinkst. | |
| 1. Du schreibst. | |
| 2. Du weinst. | |
| 3. Du schwimmst. | |
| 4. Du lachst. | |
| 5. Du telefonierst. | |

## 2. Dativartikel

**Material:** Tierpuzzle – 2x liegen jeweils 4 Tierkörper (kann geändert werden) vor
dem Kind

**Prüfer:** Reicht dem Kind einen Tierkopf

**Wem gehört denn der Kopf? (Der Kopf gehört ...)**

| Zielstruktur | Reaktion |
|---|---|
| Testitem: dem Schwein<br>dem Fisch | |
| 1. der Kuh | |
| 2. dem Schaf | |
| 3. dem Teddy | |
| 4. dem Krokodil | |
| 5. der Katze | |
| 6. dem Hund | |

## 3. Hilfsverben

**<u>Material:</u>** Bildergeschichten à 3 Bilder

**<u>Prüfer:</u>**

Zeigt auf 1. Bild: **Schau mal, hier hat der Junge nichts.**

Zeigt auf 3. Bild: **... und hier hat er ein Eis. Was hat der Junge gemacht?**

Zeigt auf 2. Bild: **Der Junge ....**

| **Zielstruktur** | **Reaktion** |
|---|---|
| <u>Testitem:</u> Der Junge hat ein Eis gekauft. | |
| 1. Das Mädchen hat gegossen. | |
| 2. Das Mädchen hat ihm die Wurst gegeben. | |
| 3. Der Junge ist hingefallen. | |
| 4. Der Ballon ist weggeflogen. | |

## 4. Artikel

### Nominativ

**Material:** Lotto

**Prüfer:** Zeigt auf die Bilder auf der Spieltafel:

**Was kommt denn da hin? (Da kommt ...)**

| Zielstruktur | Reaktion |
|---|---|
| Testitem: der Mond<br>der Baum | |
| 1. die Sonne | |
| 2. das Bett | |
| 3. die Schnecke | |
| 4. der Schuh | |
| 5. das Schwein | |
| 6. der Elefant | |

## Akkusativ

**<u>Material:</u>** Lotto und Handpuppe

**<u>Prüfer:</u>** **Jetzt müssen wir die Bilder da drauf legen.**
**Die kann uns der Fips geben.**
**Dazu müssen wir ihm sagen, was er uns geben soll.**

**Fips, gib mir ... (den Mond).**

| **Zielstruktur** | **Reaktion** |
|---|---|
| <u>Testitem:</u> den Mond<br>den Baum | |
| 1. die Sonne | |
| 2. das Bett | |
| 3. die Schnecke | |
| 4. den Schuh | |
| 5. das Schwein | |
| 6. den Elefant(en) | |

# Appendix 3 Morphology Probes (English Translation)

**Name:** **DoB.:**

## 1, Verb inflections[22]

**Third person singular:**

**Material:** picture cards - verbs

**Examiner:** **What is the girl doing here? The girl ...**

| Target | Reaction |
|---|---|
| Practice: The girl drinks. | |
| 6. The girl swings. | |
| 7. The boy plays football. | |
| 8. The girl feeds the cat. | |
| 9. The boy buys ice-cream. | |
| 10. The boy builds a tower. | |

**Second person singular:**

The examiner pantomimes the verbs and the child is asked to guess what the examiner is doing.

**Examiner:** **Look, what am I doing?** (pantomimes the verb)
**You can say now: „You ...“** (The examiner points to herself.)

If the child produces only the infinitive the examiner tries to prompt the inflected form by using corrective feedback: “Exactly. ‘You drink.’” (pointing to herself)

[22] Note: English would require present tense progressive rather than present tense in this context. However, as the German target is present tense it was also used for the translation to English.

| Target | Reaction |
|---|---|
| <u>Practice:</u> You drink. | |
| 6. You write. | |
| 7. You cry. | |
| 8. You swim. | |
| 9. You laugh. | |
| 10. You phone. | |

## 2. Dative article

**<u>Material:</u>** Animal jigsaw – 4 animal bodies are lying in front of the child (flexible quantity)

**<u>Examiner:</u>** Hands the child one of the animals' heads.

**Who does this head belong to? (This head belongs to …?)**

| Target | Reaction |
|---|---|
| <u>Practice:</u> the pig<br>the fish | |
| 7. the cow | |
| 8. the sheep | |
| 9. the Teddy | |
| 10. the crocodile | |
| 11. the cat | |
| 12. the dog | |

## 3. Auxiliaries

**Material:** picture stories à 3 pictures each

**Examiner:**

Pointing to 1st picture: **Look, this boy doesn't have anything here.**

Pointing to 3rd picture: **…but now he has an ice-cream. Why does he have an ice-cream now? What has he done?**[23]

Pointing to 2nd picture: **The boy …?**

| Target | Reaction |
|---|---|
| Practice: The boy has bought an ice-cream. | |
| 5. The girl has watered (the plant). | |
| 6. The girl has given him a sausage. | |
| 7. The boy has fallen.[24] | |
| 8. The balloon has flown away.[25] | |

## 4. Articles

### Nominative

**Material:** Bingo game

**Examiner:** Pointing successively to the pictures on the bingo board:

**Which picture goes here? (Here ...)**

[23] In contrast to the required past tense in English, present perfect is commonly used in spoken German to express events in the past.

[24] Auxiliary 'to be' in German

[25] Auxiliary 'to be' in German

| Target | Reaction |
|---|---|
| Practice: the moon<br>the tree | |
| 7. the sun | |
| 8. the bed | |
| 9. the snail | |
| 10. the shoe | |
| 11. the pig | |
| 12. the elephant | |

**Accusative**

**Material:** Bingo game and hand puppet

**Examiner:** **Now we need to put the little cards onto the board.**
**Fips can give them to us.**
**But we need to tell him what to give us.**
**Fips, give me….**

| Target | Reaction |
|---|---|
| Practice: the moon<br>the tree | |
| 1. the sun | |
| 2. the bed | |
| 3. the snail | |
| 4. the shoe | |
| 5. the pig | |
| 6. the elephant | |

# Appendix 4 Guidelines for Utterance Segmentation

## 1. Clean up

All hesitations, false starts and backtrackings are marked within mazes and, therefore, will be excluded from the MLU calculation. Only if the content of the following utterance is different to the previous incomplete one the sentence will be marked as abandoned (>).

The expressions “oh”, “ah” “na”, etc. are generally marked within mazes.
e.g. (Oh) Nein. (Na) Gut.

“Achso” is considered as single, but separate utterance.

## 2. Segmentation

One unit can consist of

- Single words if clearly indicated as one-word-utterance by context, intonation or pauses.
  e.g. as answer, in commands, etc.
- Ellipses if not in major unit cooperated.
- Simple main clauses.
- Main clauses with subordinate clause

*Complement* is used for syntactically obligatory elements, *adjunct* refers to clause elements that are optional (Garman, 1989).

a. Finite + non-finite complement or adjunct:
   e.g.: Ich hab vergessen, es auf zu bauen.

b. Finite and finite complement:

e.g.: Ich glaube dass wir da auch was drauf laden können.

c. Finite + finite adjunct:
   e.g.: Da lauf ich auch immer vorbei wenn ich zur U-bahn gehe.

<u>Special cases:</u>

*And*

The connective '*und'* (engl.: and) groups two minor text-units together if the second one cannot stand alone but is only an ellipse.

e.g.: Ich hol mir einen Stift und mal ein Bild.

If two complete sentences are linked by *und* they are counted as two utterances.

e.g.: Ich hol mir einen Stift.
Und dann mal ich ein Bild.

