Comprehension of welcher-questions in German-speaking children with and without Developmental Language Disorder

Research Report

Comprehension of *welcher*-questions in German-speaking children with and without Developmental Language Disorder

Anne D. Roesch[1] and Vasiliki Chondrogianni[1]

[1] School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK

Corresponding author:

Dr. phil. Anne Dorothée Roesch

University of Edinburgh
School of Philosophy, Psychology and Language Sciences
Dugald Stewart Building
3 Charles Street
Edinburgh EH8 9AD
Email: ad-roesch@gmx.de

Conflicts of Interest Statement: ADR and VC report no conflicts of interest.

Funding Statement: ADR and VC did not receive any funding for this study.

Abstract

Purpose. This study examined whether monolingual German-speaking preschool children with developmental language disorder (DLD) were facilitated by the presence of case-marking cues in their interpretation of German subject and object *welcher*("which")-questions, as reported for their TLD peers. We also examined whether knowledge of case-marking and/or phonological working memory modulated children's ability to revise early assigned interpretations to ambiguous questions.

Method. 63 monolingual German-speaking children with and without DLD aged between 4;0 and 5;11 years participated in an offline picture selection task targeting the comprehension of *welcher*-questions in German. We manipulated question type (subject, object), case-marking transparency, and case-marking position within the question (sentence-initial/-final).

Results. The TLD children outperformed children with DLD across conditions, and all children performed better on subject than on object *wh*-questions. Transparent and early cues elicited higher accuracy than late-arriving cues. For the DLD children, their working memory capacity explained their inability to revise early assigned interpretations to ambiguous questions, whereas their knowledge of case did not.

Conclusions. The results suggest that disambiguating morphosyntactic cues can only partly facilitate comprehension of German *welcher*-questions in children with DLD, whose poor phonological working memory rather than their knowledge of case-marking mediates performance on these structures.

Introduction

Cross-linguistic research has shown that complex structures such as relative clauses (RCs) and *wh*-questions can be acquired late by typically developing (TLD) children and cause persistent difficulties for children with developmental language disorder (DLD) acquiring English (Deevy & Leonard, 2004; Ebbels & van der Lely, 2001; Lee & Ashmore, 1983; Marinis & van de Lely, 2007; van der Lely & Battell, 2003), Danish (Jensen de Lopez, Sundahl Olsen & Chondrogianni, 2014), Italian (Adani, van der Lely, Forgiarini & Guasti, 2010; Belletti, Friedmann, Brunato & Rizzi, 2012; Guasti, Branchini & Arosio, 2012b), German (Adani *et al*., 2013; Arosio, Yatsushiro, Forgiarini, & Guasti, 2012; Roesch & Chondrogianni, 2015; 2016), Hebrew (Friedmann & Novogrodsky, 2004, 2011), Greek (Stavrakaki, 2001, 2006) and Swedish (Hansson & Nettelbladt, 2006).

In the TLD literature, studies have shown that 4-to 7-year-old TLD children – unlike adults – have great difficulty revising interpretations they assign to sentences with ambiguous early cues, even when sentence-final cues contradict their initially assigned interpretation (Choi & Trueswell, 2010; Omaki, White, Goro, Lidz & Philips, 2014). TLD children's inability to recover from this 'garden-path effect' has been linked to their limited working memory capacity, which is still developing at that age (Trueswell, Sekerina, Hill & Logrip, 1999; Hurewitz, Brown-Schmidt, Thorpe, Gleitman & Trueswell, 2001; Kidd & Bavin, 2002; Weighall, 2008). At the same time, when complex structures carry disambiguating morpho-syntactic or semantic cues such as case-marking (Arosio *et al*., 2012; Roesch & Chondrogianni, 2015, 2016; Schouwenaars *et al*., 2018), number (Adani *et al*. 2010) and gender (Guasti *et al*., 2012b), TLD children's comprehension of

these structures improves. This facilitatory cue effect has also been linked to the cue position in the sentence, with early occurring unambiguous cues facilitating disambiguation (Omaki, White, Goro, Lidz & Philips, 2014; Roesch & Chondrogianni, 2015, 2016; Schouwenaars *et al*., 2018).