*Look / Guck mal* (Klee, 1992)

If '*guck mal'* is followed by an independent clausal text-unit (i.e. it can stand alone) both units should be separated.

e.g.: Guck mal.
Da ist eine Giraffe.

Guck mal.
Was ist das denn?

If '*guck mal'* is followed by a dependent clausal text-unit both units should be grouped together as one utterance.

e.g.: Guck mal eine Giraffe.
Guck mal was das ist.

## 3. Word segmentation

Names will be considered as one morpheme: e.g.  OmaErika, TanteUte
Aber: Herr Müller

Adverbs containing the word stem '*-hin'* or *'-mal'*  are segmented into two morphemes: e.g. noch mal, ein mal, dort hin, da hin

# Appendix 5 SALT Transcription Convention for German

## Verb Phrase

### Present Tense

| | | | |
|---|---|---|---|
| **1st person sing.:** | stem + /1e | **1st person pl.:** | stem+ /1en |
| **2nd person sing.:** | stem+ / st | **2nd person pl.:** | stem+ /2t |
| **3rd person sing.:** | stem+ / 3t | **3rd person pl.:** | stem+ /3en |

**Infinitive:** stem+ /en

### Present Perfect

→ Auxiliary + Past participle

Past participle:

a) weak verbs: ge+ /stem+ /t2
b) strong verbs: ge+ |stem+ /changed-stem+ /en2 (+ vowel change in tem)

Exp.: ge|geh/gang/en2

### Past Tense

| | | | |
|---|---|---|---|
| **1st person sing.:** | stem+ /1te | **1st person pl.:** | stem+ /1ten<br>stem+ /Pen |
| **2nd person sing.:** | stem+ /Pst | **2nd person pl.:** | stem+ /tet<br>stem+ /Pt |
| **3rd person sing.:** | stem+ /3te | **3rd person pl.:** | stem+ /3ten<br>stem+ /P3en |

## Conditional

| | |
|---|---|
| **1st person sing. / 3rd person sing.:** | stem+ /Cte |
| **2nd person sing.:** | stem+ /Ctest |
| **1st person pl. / 3rd person pl.:** | stem+ /Cten |
| **2nd person pl.:** | stem+ /Ctet |

## Imperative /eI

## Noun Phrase

### Plural

/nPl

/enPl

/sPl

/ePl

/erPl

/ienPl

/ Pl (vowel change)

/0Pl (zero ending)

### Forming adverb

/z (e.g. mittag/z)

Other derivations:

/in (feminine form)

## Case

**Dative**

Singular: /enCNoun

Plural: /nC

**Accusative**

Singular: /enCNoun

Plural: /nC

**Genitive**

Only in singular: /esGen

Possesive: /s

# Articles

## Gender

Definite article: der (masc.), die (fem.), das (neuter)
Infinite article: ein (masc. & neuter)    single morpheme
eine (fem.)

Instead of noun: ein/erM
ein/sN

## Case

### Dative

Definite article: dem (masc. & neuter), der (fem.), den (plural) → single morph.
den/enD (plural-pronoun)

Infinite article: ein/emDM (masc.)
ein/er D (fem.)
ein/emDN (neuter)

### Accusative

Definite article: den (masc.), die (fem.), das (neuter), den (plural) → single m.
Infinite article: ein/enA (masc.)
eine (fem.)
ein (neuter)    single morpheme

### Genitive

Definite article: des (masc. & neuter), der (fem.), den (plural) → single morph.
Infinite article: ein/esGen (masc. & neuter)
ein/erGen (fem.)

# Adjectives

## Gender

Following indefinite article:

| Gender | Nominative | Accusative | Dative |
|---|---|---|---|
| Masculine | /erG | /enC | /emDAd |
| Feminine | /eG | /eGA | /erDAd |
| Neuter | /es | /esA | /emDAd |
| Plural | /ePlAd | /ePlAdA | /ePlAdD/N |

Following definite article:

| Gender | Nominative | Accusative | Dative |
|---|---|---|---|
| Masculine | /E | /enCw | /enCw |
| Feminine | /E | /E | /enCw |
| Neuter | /E | /E | /enCw |
| Plural | /ePlAd /N | /ePlAdA /N | /ePlAdD/N |

Possessive pronouns are marked as adjectives.

## Comparison

Comparative: /er

Superlative: /ste

## Prefixes

**Separable prefixes:**

Counted as extra morpheme.

## Contractions

| | | | |
|---|---|---|---|
| **‘s** | contracted pronoun ‘es’ | **‘M** | dem |
| **‘S** | ‘ist’ as Copula ‘sein’ | **‘N** | den |
| **/S** | contracted ‘das’ | **zu/’M** | zum |
| **‘ne** | eine | **an/’M** | am |
| **‘n** | ein | **in/’M** | im |
| **‘n/emD(M/N)** | Contracted dative determiner einem | **/s** | Possesive s |
| **‘n/erD** | Contracted dative determiner einer | **‘nn** | denn / dann |

## Compounds

Counted as one morpheme if the compound is frequently used.

Counted as two morphemes if the two nouns are more commonly used separately than as compound.

“Bindungs-n”: /n/

“Bindungs-s”: /ß/

## Correct marking

| | |
|---|---|
| CS | Copula “sein” |
| M | Modal verb |
| AuxS | Auxiliary “sein” |
| AuxH | Auxiliary “haben” |
| AuxW | Auxiliary “werden” - passive |
| VP1 | Verb position: first element |
| VP2 | Verb position: second element |
| VP02 | Verb position: second element but first was omitted |
| VPF | Verb position: clause final element |
| CN | Case-marking: nominative |
| CN-M | Case-marking: nominative – masculine definite article |
| CN-MI | Case-marking: nominative – masculine indefinite article |
| CN-Pl | Case-marking: nominative – plural |
| CN-P-M | Case-marking: nominative – masculine definite article instead of pronoun |
| CA | Case-marking: accusative (and further coding as above following example nominative) |
| CD | Case-marking: dative (and further coding as above following example nominative) |
| CAP | Case-marking: accusative personal pronoun (e.g. ‘mich’) |
| CDP | Case-marking: dative personal pronoun (e.g. ‘mir’) |

## Error Codes

| **Code** | Error |
|---|---|
| EG | Wrong gender |
| EG-X-XX | Format: EG-Case-target gender+wrong gender<br>Example: EG-N-FM (nominative feminine substituted with masculine) |
| EGAd | Error in gender marking in adjective (e.g. ‘linke’ –‘links’) |
| EC | Wrong case |

| | |
|---|---|
| ECXX-X | Format: EC+target case+wrong case-gender<br>Example: ECDA-M: Dative masculine article substituted with accusative |
| ECAd | Error in case marking in adjective |
| EC+G-CC-GG | Error in case and gender marking (first target, second substitute) |
| EPl | Wrong plural |
| EPlVV | Error in vowel change of irregular plural inflection |
| EN | Error number (e.g. determiner) |
| EL | Wrong lexical item |
| ELF | Wrong lexical item – wrong function word |
| EVIVV-PT | Error irregular: vowel change – in present tense verb inflection |
| EVIVV-PP | Error irregular: vowel change – in present participle |
| EVT | Wrong tense |
| EVI | Wrong verb inflection |
| EVIV | Infinite verb |
| EVAux–HS/SH | Wrong auxiliary (first target, second substitute) |
| EVPP | Error verb – error in past participle |
| EVbast | Error verb – only bare stem of irregular verbs |
| EVsub | Error verb – verb substituted (e.g. 'bo' = trinken (to drink)) |
| EVpassive | Error verb – error in passive construction |
| EVO | Omission verb |
| EVOIV | Error verb – omission infinitive |
| EVOI | Error verb – omission inflected verb |
| EVOM | Omission modal verb |
| EVOCS | Omission copula "sein" |
| EVOAux | Omission auxiliary |
| EVOPP | Omission past participle |
| EVOge | Omission prefix /ge-/ in past participle |
| EVP | Error verb position |