To date, our understanding regarding how children with DLD comprehend temporarily ambiguous questions and whether they can revise their early interpretation remains incomplete. In this study, we addressed this gap by investigating the comprehension of subject and object *welcher*("which")-questions in German-speaking TLD and DLD preschool children. By focusing on four- to five-year-old children with DLD, we examined whether cue facilitation occurs at this age, as it has been reported for their TLD age-matched peers, and whether knowledge of case-marking or working memory modulate recovery from the 'garden-path' effect in both the language impaired and the TLD populations.

***Wh*-questions in German**

German is a morphologically rich language that marks gender, number, and case on determiners and nouns. In declarative sentences, SVO is considered the most canonical and frequent word order (Haider, 2010), as in (1a). German is also a V2 language in which the verb always occupies the second position in declarative main clauses and agrees with the subject regardless of its position in the sentence (Grewendorf, Hamm & Sternefeld, 1987; Grewendorf, 2002). Further, German allows object topicalization (OVS), as in (1b), although it is generally less frequent.

1a. Der Affe jagt den Hasen.

The$_{NOM}$ monkey$_{NOM}$ chases the$_{ACC}$ rabbit$_{ACC}$

'The monkey chases the rabbit.'

1b. Den Hasen jagt der Affe.

The$_{ACC}$rabbit$_{ACC}$ chases the$_{NOM}$ monkey$_{NOM}$

'The monkey chases the rabbit.'

In German, thematic roles are expressed overtly via case marking on the determiner and/or the noun (Jeuk, 2008; Köpcke, 2003), as in (1a & b). Subjects carry nominative case, whereas direct objects carry accusative. German also has three genders (i.e., masculine, feminine, and neuter) marked on the determiner and sometimes also on the noun. There are masculine nouns, which do not carry overt case marking (e.g. as in *der$_{NOM}$ Hund$_{Ø}$–den$_{ACC}$ Hund$_{Ø}$,* 'the dog'), whereas on other masculine nouns case marking is obligatory (e.g. *der$_{NOM}$ Bär – den$_{ACC}$ Bären$_{ACC}$*, 'the bear'). In feminine and neuter nouns, syncretism between the nominative and the accusative case in each gender in the singular makes syntactic roles indistinguishable from one another, e.g., Feminine singular: *die$_{NOM/ACC}$ Maus$_{Ø}$,* 'the mouse'; Neuter singular: *das $_{NOM/ACC}$ Pferd$_{Ø}$*, 'the horse'. Given that German is a V2 language, the correct interpretation of the different syntactic roles within a sentence is contingent upon the ability to understand case marking, as (1b) demonstrates.

Wh-questions are complex structures that involve displacement of constituents (Chomsky & Lasnik, 1995). In a subject *wh*-question (as in 2a), the subject *welcher Elefant* remains in a

sentence initial position even after movement (Haider, 2010) causing no change of the canonical word order (SVO) of the sentence. In an object *wh*-question (as in 2b), however, the object leaves its original sentence-final position (marked by a silent trace [t_i]) to move to the sentence-initial landing site. In this respect, constituent movement in object *wh*-questions creates a non-canonical word order (OSV) compared to subject *wh*-questions (SVO). To felicitously interpret *wh*-questions in German, one needs to be able to make use of case-marking, as word order may not offer a facilitatory cue.

2a.	*Welch**er***	*Elefant$_i$ [t_i]*	*wäscht*	*de**n***	*Bär**en**?*
	Which$_{NOM}$	elephant$_{Ø}$	washes	the$_{ACC}$	bear$_{ACC}$?
	‘Which elephant is washing the bear?’				

2b.	*Welch**en***	*Elefant**en**$_i$*	*wäscht*	*de**r***	*Bär [t_i]?*
	Which$_{ACC}$	elephant$_{ACC}$	washes	the$_{NOM}$	bear$_{Ø}$?
	‘Which elephant is the bear washing?’				

Acquisition of German *wh*-questions in TD children and children with DLD

Wh-questions have been reported to emerge in German-speaking TD children at the age of 1;7 years and involve structures with the infinitival form of the verb (Clahsen, Kursawe & Penke, 1995; Clahsen, 1982 in Schrey-Dern, 2006). Once children reach a mean length of utterance (MLU) of >1.75 words, they start producing finite verbs in a V2 position, which resembles adult-

like subject-verb inversion (Clahsen, 1988; Clahsen & Penke, 1992; Wexler, 1994). With the acquisition of the verb's finiteness (around 3 years of age), fronted *wh*-elements also appear (Penner, 1994; Tracy, 1994).

The ability to comprehend simple *wh*-questions, such as *wo/ was/ wer (ist) …?* ("where, what, who (is) …?") also emerges in TLD children aged between 3- and 4-years old, whereas it remains challenging for children with DLD of similar age (Penner & Kölliker Funk, 1998; Siegmüller, Herzog & Hermann, 2005; Schulz, Tracy & Wenzel, 2008; Schulz & Wenzel, 2007; Schulz, 2007). Importantly, the number and position of case-marking cues has been shown to modulate children's ability to felicitously comprehend non-referential *wer* ("who")-questions. In a previous study with five-year-old monolingual German-speaking children (Roesch & Chondrogianni, 2015), we manipulated non-referential *wer/wen?* ("who$_{NOM/ACC}$?") subject and object *wh*-questions carrying case-marking cues on the *wh*-element (*wer$_{NOM}$* and *wen$_{ACC}$*), or on both the *wh*-element and the second NP (NP2), e.g. *Wer$_{NOM}$ schiebt den$_{MASC\cdot ACC}$ Löwen $_{MASC\cdot ACC}$/die$_{FEM}$ Ente$_{FEM}$?* ("Who is pushing the lion/the duck?" for subject *wh*-questions), or *Wen$_{ACC}$ schiebt der$_{NOM}$ Löwe$_{NOM}$/die$_{FEM}$ Ente?* ("Who is the lion/ duck pushing?") for object *wh*-questions. For both the DLD and the TLD children, the more cues were available, the higher their accuracy was. Yet, children with DLD did not reach same accuracy rates as their TLD peers. The children with DLD in our study also performed significantly better on subject versus object *wh*-questions, similarly to what has been reported for children with DLD in other languages (De Vincenci, Arduino, Ciccarelli & Job, 1999; Ebbels & van der Lély, 2001; Ervin-Tripp, 1970; Friedmann & Novogrodsky, 2011; Guasti *et al*. 2012a,b; Marinis & van der Lély, 2007; O'Grady,

2005; Stavrakaki, 2001, 2006; Tyack & Ingram, 1977; Van der Lély & Battell, 2003), although their performance is reduced compared to that of their TLD age-matched peers. Cue facilitation in TLD children was also independently found in an eye-tracking study examining *welcher*-questions in 7- to 10-year-old TLD German children, where children of this age showed adult-like performance on these structures when full cues were available (Schouwenaars *et al.,* 2018).

Two questions emerge at this point: (i) what makes complex structures such as *wh*-questions, and especially object *wh*-questions, hard to acquire for children with DLD, and (ii) in what way does the presence of cues facilitate the comprehension of these structures. We turn to these two questions in the following section.

The development of cue comprehension in TLD German children

Various studies have examined how TLD German-speaking children acquire cues in simple and complex sentences. Brandt, Lieven and Tomasello (2016) investigated the use of word order and case-marking in 3- to 6-year-old TLD German children in short SVO and OVS sentences as well as in subject and object RCs. Whilst younger children were reported to rely stronger on word order than on case-marking, older children showed adult-like competence by giving case-marking precedence over word order. Similar results were reported by Dittmar, Abbot-Smith, Lieven and Tomasello (2008) investigating the comprehension of simple sentences manipulated in terms of word order or case-marking cues in 2- to 7-year-old children and by Lindner (2003) examining German TD children aged 2;2 to 9;10 years and their comprehension strategies in subject-object sentences manipulated in terms of animacy, word order, verb-agreement and case-marking. In

sum, with age, TLD children become sensitive to distributed cues (case-marking and verb-agreement) rather than focussing on local cues (i.e., animacy and 1st NP), and can also be primed to focus on distributed cues even in low frequency structures such as relative clauses, resulting in increased performance on such structures (Brandt, Nitschke & Kidd, 2017).