| | |
|---|---|
| EVPF | Final verb position |
| EVIT | Wrong verb inflection in tense marking |
| EV2PSt | 2. PS only /s/ |
| EVAS | Error verb argument structure |
| EOAd | Omission adverb |
| EOPart | Omission particle |
| EOQW | Omission question word |
| EOS | Omission subject |
| EOO | Omission object |
| EOPrep | Omission preposition |
| EOP | Omission pronoun |
| EOC | Omission conjunction word |
| EWO | Wrong word order |
| ED | Determiner (definite/ indefinite → congruence with adjective) |
| EDO-Case+Gender | Omission determiner |
| EDsub | Determiner substituted (e.g. /de/) |
| EAd | Adjective- congruence with noun |
| EPr | Error in use of pronouns (e.g. ihr/es instead of ihr) |
| EPrSub | Pronoun substituted |
| Ecom | Error in comparison – e.g. wrong comparative |
| EPFCD | Final consonant deletion |
| EPDUS | Deletion of unstressed syllable |
| EPSubUS | Unstressed syllable substituted – e.g. ka'putt -> nputt |
| EP-R-H (example) | Phonological error: e.g. /r/ substituted with /h/ |
| ESubW | Word substitute- e.g. "nicht" -> n |
| Ege | Prefix /ge/ replaced (e.g. [illegible] ) |
| EPCC | Error in phonology of consonant cluster |
| Unf | Disfluency |

## Ambiguous cases

| | |
|---|---|
| AmCS | Ambiguous whether contracted form of copula 'sein', "ist", was produced |
| Am3PS | Ambiguous whether 3rd person singular verb inflection was produced |
| Am2PSt | Ambiguous whether 2nd person singular was signalled by /-st/ or only by /-s/ |
| Am2PS | Ambiguous whether inflection or not – verb stem ending with /s/ |
| AmPPt | Ambiguous whether word final /t/ of participle perfect is produced |
| AmVI | Ambiguous whether verb is inflected / which inflection |
| AmC | Ambiguous whether case was signalled |
| AmG | Ambiguous whether gender was signalled |
| AmC/G | Ambiguous whether gender or case correct / incorrect |
| AmD | Ambiguous whether determiner present (e.g. in/'n) |
| AmPl | Ambiguous whether plural was signalled / correct plural |
| AmPos | |
| AmVP | Verb position ambiguous |
| AmWE | |
| AmV | Ambiguous whether word functions as verb |
| AmL | Ambiguous lexical (which function) |
| AmAuxS | Ambiguous whether contracted form of auxiliary 'sein', "ist", was produced |
| Cind | Case marking indeterminable |

## Verb roots necessary / inflections not separated

| | | |
|---|---|---|
| haben | können | mögen |
| sein | wollen | |
| werden | müssen | |
| dürfen | möchten | |

If past tense inflected like other verbs → inflection counted in past tense.

E.g.: woll/1te; muss/3te; moch/tet; durf/3ten

# Appendix 6 Letter to Parents of Participants in Clinical Groups

Sehr geehrte Eltern,

nach drei Jahren praktischer Tätigkeit in der Logopädischen Praxis Grosstück in Hamburg bin ich nach England gegangen, um mich dort an der Universität Newcastle upon Tyne weiterzubilden. Nun schreibe ich an meiner Doktorarbeit. Ziel dieser Forschungsarbeit ist es, einen Beitrag zur Suche eines Leitsymptoms für Spezifische Sprachentwicklungsstörungen (SSES) zu leisten. Studien in anderen Sprachen haben gezeigt, dass die meisten Kinder mit einer SSES auch Probleme im grammatikalischen Bereich zeigen. Eine Fortführung dieser Studien könnten die Diagnose von SSES wesentlich erleichtern und zeitlich verkürzen, was allen Beteiligten (Kindern, Therapeuten, Eltern) zu Gute kommen würde.

Aus diesem Grunde möchte ich Sie bitten, mit Ihrem Kind an meiner Studie teilzunehmen. Ich habe im letzten Jahr Kinder getestet, die Sprachtherapie benötigen, und möchte ihre Daten nun mit Daten von sprachunauffälligen Kindern vergleichen, die ungefähr im gleichen Sprachentwicklungsalter sind. Dies würde bedeuten, dass Sie Ihr Einverständnis zu einer ausführlichen logopädischen Diagnostik bei ihrem Kind geben. Außer den Tests, die notwendig sind, um den genauen Sprachstatus ihres Kindes zu bestimmen, würde eine 20minütige Spontansprachaufnahme gemacht und später analysiert werden. Für Sie und Ihr Kind bestünde der Vorteil, dass Sie durch diese Untersuchungen einen Überblick über den sprachlichen Entwicklungsstand Ihres Kindes erhalten. Für den Fall, dass bei Ihrem Kind sprachliche Auffälligkeiten in irgendeiner Form im Rahmen der Untersuchung festgestellt würden, erhalten Sie eine Beratung, was am besten zu tun ist. Die Daten der Studie werden zeigen können, ob Auffälligkeiten in der Grammatik (z.B. Artikelgebrauch, Verbendungen) es vermögen, zwischen Kindern mit und ohne SSES zu unterscheiden. Eine genaue Diagnostik und das bessere Verstehen von Sprachstörungen sind ein wichtiger Schritt in der Forschung, um Kindern mit Sprachproblemen besser und gezielter helfen zu können. Mit Ihrer Unterstützung würden Sie einen wichtigen Beitrag leisten.

Selbstverständlich werden alle erhobenen Daten streng anonym behandelt und gespeichert. Der Name Ihres Kindes wird nicht verwendet werden. Außerdem steht Ihnen natürlich jederzeit das Recht zu, von dem Forschungsprojekt zurückzutreten.

Wenn Sie einverstanden sind, an der Studie teilzunehmen und Ihr Kind **keine Ihnen bekannte Hörstörung** hat, bitte ich Sie, untenstehende Einverständniserklärung zu unterschreiben.

Mit freundlichen Grüßen und bestem Dank, J. Girndt, MSc

---

Ich erkläre mich hiermit einverstanden, dass die Daten meiner Tochter / meines Sohn ________________________________________(Name)________________ (Geb.) für Forschungszwecke ohne Namensnennung verwandt werden dürfen. Während der Erhebung der Testdaten darf eine Audioaufnahme der Äußerungen meines Kindes gemacht werden.

_______________________________ ___________________________________

Ort, Datum Unterschrift

# Appendix 7 Letter to Parents of Participants in Control Group

Sehr geehrte Eltern,

nach drei Jahren praktischer Tätigkeit in der Logopädischen Praxis Grosstück in Hamburg bin ich nach England gegangen, um mich dort an der Universität Newcastle upon Tyne weiterzubilden. Nun schreibe ich an meiner Doktorarbeit. Ziel dieser Forschungsarbeit ist es, einen Beitrag zur Suche eines Leitsymptoms für Spezifische Sprachentwicklungsstörungen (SSES) zu leisten. Studien in anderen Sprachen haben gezeigt, dass die meisten Kinder mit einer SSES auch Probleme im grammatikalischen Bereich zeigen. Eine Fortführung dieser Studien könnten die Diagnose von SSES wesentlich erleichtern und zeitlich verkürzen, was allen Beteiligten (Kindern, Therapeuten, Eltern) zu Gute kommen würde.