Why are object wh-questions difficult to acquire for children with DLD?

Difficulties with the comprehension of object *wh*-questions have been attributed to a number of factors related to (i) to the limited working memory capacity of children with DLD and their problems with processing and integrating morphosyntactic information (Deevy & Leonard, 2006; Leonard, 2014), to (ii) the (non-)canonical ordering of constituents in object questions (Friedmann & Novogrodsky, 2004, 2007), and to (iii) problems with performing syntactic operations such as movement in children with DLD (van der Lely & Battel, 2003).

According to the *Linguistic Processing Account* (Deevy & Leonard, 2004), children with DLD have problems interpreting object *wh*-questions due to their limited working memory capacity. Deevy & Leonard (2006) showed that English-speaking with DLD displayed poorer performance on object compared to subject *wh*-questions. This is because in object *wh*-questions the moved constituent needs to be interpreted at its trace position and not where it surfaces linearly in the sentence. As such, children need to hold more information in working memory before they can interpret the trace. This is not the case for subject *wh*-questions where the linear word order matches on to the correct sentence interpretation. Given that children with DLD have limited working memory capacity, they will have difficulty retaining in memory the information prior to

the interpretation of the object at its trace position, and will, hence, perform more poorly on this condition. Thus, according to the *Linguistic Processing Account*, there is a link between limited working memory capacity and poor performance on complex structures, which has been independently established in children with DLD (Adams & Gathercole, 2000; Delage & Frauenfelder, 2019; Gillam et al., 2017).

A widely reported finding in the TLD and DLD literature is that children perform better on subject than on object questions, what is known as the subject-object asymmetry. This asymmetry has been explained in different ways. According to the *Canonicity Hypothesis* (Friedmann and Novogrodsky, 2004, 2007), children are more successful at interpreting or comprehending subject *wh*-questions because the syntactic order of constituents (NP1: subject, NP2: object) matches the linear order of thematic roles (NP1: agent, NP2: patient). If children assign a linear canonical interpretation to the sentence, they will be successful at interpreting subject questions because of the match between thematic role and question type. In the case of object questions, however, linear word order and thematic roles do not match, as the object argument has been displaced from its original position to a topicalized landing site at the beginning of the sentence. If children interpret an object *wh*-question in a linear manner, they will be misled to assign a subject interpretation to the first constituent, leading to low performance on object *wh*-questions. In the Friedman & Novogrodksy (2004) study, Hebrew-speaking DLD children were unable to interpret object *wh*-questions as non-canonical structures, contrary to their TD peers.

According to the *Representational Deficit of Dependent Relations* (RDDR) (van der Lely & Battel, 2003), *wh*-questions are problematic for children with DLD, because they are unable to form syntactic dependencies between displaced constituents and their traces. As a result, children with DLD are expected to show optionality in the comprehension of *wh*-questions, which will manifest itself as chance performance. Following the RDDR, children with DLD should have more problems with object compared with subject questions, because, for this account as well, linear word order may give rise to more target-like interpretations.

Does the position of the cue matter?

Recent studies with TLD children and adults have reported that the position of the cue in the sentence affects comprehension and processing. Choi and Trueswell (2010) tested their *Verbal Processing Constraints Account* by investigated whether Korean-speaking 4-to 5-year-old children and adults were able to revise their initial interpretation of ambiguous sentences, when a sentence-final cue contradicts that initial interpretation. Their results suggested that, unlike adults, who were able to revise their initial erroneous interpretation, children persisted with their initial interpretation. Similar findings were reported by Omaki *et al*. (2014) for Japanese- and English-speaking children, when processing sentences with late-arriving cues. Taken together, these results suggest that sentences with ambiguous cues at the beginning of the sentence give rise to 'garden-path effects' (Frazier, 1987; Frazier & Fodor, 1978) and require revision of initially assigned interpretations. This process is costly for working memory, as the comprehender needs to store linguistic material in memory that needs to be updated to accommodate the new and

correct interpretation. This account then predicts that even TLD children of a young age will have difficulty revising sentences with late arriving cues and that this inability to revise the initially assigned interpretation will be linked to young TLD children's limited working memory capacity. In our study, we test this prediction by examining whether 4- to 5-year-old children are able to revise their interpretation of sentences with late arriving cues and we examine what the contribution of working memory is when performing these revisions. Furthermore, we extend the predictions of the *Verbal Processing Constraints Account* to children with DLD by examining whether they were facilitated or hindered by the presence of early or late disambiguating cues respectively, and whether their performance was influenced by their working memory capacity.

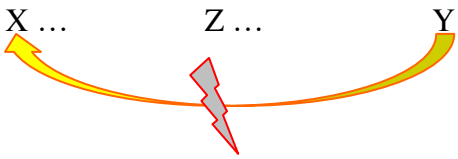
Does the presence of cues facilitate comprehension of complex structures children?

A number of studies have shown that the presence of semantic and morphosyntactic cues can improve TLD children's comprehension of complex structures. This facilitation has been reported when the two NPs in a complex structure carry distinctive morpho-syntactic or semantic cues, e.g., gender or number (Adani *et al.,* 2013; Belletti *et al*., 2012), or where NPs are of a different lexical surface type (Gordon, Hendrick & Johnson, 2001), e.g., lexically descriptive NP versus indexical pronouns or proper names). In the context of *which* object questions (example 2) such as *Which dog is the donkey chasing?*, both NPs are lexical and appear in a non-canonical word order (due to object topicalization and subject-verb inversion in question formation). In (2), the argument Z (here the noun 'the donkey') intervenes between the trace in Y and its landing site in X, creating what is coined in linguistic terms as an 'intervention effect'. Thus, on the level of

interpretation, the two arguments *which*-N ('which dog' = X) and the NP ('the donkey' = Z) compete for selection of the subject role.

Intervener

(2) Which dog is the donkey chasing [$Y_{x\text{-}origin}$] ?



Z intervenes in the relation between X and Y

However, when the intervener carries features that differ from those of X and Y in terms of gender, number or type of NP, the intervention effect can be overridden. In the case of German, the presence of case and gender may render the two nouns in a *wh*-question sufficiently distinct to override intervention effects and lead to an improvement of performance. In our previous studies with bilingual 4- to 5-year-old French-German TLD children (Roesch & Chondrogianni, 2016) and monolingual German 5-year-old children with and without DLD (Roesch & Chondrogianni, 2015), we found that when the two NPs in non-referential *wh*-question were unambiguously marked for case, German-speaking five-year-old children could felicitously interpret them. In the present study, we expanded the focus of our inquiry by investigating how 4- to 5-year-old children with and without DLD comprehend referential *wh*-questions with different types of case-marking cues, and how age, knowledge of case marking, and working memory modulates their performance.

Present study

In the present study, we examined whether German-speaking preschool children with DLD are sensitive to the type and position of morphosyntactic (case) cues when interpreting *welcher*-questions, as previously reported for TLD children (Roesch & Chondrogianni, 2015, 2016), and

how their comprehension is modulated by age and working memory. Our research questions were the following:

1. Does question type, presence and position of cues affect accuracy of performance when 4-to 5-year-old children with or without DLD comprehend *welcher*-questions?

2. Does question type and cue type affect error patterns during comprehension of subject and object *welcher*-questions in children with or without DLD?

3. How does age, knowledge of case marking and phonological working memory affect comprehension of temporarily ambiguous object *welcher*-questions in TLD and DLD children?