Aus diesem Grunde möchte ich Sie bitten, mit Ihrem Kind an meiner Studie teilzunehmen. Ich habe im letzten Jahr Kinder getestet, die Sprachtherapie benötigen, und möchte ihre Daten nun mit Daten von sprachunauffälligen Kindern vergleichen, die ungefähr im gleichen Sprachentwicklungsalter sind. Dies würde bedeuten, dass Sie Ihr Einverständnis zu einer ausführlichen logopädischen Diagnostik bei ihrem Kind geben. Außer den Tests, die notwendig sind, um den genauen Sprachstatus ihres Kindes zu bestimmen, würde eine 20minütige Spontansprachaufnahme gemacht und später analysiert werden. Für Sie und Ihr Kind bestünde der Vorteil, dass Sie durch diese Untersuchungen einen Überblick über den sprachlichen Entwicklungsstand Ihres Kindes erhalten. Für den Fall, dass bei Ihrem Kind sprachliche Auffälligkeiten in irgendeiner Form im Rahmen der Untersuchung festgestellt würden, erhalten Sie eine Beratung, was am besten zu tun ist. Die Daten der Studie werden zeigen können, ob Auffälligkeiten in der Grammatik (z.B. Artikelgebrauch, Verbendungen) es vermögen, zwischen Kindern mit und ohne SSES zu unterscheiden. Eine genaue Diagnostik und das bessere Verstehen von Sprachstörungen sind ein wichtiger Schritt in der Forschung, um Kindern mit Sprachproblemen besser und gezielter helfen zu können. Mit Ihrer Unterstützung würden Sie einen wichtigen Beitrag leisten.

Selbstverständlich werden alle erhobenen Daten streng anonym behandelt und gespeichert. Der Name Ihres Kindes wird nicht verwendet werden. Außerdem steht Ihnen natürlich jederzeit das Recht zu, von dem Forschungsprojekt zurückzutreten.

Wenn Sie einverstanden sind, an der Studie teilzunehmen und Ihr Kind **keine Ihnen bekannte Hörstörung** hat, bitte ich Sie, untenstehende Einverständniserklärung zu unterschreiben.

Mit freundlichen Grüßen und bestem Dank, J. Girndt, MSc

---

Ich erkläre mich hiermit einverstanden, dass die Daten meiner Tochter / meines Sohn ________________________________________(Name)________________ (Geb.) für Forschungszwecke ohne Namensnennung verwandt werden dürfen. Während der Erhebung der Testdaten darf eine Audioaufnahme der Äußerungen meines Kindes gemacht werden.

_____________________ ___________________________________

Ort, Datum Unterschrift

# Appendix 8 Diagnostic Accuracy of Morphemes and Composites

**Table 41.** Diagnostic Accuracy of Elicited 3rd Person Singular Inflection in Present Tense

| Score | Sensitivity | Specificity | LR+ | LR- | Overall Accuracy |
|---|---|---|---|---|---|
| 0.00 | 0.08 | 1.00 | n.a. | 0.92 | 60% |
| 0.20 | 0.15 | 1.00 | n.a. | 0.85 | 63% |
| 0.25 | 0.23 | 1.00 | n.a. | 0.77 | 67% |
| 0.40 | 0.38 | 1.00 | n.a. | 0.62 | 73% |
| 0.60 | 0.46 | 0.94 | 7.85 | 0.57 | 73% |
| 0.67 | 0.54 | 0.88 | 4.58 | 0.52 | 73% |
| **0.75** | **0.62** | **0.88** | **5.23** | **0.44** | **77%** |
| 0.80 | 0.69 | 0.76 | 2.94 | 0.40 | 73% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 42.** Diagnostic Accuracy of Present Tense Composite from Elicited 2nd and 3rd Person Singular

| Score | Sensitivity | Specificity | LR+ | LR- | Overall Accuracy |
|---|---|---|---|---|---|
| 0.00 | 0.08 | 1.00 | n.a. | 0.92 | 60% |
| 0.14 | 0.15 | 1.00 | n.a. | 0.85 | 63% |
| 0.33 | 0.23 | 1.00 | n.a. | 0.77 | 67% |
| 0.40 | 0.31 | 1.00 | n.a. | 0.69 | 70% |
| 0.50 | 0.38 | 1.00 | n.a. | 0.62 | 73% |
| **0.56** | **0.46** | **1.00** | **n.a.** | **0.54** | **77%** |
| 0.60 | 0.46 | 0.88 | 3.92 | 0.61 | 70% |
| 0.67 | 0.46 | 0.82 | 2.62 | 0.65 | 67% |
| 0.70 | 0.46 | 0.76 | 1.96 | 0.70 | 63% |
| 0.71 | 0.54 | 0.76 | 2.29 | 0.60 | 67% |
| 0.75 | 0.69 | 0.76 | 2.94 | 0.40 | 73% |
| 0.78 | 0.69 | 0.71 | 2.35 | 0.44 | 70% |
| 0.80 | 0.85 | 0.53 | 1.80 | 0.29 | 67% |
| 0.83 | 0.92 | 0.53 | 1.96 | 0.15 | 70% |
| 0.89 | 0.92 | 0.47 | 1.74 | 0.16 | 67% |
| 0.90 | 0.92 | 0.35 | 1.43 | 0.22 | 60% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 43.** Diagnostic Accuracy of Present Tense Composite –*s* from Elicited 2nd and 3rd Person Singular

| Score | Sensitivity | Specificity | LR+ | LR- | Overall Accuracy |
|---|---|---|---|---|---|
| 0.00 | 0.08 | 1.00 | n.a. | 0.92 | 60% |
| 0.14 | 0.15 | 1.00 | n.a. | 0.85 | 63% |
| 0.40 | 0.23 | 1.00 | n.a. | 0.77 | 67% |
| 0.67 | 0.31 | 1.00 | n.a. | 0.69 | 70% |
| 0.71 | 0.38 | 1.00 | n.a. | 0.62 | 73% |
| 0.78 | 0.54 | 1.00 | n.a. | 0.46 | 80% |
| 0.80 | 0.62 | 1.00 | n.a. | 0.38 | 83% |
| 0.83 | 0.69 | 1.00 | n.a. | 0.31 | 87% |
| **0.88** | **0.85** | **0.94** | **14.38** | **0.16** | **90%** |
| 0.90 | 0.92 | 0.76 | 3.92 | 0.10 | 83% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 44.** Diagnostic Accuracy of Present Tense Composite from Spontaneous Production

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.278 | 0.07 | 1.00 | n.a. | 0.93 | 58% |
| 0.500 | 0.14 | 1.00 | n.a. | 0.86 | 61% |
| 0.583 | 0.21 | 1.00 | n.a. | 0.79 | 65% |
| 0.600 | 0.36 | 1.00 | n.a. | 0.64 | 71% |
| 0.667 | 0.36 | 0.94 | 6.07 | 0.68 | 68% |
| 0.722 | 0.43 | 0.94 | 7.29 | 0.61 | 71% |
| 0.750 | 0.43 | 0.88 | 3.64 | 0.65 | 68% |
| 0.786 | 0.57 | 0.88 | 4.86 | 0.49 | 74% |
| 0.813 | 0.64 | 0.88 | 5.46 | 0.40 | 77% |
| 0.846 | 0.71 | 0.88 | 6.07 | 0.32 | 81% |
| **0.850** | **0.79** | **0.88** | **6.68** | **0.24** | **84%** |
| 0.857 | 0.79 | 0.82 | 4.45 | 0.26 | 81% |
| 0.862 | 0.79 | 0.76 | 3.34 | 0.28 | 77% |
| 0.867 | 0.79 | 0.71 | 2.67 | 0.30 | 74% |
| 0.875 | 0.79 | 0.65 | 2.23 | 0.33 | 71% |
| 0.941 | 0.86 | 0.65 | 2.43 | 0.22 | 74% |
| 0.963 | 0.86 | 0.59 | 2.08 | 0.24 | 71% |
| 1.000 | 1.00 | 0.00 | 1.00 | n.a. | 45% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 45.** Diagnostic Accuracy of Present Tense + Copula Composite from Spontaneous Production

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.00 | 0.07 | 1.00 | n.a. | 0.93 | 56% |
| 0.49 | 0.13 | 1.00 | n.a. | 0.87 | 59% |
| 0.58 | 0.20 | 1.00 | n.a. | 0.80 | 63% |
| 0.75 | 0.27 | 0.94 | 4.53 | 0.78 | 63% |
| 0.79 | 0.33 | 0.94 | 5.67 | 0.71 | 66% |
| 0.83 | 0.40 | 0.94 | 6.80 | 0.64 | 69% |
| 0.87 | 0.47 | 0.94 | 7.93 | 0.57 | 72% |
| 0.87 | 0.53 | 0.94 | 9.07 | 0.50 | 75% |
| 0.88 | 0.60 | 0.94 | 10.20 | 0.43 | 78% |
| 0.89 | 0.73 | 0.94 | 12.47 | 0.28 | 84% |
| 0.89 | 0.80 | 0.94 | 13.60 | 0.21 | 88% |
| **0.92** | **0.87** | **0.94** | **14.73** | **0.14** | **91%** |
| 0.93 | 0.93 | 0.88 | 7.93 | 0.08 | 91% |
| 0.93 | 0.93 | 0.82 | 5.29 | 0.08 | 88% |
| 0.93 | 0.93 | 0.76 | 3.97 | 0.09 | 84% |
| 0.95 | 0.93 | 0.71 | 3.17 | 0.09 | 81% |
| 0.96 | 0.93 | 0.65 | 2.64 | 0.10 | 78% |
| 0.98 | 0.93 | 0.59 | 2.27 | 0.11 | 75% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 47% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 46.** Diagnostic Accuracy of Elicited Accusative in Definite Articles

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.00 | 0.15 | 1.00 | n.a. | 0.85 | 63% |
| 0.33 | 0.23 | 1.00 | n.a. | 0.77 | 67% |
| 0.50 | 0.38 | 0.94 | 6.54 | 0.65 | 70% |
| 0.67 | 0.62 | 0.88 | 5.23 | 0.44 | 77% |
| **0.80** | **0.69** | **0.88** | **5.88** | **0.35** | **80%** |
| 0.83 | 0.85 | 0.65 | 2.40 | 0.24 | 73% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 47. Diagnostic Accuracy of Elicited Dative in Definite Articles**

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| **0.00** | **0.77** | **0.82** | **4.36** | **0.28** | **80%** |
| 0.17 | 0.85 | 0.59 | 2.05 | 0.26 | 70% |
| 0.33 | 0.92 | 0.53 | 1.96 | 0.15 | 70% |
| 0.50 | 1.00 | 0.29 | 1.42 | 0.00 | 60% |
| 0.83 | 1.00 | 0.24 | 1.31 | 0.00 | 57% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 48.** Diagnostic Accuracy of Case Composite from Elicited Definite Articles

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.00 | 0.08 | 1.00 | n.a. | 0.92 | 60% |
| 0.28 | 0.23 | 1.00 | n.a. | 0.77 | 67% |
| 0.39 | 0.31 | 0.94 | 5.23 | 0.74 | 67% |
| 0.44 | 0.46 | 0.88 | 3.92 | 0.61 | 70% |
| 0.50 | 0.54 | 0.88 | 4.58 | 0.52 | 73% |
| 0.53 | 0.62 | 0.88 | 5.23 | 0.44 | 77% |
| 0.56 | 0.69 | 0.88 | 5.88 | 0.35 | 80% |
| **0.59** | **0.77** | **0.88** | **6.54** | **0.26** | **83%** |
| 0.61 | 0.85 | 0.82 | 4.79 | 0.19 | 83% |
| 0.67 | 0.92 | 0.71 | 3.14 | 0.11 | 80% |
| 0.72 | 0.92 | 0.53 | 1.96 | 0.15 | 70% |
| 0.78 | 1.00 | 0.47 | 1.89 | 0.00 | 70% |
| 0.82 | 1.00 | 0.41 | 1.70 | 0.00 | 67% |
| 0.83 | 1.00 | 0.29 | 1.42 | 0.00 | 60% |
| 0.94 | 1.00 | 0.18 | 1.21 | 0.00 | 53% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 49.** Diagnostic Accuracy of Spontaneous Production of Nominative Articles

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.000 | 0.06 | 1.00 | n.a. | 0.94 | 53% |
| 0.111 | 0.12 | 1.00 | n.a. | 0.88 | 56% |
| 0.125 | 0.13 | 1.00 | n.a. | 0.88 | 58% |
| 0.167 | 0.24 | 1.00 | n.a. | 0.76 | 62% |
| 0.667 | 0.24 | 0.94 | 4.00 | 0.81 | 59% |
| 0.727 | 0.24 | 0.88 | 2.00 | 0.87 | 56% |
| 0.760 | 0.29 | 0.88 | 2.50 | 0.80 | 59% |
| 0.769 | 0.35 | 0.88 | 3.00 | 0.73 | 62% |
| 0.789 | 0.41 | 0.88 | 3.50 | 0.67 | 65% |
| 0.800 | 0.47 | 0.88 | 4.00 | 0.60 | 68% |
| 0.857 | 0.53 | 0.88 | 4.50 | 0.53 | 71% |
| 0.889 | 0.59 | 0.88 | 5.00 | 0.47 | 74% |
| 0.900 | 0.65 | 0.82 | 3.67 | 0.43 | 74% |
| **0.905** | **0.71** | **0.82** | **4.00** | **0.36** | **76%** |
| 0.909 | 0.76 | 0.76 | 3.25 | 0.31 | 76% |
| 0.933 | 0.76 | 0.71 | 2.60 | 0.33 | 74% |
| 0.957 | 0.82 | 0.71 | 2.80 | 0.25 | 76% |
| 0.966 | 0.82 | 0.65 | 2.33 | 0.27 | 74% |
| 1.000 | 1.00 | 0.00 | 1.00 | n.a. | 50% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 50.** Diagnostic Accuracy of Spontaneous Production of Dative Articles

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.00 | 0.50 | 1.00 | n.a. | 0.50 | 79% |
| 0.25 | 0.58 | 1.00 | n.a. | 0.42 | 82% |
| 0.29 | 0.67 | 1.00 | n.a. | 0.33 | 86% |
| **0.30** | **0.75** | **1.00** | **n.a.** | **0.25** | **89%** |
| 0.33 | 0.75 | 0.81 | 4.00 | 0.31 | 79% |
| 0.42 | 0.75 | 0.75 | 3.00 | 0.33 | 75% |
| 0.43 | 0.75 | 0.69 | 2.40 | 0.36 | 71% |
| 0.50 | 0.92 | 0.63 | 2.44 | 0.13 | 75% |
| 0.60 | 0.92 | 0.56 | 2.10 | 0.15 | 71% |
| 0.67 | 1.00 | 0.38 | 1.60 | 0.00 | 64% |
| 0.70 | 1.00 | 0.31 | 1.45 | 0.00 | 61% |
| 0.75 | 1.00 | 0.19 | 1.23 | 0.00 | 54% |
| 0.78 | 1.00 | 0.13 | 1.14 | 0.00 | 50% |
| 0.80 | 1.00 | 0.06 | 1.07 | 0.00 | 46% |
| 0.92 | 1.00 | 0.00 | 1.00 | n.a. | 43% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 51.** Diagnostic Accuracy of Spontaneous Production of Dative Definite Articles