The different accounts make both converging and diverging predictions as to how subject and object questions are expected to be acquired by German-speaking children with DLD.

With respect to the subject-object asymmetry, all theoretical accounts predict that children with DLD will perform better on subject than on object questions. However, the source of the performance differs. For the *Canonicity account* (Friedmann & Novogrodsky, 2004; 2007), this asymmetry is because in subject *wh*-questions, linear word order matches thematic role assignment and thus leads to the right interpretation of the sentence, despite both structures involving movement. However, in object *wh*-questions, there is a mismatch between linear word order and thematic role. Therefore, if children with DLD use the linear word order to assign thematic roles in object *wh*-questions, they are expected to arrive to the reverse interpretation and

hence, mainly make reversal errors in terms of error patterns.

For the *Linguistic Processing Theory* (Deevy & Leonard, 2004; Leonard, 2014), German-speaking children's with DLD problems with object *welcher*-question may arise from their limited working memory capacity. Specifically, the limited working memory capacity of children with DLD will be overstretched when comprehending object *wh*-question, as they need to retain verbal information in working memory before they can interpret the object in its original position; children with DLD will have to memorise a great load of unanalysed verbal input and will possibly adopt a guessing strategy when confronted with object *wh*-questions. For the *RDDR* (van Der Lely, 2005), children with DLD are expected to show chance performance on the comprehension of object *wh*-questions, as movement operations are optional in this population. This will contrast with the performance of the TD children who are expected to perform above chance, as movement is not optional syntactic operation for them.

Finally, predictions regarding the effect of the presence or position of disambiguating (case) cues on accuracy are addressed from a psycholinguistic perspective, by the *Verbal Processing Constraint* Account (Choi & Trueswell, 2010; Omaki *et al.,* 2014) and in linguistic terms by the *Intervention Hypothesis* (Friedmann & Novogrodsky, 2011).

Following the *Verbal Processing Constraints Account* (Choi & Trueswell, 2010; Omaki *et al.,* 2014), children will initially assign a sentence interpretation following linear word order and will have difficulties revising their premature interpretation of the clause when a later arriving cue contradicts it. Therefore, *welcher*-questions carrying cues initially, that is on the *welcher*-element (i.e. in case of the 'double cues' and the '*wh*-cue' conditions in our experiment, see

below) will elicit higher accuracy rates than *wh*-questions carrying cues on the 2nd NP only (i.e. in case of the 'NP-cue' condition). More specifically, for *welcher*-questions carrying only sentence-final cues, children are more likely to interpret these as subject *wh*-questions only, because these correspond the prototypical, canonical and linear SVO word order, and have difficulty revising their initial interpretation. This will be also reflected in children's error patterns, which will mostly consist of reversal errors for object NP-cue *wh*-questions.

The *Intervention Hypothesis* (Friedmann & Novogrodsky, 2011) predicts that the distinct morphological marking on articles and nouns in the *welcher*-phrase and NP2 may lead to the overriding of intervention effects and facilitate children's comprehension of *wh*-questions (Adani *et al*., 2013; Belletti, *et al*., 2012, Gordon *et al*., 2001; for relative clauses).

Method

Participants

Sixty-three German-speaking children with and without DLD aged between 4;0 and 5;11 years participated in the study. There were 16 four-year-old children with DLD (mean age: 55.06 months; range: 49-59; SD: 3.13) and 17 typically developing (TLD) age-matched controls (mean age: 51.7 months; range: 48-58; SD: 3.63). In the 5-year-old group, there were 15 children with DLD (mean age: 66.07 months; range: 61-71 in months; SD: 3.43) and 15 TLD age-matched peers (mean age: 64.1 months; range: 61-69; SD: 2.63). The TLD children were recruited from nurseries in Wiesbaden (Mid-West Germany) and in Essen (Northern Germany). According to the parents of the TLD children, they had no history of DLD, mental or psychological disorders. The children