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.00 | 0.33 | 1.00 | n.a. | 0.67 | 76% |
| 0.25 | 0.44 | 1.00 | n.a. | 0.56 | 80% |
| 0.29 | 0.56 | 1.00 | n.a. | 0.44 | 84% |
| **0.30** | **0.67** | **1.00** | **n.a.** | **0.33** | **88%** |
| 0.33 | 0.67 | 0.94 | 10.67 | 0.36 | 84% |
| 0.40 | 0.67 | 0.88 | 5.33 | 0.38 | 80% |
| 0.42 | 0.67 | 0.81 | 3.56 | 0.41 | 76% |
| 0.50 | 0.89 | 0.63 | 2.37 | 0.18 | 72% |
| 0.67 | 0.89 | 0.50 | 1.78 | 0.22 | 64% |
| 0.75 | 0.89 | 0.44 | 1.58 | 0.25 | 60% |
| 0.78 | 0.89 | 0.38 | 1.42 | 0.30 | 56% |
| 0.80 | 1.00 | 0.25 | 1.33 | 0.00 | 52% |
| 0.86 | 1.00 | 0.19 | 1.23 | 0.00 | 48% |
| 0.88 | 1.00 | 0.13 | 1.14 | 0.00 | 44% |
| 1.00 | 1.00 | 0.00 | 1.00 | n.a. | 36% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 52.** Diagnostic Accuracy of Case Composite from Spontaneous Production of Articles

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.000 | 0.06 | 1.00 | n.a. | 0.94 | 53% |
| 0.111 | 0.12 | 1.00 | n.a. | 0.88 | 56% |
| 0.125 | 0.18 | 1.00 | n.a. | 0.82 | 59% |
| 0.200 | 0.24 | 1.00 | n.a. | 0.76 | 62% |
| 0.606 | 0.29 | 1.00 | n.a. | 0.71 | 65% |
| 0.636 | 0.35 | 1.00 | n.a. | 0.65 | 68% |
| 0.667 | 0.41 | 0.94 | 7.00 | 0.63 | 68% |
| 0.684 | 0.41 | 0.88 | 3.50 | 0.67 | 65% |
| 0.731 | 0.47 | 0.88 | 4.00 | 0.60 | 68% |
| 0.737 | 0.47 | 0.82 | 2.67 | 0.64 | 65% |
| 0.745 | 0.53 | 0.82 | 3.00 | 0.57 | 68% |
| 0.750 | 0.59 | 0.82 | 3.33 | 0.50 | 71% |
| 0.763 | 0.65 | 0.82 | 3.67 | 0.43 | 74% |
| **0.792** | **0.71** | **0.82** | **4.00** | **0.36** | **76%** |
| 0.808 | 0.71 | 0.76 | 3.00 | 0.38 | 74% |
| 0.818 | 0.71 | 0.71 | 2.40 | 0.42 | 71% |
| 0.821 | 0.71 | 0.65 | 2.00 | 0.45 | 68% |
| 0.833 | 0.71 | 0.59 | 1.71 | 0.50 | 65% |
| 0.839 | 0.76 | 0.59 | 1.86 | 0.40 | 68% |
| 0.840 | 0.82 | 0.59 | 2.00 | 0.30 | 71% |
| 0.852 | 0.88 | 0.59 | 2.14 | 0.20 | 74% |
| 0.867 | 0.88 | 0.53 | 1.88 | 0.22 | 71% |
| 0.875 | 0.88 | 0.47 | 1.67 | 0.25 | 68% |
| 0.893 | 0.88 | 0.41 | 1.50 | 0.29 | 65% |
| 0.900 | 0.88 | 0.35 | 1.36 | 0.33 | 62% |
| 0.905 | 0.88 | 0.29 | 1.25 | 0.40 | 59% |
| 0.913 | 0.88 | 0.24 | 1.15 | 0.50 | 56% |
| 0.915 | 0.94 | 0.24 | 1.23 | 0.25 | 59% |
| 0.929 | 0.94 | 0.18 | 1.14 | 0.33 | 56% |
| 0.935 | 1.00 | 0.12 | 1.13 | 0.00 | 56% |
| 0.947 | 1.00 | 0.06 | 1.06 | 0.00 | 53% |
| 0.986 | 1.00 | 0.00 | 1.00 | n.a. | 50% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

**Table 53.** Diagnostic Accuracy of Spontaneous Production of Nominative Pronouns

| ***Score*** | ***Sensitivity*** | ***Specificity*** | ***LR+*** | ***LR-*** | ***Overall Accuracy*** |
|---|---|---|---|---|---|
| 0.000 | 0.13 | 1.00 | n.a. | 0.88 | 58% |
| 0.778 | 0.19 | 1.00 | n.a. | 0.81 | 61% |
| 0.821 | 0.25 | 1.00 | n.a. | 0.75 | 64% |
| 0.857 | 0.31 | 1.00 | n.a. | 0.69 | 67% |
| 0.889 | 0.38 | 1.00 | n.a. | 0.63 | 70% |
| 0.905 | 0.44 | 1.00 | n.a. | 0.56 | 73% |
| 0.913 | 0.50 | 1.00 | n.a. | 0.50 | 76% |
| 0.921 | 0.50 | 0.94 | 8.50 | 0.53 | 73% |
| 0.943 | 0.56 | 0.94 | 9.56 | 0.46 | 76% |
| **0.959** | **0.63** | **0.94** | **10.63** | **0.40** | **79%** |
| 0.962 | 0.63 | 0.88 | 5.31 | 0.43 | 76% |
| 0.966 | 0.63 | 0.82 | 3.54 | 0.46 | 73% |
| 0.969 | 0.63 | 0.76 | 2.66 | 0.49 | 70% |
| 0.972 | 0.63 | 0.71 | 2.13 | 0.53 | 67% |
| 0.974 | 0.63 | 0.65 | 1.77 | 0.58 | 64% |
| 0.974 | 0.63 | 0.59 | 1.52 | 0.64 | 61% |
| 0.979 | 0.69 | 0.59 | 1.67 | 0.53 | 64% |
| 0.985 | 0.69 | 0.53 | 1.46 | 0.59 | 61% |
| 0.986 | 0.75 | 0.53 | 1.59 | 0.47 | 64% |
| 1.000 | 1.00 | 0.00 | 1.00 | n.a. | 48% |

*Note.* Overall accuracy = number of correct classified participants / total number of participants

# Appendix 9 Overview Groups with Outliers Excluded

**Table 54.** Overview over Participant Groups with Outliers in- and excluded from SLI group

| ***Group*** | ***SLI*** | | ***PI*** | ***ND-A*** | ***ND-L*** |
|---|---|---|---|---|---|
| ***n*** | *17* | *13* | *16* | *17* | *16* |
| ***Age*** | 4;00 (0;91) Range: 3;0-5;1 | 4;1 (0;88) Range: 3;0-5;1 | 4;0 (0;88) Range: 3;0-5;1 | 4;0 (0;88) Range: 2;11-5;3 | 3;3 (0;76) Range: 2;0-3;11 |
| ***MLU-M*** | 2.87 $(0.95)_{a,b}$ | 3.32 $(0.50)_{c}$ | 3.71 $(0.77)_{a}$ | 3.93 $(0.46)_{b,c}$ | 3.39 (0.68) |
| ***MLU-W*** | 2.48 $(0.76)_{a,b}$ | 2.83 $(0.46)_{c}$ | 3.25 $(0.59)_{a}$ | 3.35 $(0.36)_{b,c}$ | 2.96 (0.51) |
| ***NDW*** | 82.24 $(33.35)_{a,b}$ | 97.15 $(20.29)_{c}$ | 110.07 $(20.58)_{a}$ | 121.41 $(16.48)_{b,c,d}$ | 101.75 $(23.10)_{d}$ |

*Note.* Standard deviations in brackets. Means in a row sharing subscripts are significantly different when analysed with ANOVA.

SLI = Specific Language Impairment Group; PI = Phonologically Impaired Group;

ND-A = Normally developing, age-matched group;

ND-L = Normally developing, language-matched group, n = Number of participants

MLU-M = Mean length of utterance in morphemes; MLU-W = Mean length of utterance in words

NDW = Number of different words in first 100 complete and intelligible utterances

Significant differences between the SLI group and the ND-A group remain for MLU-M ($p$ = .012), MLU-W ($p$ = .031) and NDW ($p$ = .011) even if the four weakest children of the SLI group were excluded.

# Appendix 10 t-values for SLI relative to ND-A group

**Table 55.** SLI-Group in Elicitation Tasks: Percentiles & *t*-Value Relative to ND-A Group

| *SLI-Group* | | *Morphemes* | | | | | | *Composite Scores* | |
|---|---|---|---|---|---|---|---|---|---|
| **Subject** | **Age** | **3rdPS** | **2ndPS** | **2ndPS *-s/-st*** | **Nom** | **Acc** | **Dative** | **PrT *-s*** | **Case** |
| SLI1 | 3;0 | 0,1 | | | 0,0 | 0,0 | 12,8 | 0,0 | 0,0 |
| SLI2 | 3;0 | 0,0 | | | | | 12,8 | | |
| SLI3 | 3;1 | 0,0 | | | 0,0 | 0,0 | 12,8 | | |
| SLI4 | 3;1 | 0,0 | 0,9 | 0,0 | 0,0 | 0,0 | 12,8 | | |
| SLI5 | 3;1 | 68,6 | 0,9 | 0,0 | 26,8 | 30,4 | 12,8 | 0,5 | 13,5 |
| SLI6 | 3;7 | 0,0 | 0,9 | 0,0 | 1,0 | 6,1 | 23,1 | 0,0 | 5,1 |
| SLI7 | 3;7 | 0,0 | 0,9 | 0,0 | 0,1 | 0,0 | 53,9 | 0,0 | 0,9 |
| SLI8 | 4;0 | 0,1 | 28,5 | 59,2 | 63,2 | 30,4 | 12,8 | 0,0 | 20,5 |
| SLI9 | 4;1 | 0,0 | | | | | | | |
| SLI10 | 4;1 | 68,6 | 78,6 | 59,2 | 63,2 | 71,8 | 37,4 | 69,5 | 51,5 |
| SLI11 | 4;1 | 68,6 | 0,9 | 0,0 | 63,2 | 6,1 | 12,8 | 0,0 | 10,9 |
| SLI12 | 4;3 | 16,9 | | | 63,2 | 0,1 | 12,8 | 0,1 | 5,1 |
| SLI13 | 4;8 | 9,7 | 47,7 | 59,2 | 0,1 | 0,7 | 12,8 | 2,7 | 0,9 |
| SLI14 | 4;10 | 3,3 | 54,6 | 59,2 | 63,2 | 71,8 | 12,8 | 2,7 | 29,4 |
| SLI15 | 5;0 | 68,6 | 28,5 | 0,1 | 63,2 | 0,7 | 12,8 | 6,9 | 8,5 |
| SLI16 | 5;0 | 1,3 | 18,3 | 59,2 | 63,2 | 23,3 | 12,8 | 0,0 | 17,3 |
| SLI17 | 5;1 | 0,0 | 11,0 | 59,2 | 1,0 | 6,1 | 12,8 | 0,0 | 3,0 |

*Note.* Percentiles from modified *t*-test for elicited data relative to age-matched controls

3rdPS=3$^{rd}$ person singular present tense; Nom = nominative articles; Acc = accusative articles; Dat = dative articles; PrT –*s* = present tense composite (with 2ndPS –*s*); Case = case composite from articles

**Table 56.** SLI-Group in Spontaneous Language: Percentiles & *t*-Value Relative to ND-A Group

| *SLI-Group* | | *Morphemes* | | | | | | *Composite Scores* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Subject** | **Age** | **3rd PS** | **Cop** | **Pron Nom** | **Nom** | **Acc** | **Dat** | **3rd+2nd PS** | **PrT** | **PrT+ Cop** | **Case** |
| SLI1 | 3;0 | 8,7 | 0,0 | 0,0 | 0,0 | 2,4 | 0,5 | 17,4 | 0,4 | 0,0 | 0,0 |
| SLI2 | 3;0 | | | 0,0 | 0,0 | | | | | | 0,0 |
| SLI3 | 3;1 | | | 77,0 | 0,0 | | | | | 0,0 | 0,0 |
| SLI4 | 3;1 | 75,5 | 0,0 | 0,0 | 67,9 | | | 75,1 | 55,0 | 0,8 | 14,1 |
| SLI5 | 3;1 | 0,2 | 0,1 | 0,0 | 32,4 | 3,0 | 0,5 | 1,3 | 0,3 | 0,0 | 12,9 |
| SLI6 | 3;7 | 6,7 | 0,0 | 0,0 | 4,8 | 0,0 | 0,5 | 14,7 | 3,8 | 0,2 | 1,0 |
| SLI7 | 3;7 | 0,0 | 0,0 | 0,0 | 7,7 | 86,0 | 31,7 | 0,0 | 0,0 | 0,0 | 25,3 |
| SLI8 | 4;0 | | 59,4 | 16,0 | 67,9 | 44,4 | 7,0 | 0,0 | 0,1 | 0,0 | 10,3 |
| SLI9 | 4;1 | | | | 0,0 | | | | | | 0,0 |
| SLI10 | 4;1 | 5,2 | 33,3 | 77,0 | 67,9 | 86,0 | 63,5 | 23,8 | 15,1 | 29,3 | 78,8 |
| SLI11 | 4;1 | 0,2 | 59,4 | 5,1 | 52,6 | 86,0 | 5,2 | 1,1 | 0,4 | 6,2 | 72,4 |
| SLI12 | 4;3 | 27,1 | 59,4 | 77,0 | 35,5 | 86,0 | 7,9 | 36,8 | 23,1 | 17,8 | 17,1 |
| SLI13 | 4;8 | 75,5 | 0,2 | 0,2 | 9,0 | 30,5 | 0,5 | 75,1 | 74,4 | 8,5 | 3,3 |
| SLI14 | 4;10 | 59,7 | 59,4 | 77,0 | 5,6 | 30,5 | 0,5 | 63,0 | 24,1 | 8,0 | 1,8 |
| SLI15 | 5;0 | 23,0 | 11,7 | 55,5 | 20,0 | 74,2 | 0,5 | 5,0 | 10,4 | 9,3 | 47,9 |
| SLI16 | 5;0 | 75,5 | 59,4 | 0,4 | 34,0 | 24,6 | 0,5 | 75,1 | 74,4 | 75,0 | 42,6 |
| SLI17 | 5;1 | 2,1 | 59,4 | 44,0 | 28,9 | 47,9 | 31,7 | 6,6 | 10,4 | 15,6 | 43,1 |