with DLD were recruited from speech and language therapists in Wiesbaden (Mid-West Germany). Apart from the usual speech and language therapist reports confirming the children's impairment status, we conducted a range of assessments assessing the children's verbal, non-verbal and auditory memory abilities. Therefore, the inclusion criteria for the children with DLD consisted of a clinical diagnosis of language impairment and performance of at least one standard deviation below the mean in one or more language assessments. More specifically, we used three subtests from the '*Linguistische Sprachstandserhebung – Deutsch als Zweitsprache*' (LiSeDaZ; Schulz & Tracy, 2011) which were the following: (i) the comprehension of morpho-syntax further involving the comprehension of inflected verbs (Verbbedeutung), *wh*-questions (Verständnis von W-Fragen) and negation (Negation), and (ii) the production of case-marking (elicited production task). Note, that within the subtest 'production of case-marking' the accusative and the dative case were tested, and although a composite score could have been computed, we kept these scores apart for a more detailed analysis. To examine the 'phonological loop' component, which stores acoustic and verbal information in Baddeley's (2003) working memory model, we used two simple span tasks from the SETK 3-5 ('*Sprachentwicklungstest für drei- bis fünfjährige Kinder*'; Grimm, Aktas & Frevert, 2010) (i) the word span task involving monosyllabic word strings (i.e., 2-6 words per string and 10-word strings in total), and (ii) the nonword repetition task. Both tasks have been shown to correlate with language abilities, although the nonword repetition task has been shown to be a better predictor for performance of complex syntactic structures (Gillam et al., 2017). Simple span tasks have also been shown to be able to capture developmental differences in working memory in children as young as the ones tested here (e.g., Vugs et al.,

2017). To measure non-verbal abilities, we used the '*Wechsler Preschool and Primary Scale of Intelligence'* (WPPSI-3; Wechsler, 2011). Exclusion criteria for all groups were performance below one standard deviation on the non-verbal intelligence task (WPPSI-3, Wechsler, 2011), as well as a history of hearing impairment, frank neurological impairment, psycho-emotional disturbance, or diagnosis of autism. All TLD and DLD children were within the normal range for non-verbal intelligence (WPPSI-3, Wechsler, 2011). Table 1 shows the raw scores on the five subtests of the LiSeDaZ (Schulz & Tracy, 2011) and the SETK 3-5 (Grimm, Aktas & Frevert, 2010) for the children with DLD and their TLD controls. All numbers indicate raw scores, as the two groups were closely matched on age.

INSERT TABLE 1 HERE

An independent samples t-test indicated that the 4-year-old preschool children with DLD were outperformed by their TLD age-matched peers in the production of accusative case-marking. Significance was approached in the comprehension of verbs, while no significant differences were found in the comprehension of negation or the verbal span task. Significant differences in performance between the 5-year-old TLD children and age-matched children with DLD were found across tasks. The only exception, where there was no significant difference, was in the comprehension of verbs.

Experimental Tasks

The task administered in the present study was the picture-pointing task used in our previous studies with TD children (Roesch & Chondrogianni, 2015; 2016). In this task, children were

shown a picture panel with three animals performing the same action on each other. The two animals on the right and the left side of the template were of the same kind, while the middle animal was different (*see* Figure 1).

INSERT FIGURE 1 HERE

To examine whether children comprehension of *welcher*-questions is affected by morphosyntactic factors, we manipulated question type (i.e., subject or object *wh*-questions), as well as the type and position of case-marking cues to examine their effect in comprehension. We created three cue conditions: (i) case-marking carried by the *wh*-element and NP2 (henceforth referred to as 'double cues'), (ii) case-marking carried only by the *wh*-element (henceforth referred to as *early*-cue), and (iii) case-marking only occurring on NP2 (henceforth: *late*-cue). All conditions are shown in Examples 3 – 5 below.