*Note.* Percentiles from modified *t*-test for spontaneous data relative to age-matched controls

3rdPS = 3rd person singular present tense; Cop = Copula; Pron = Pronoun; Nom = nominative articles; Acc = accusative articles; Dat = dative articles; 3rd+2ndPS = composite 3rd+2nd person singular present tense; PrT = present tense composite; PrT + Cop = PrT including copula; Case = case composite from articles

# Appendix 11 t-values for SLI relative to ND-L group

**Table 57.** SLI-Group in Elicitation Tasks: Percentiles & *t*-Value Relative to ND-L Group

| *SLI-Group* | | *Morphemes* | | | | | | *Composite Scores* | |
|---|---|---|---|---|---|---|---|---|---|
| **Subject** | **Age** | **3rdPS** | **2ndPS** | **2ndPS *-s/-st*** | **Nom** | **Acc** | **Dative** | **PrT *-s*** | **Case** |
| SLI1 | 3;0 | 13,6 | | | 0,0 | 1,5 | 25,5 | 6,5 | 0,0 |
| SLI2 | 3;0 | 1,5 | | | | | 25,5 | | |
| SLI3 | 3;1 | 1,5 | | | 0,0 | 1,5 | 25,5 | | |
| SLI4 | 3;1 | 1,5 | 2,0 | 0,0 | 0,0 | 1,5 | 25,5 | | |
| SLI5 | 3;1 | 76,5 | 2,0 | 0,0 | 15,3 | 62,2 | 25,5 | 49,9 | 33,3 |
| SLI6 | 3;7 | 4,9 | 2,0 | 0,0 | 0,0 | 41,3 | 56,1 | 1,1 | 11,9 |
| SLI7 | 3;7 | 1,5 | 2,0 | 0,0 | 0,0 | 1,5 | 95,5 | 0,4 | 1,4 |
| SLI8 | 4;0 | 13,6 | 32,1 | 60,0 | 69,1 | 62,2 | 25,5 | 33,3 | 48,2 |
| SLI9 | 4;1 | 1,5 | | | | | | | |
| SLI10 | 4;1 | 76,5 | 76,4 | 60,0 | 69,1 | 79,8 | 83,0 | 72,5 | 86,4 |
| SLI11 | 4;1 | 76,5 | 2,0 | 0,0 | 69,1 | 41,3 | 25,5 | 41,9 | 27,0 |
| SLI12 | 4;3 | 54,5 | | | 69,1 | 10,6 | 25,5 | 45,1 | 11,9 |
| SLI13 | 4;8 | 48,3 | 49,2 | 60,0 | 0,0 | 22,9 | 25,5 | 55,9 | 1,4 |
| SLI14 | 4;10 | 38,3 | 55,2 | 60,0 | 69,1 | 79,8 | 25,5 | 55,9 | 63,5 |
| SLI15 | 5;0 | 76,5 | 32,1 | 0,3 | 69,1 | 22,9 | 25,5 | 59,4 | 20,8 |
| SLI16 | 5;0 | 30,8 | 22,5 | 60,0 | 69,1 | 58,1 | 25,5 | 41,9 | 41,9 |
| SLI17 | 5;1 | 6,4 | 15,0 | 60,0 | 0,0 | 41,3 | 25,5 | 27,3 | 6,2 |

*Note.* Percentiles from modified *t*-test for elicited data relative to language-matched controls

3rdPS = 3rd person singular present tense; Nom = nominative articles; Acc = accusative articles; Dat = dative articles; PrT –*s* = present tense composite (with 2ndPS –*s*); Case = case composite from articles

**Table 58.** SLI-Group in Spontaneous Language: Percentiles & *t*-Value Relative to ND-L Group

| *SLI-Group* | | *Morphemes* | | | | | | *Composite Scores* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Subject** | **Age** | **3rd PS** | **Cop** | **Pron Nom** | **Nom** | **Acc** | **Dat** | **3rd+2nd PS** | **PrT** | **PrT + Cop** | **Case** |
| SLI1 | 3;0 | 20,2 | 0,0 | 0,0 | 0,0 | 22,0 | 6,5 | 33.3 | 5.6 | 0.1 | 0.1 |
| SLI2 | 3;0 | | | 0,0 | 0,0 | | | | | | 0.0 |
| SLI3 | 3;1 | | 0,0 | 67,0 | 0,0 | | | | | 0.0 | 0.0 |
| SLI4 | 3;1 | 75,0 | 0,0 | 11,3 | 72,5 | | | 72.1 | 63.1 | 14.0 | 50.2 |
| SLI5 | 3;1 | 1,9 | 4,1 | 18,5 | 46,3 | 24,1 | 6,5 | 9.4 | 4.7 | 2.5 | 48.7 |
| SLI6 | 3;7 | 17,4 | 0,0 | 34,5 | 14,9 | 0,6 | 6,5 | 30.7 | 17.6 | 6.9 | 15.7 |
| SLI7 | 3;7 | 0,0 | 0,1 | 26,3 | 19,8 | 86,5 | 77,9 | 0.9 | 0.1 | 0.0 | 61.9 |
| SLI8 | 4;0 | | 65,3 | 55,3 | 72,5 | 65,3 | 40,7 | 0.5 | 1.9 | 3.2 | 44.7 |
| SLI9 | 4;1 | | | | 0,0 | | | | | | 0.0 |
| SLI10 | 4;1 | 14,8 | 49,7 | 67,0 | 72,5 | 86,5 | 93,3 | 38.7 | 33.9 | 50.0 | 90.2 |
| SLI11 | 4;1 | 1,7 | 65,3 | 50,4 | 61,8 | 86,5 | 34,3 | 8.6 | 5.6 | 29.1 | 87.6 |
| SLI12 | 4;3 | 39,3 | 65,3 | 67,0 | 48,8 | 86,5 | 43,3 | 48.0 | 41.3 | 41.9 | 53.9 |
| SLI13 | 4;8 | 75,0 | 4,9 | 39,0 | 21,9 | 56,9 | 6,5 | 72.1 | 75.0 | 32.4 | 27.8 |
| SLI14 | 4;10 | 63,8 | 65,3 | 67,0 | 16,3 | 56,9 | 6,5 | 64.4 | 42.2 | 31.7 | 21.2 |
| SLI15 | 5;0 | 35,8 | 31,6 | 63,1 | 34,9 | 80,3 | 6,5 | 18.4 | 28.4 | 33.4 | 76.6 |
| SLI16 | 5;0 | 75,0 | 65,3 | 41,4 | 47,6 | 52,7 | | 72.1 | 75.0 | 73.9 | 73.7 |
| SLI17 | 5;1 | 8,6 | 65,3 | 61,2 | 43,3 | 67,2 | 77,9 | 21.1 | 28.4 | 40.0 | 74.0 |

*Note.* Percentiles from modified *t*-test for spontaneous data relative to language-matched controls

3rdPS = 3rd person singular present tense; Cop = Copula; Pron = Pronoun; Nom = nominative articles; Acc = accusative articles; Dat = dative articles; 3rd+2ndPS = composite 3rd+2nd person singular present tense; PrT = present tense composite; PrT + Cop = PrT including copula; Case = case composite from articles