	Subject *wh*-questions	**Object *wh*-questions**
(3) *Double cues*	**Welcher** Esel schiebt **den** Hund?	**Welchen** Esel schiebt **der** Hund?
	$\text{Which}_{\text{NOM}}$ donkey pushes the_{ACC} dog?	$\text{Which}_{\text{ACC}}$ donkey pushes the_{NOM} dog?
	Which donkey is pushing the dog?	*Which donkey is the dog pushing?*
(4) *Early (wh-)*	**Welcher** Hase schiebt die Gans?	**Welchen** Hasen schiebt die Gans?
cue	$\text{Which}_{\text{NOM}}$ rabbit pushes $\text{the}_{\text{-Ø}}$ goose?	$\text{Which}_{\text{ACC}}$ rabbit pushes $\text{the}_{\text{-Ø}}$ goose?
	Which rabbit is pushing the goose?	*Which rabbit is the goose pushing?*

(5) *Late (NP-) cue*	Welches Pferd schiebt **den Elefanten?**	Welches Pferd schiebt **der** Elefant?
	Which-ø horse pushes the$_{ACC}$ elephant$_{ACC}$?	Which-ø horse pushes the$_{NOM}$ elephant?
	Which horse is pushing the elephant?	*Which horse is the elephant pushing?*

The tasks were introduced to the children using two practice items. There were six target *wh*-questions and four distractor items per condition, amounting to a total of 60 *wh*-questions (excluding the two practice items). Within the experimental properties, target answers always involved one of the side animals, whereas distractor items always targeted the middle animal. Therefore, the distractor items were created to spread the children's attention to all depicted animals as possible target responses.

There were four different response types in coding: correct, reverse, distractor or no answer. According to this, the child's answer was coded as *Correct* if showing the correct item, as *Reverse* if showing the *reverse* item (i.e., nominative instead of the accusative or vice versa), or as *Distractor*, if showing the middle animal instead of the actual target animal on the left or the right of the picture. These codes were intended to allow an error pattern analysis.

Procedure

Testing consisted of three separate 30-minute sessions per child. Children were tested in a quiet room in their schools or homes. The baseline tasks were administered in the first session, while in the second and third session the comprehension. Prior to testing, a background questionnaire was completed by the parents to gather background information such as early language

development of their children, as well as whether their child was diagnosed with DLD and had received intervention, and whether family members showed any symptoms of language impairments. Participants were excluded, if they failed to complete one of the parts of data collection, i.e., if the parental questionnaire was not on hand prior to testing, or if the baseline or experimental tasks were incomplete). This resulted in a total of fourteen TLD children and six children with DLD being excluded from the study.

Results

Statistical analysis used lme4 statistical package in R (R Core Team, 2019). To investigate accuracy, we ran generalized mixed-effects logistic regression because of the binary nature of the data (1=correct, 0=incorrect). Predictors were entered into the model in a stepwise fashion and predictors that did not improve the model fit were excluded from the final model. Model comparisons were ran using likelihood ratios until the optimal model was identified. Where possible, we included the maximal random effect structure of the model (Baayen, Davidson, & Bates, 2008). Error types were investigated using multinomial logistic regression for each condition separately. Pairwise comparisons between levels of individual factors were carried out by changing the reference level. Data visualisation was carried out using ggplot2.

Accuracy on *wh*-questions

Figure 2 presents the accuracy rate on referential subject and object *wh*-questions with 'double cues', 'early cues' or 'late cues' in the 4- and 5-year-old pre-schoolers with DLD and their TLD controls.

INSERT FIGURE 2 HERE

To investigate whether the TLD children and the children with DLD differed on *wh*-questions in terms of question type, number and position of case-marking cue, we ran a mixed effects logistic regression with Question Type (**Subject**, Object) and Cue (e.g., 'double cues', **'early cue'** and 'late cue'), Group (**DLD**, TLD) and Age (4- and 5-year olds) as fixed effects (levels in bold are the reference levels). The optimal model is presented in Table 2.

INSERT TABLE 2

To better understand the interaction between Age, Group and Cue, we ran mixed effects logistic regressions for each group separately. For the TLD group, there was a significant interaction between Cues and Age (5-year-olds: *double cues*: $E. = -1.96$, $S.E. = 0.52$, $z = -3.79$, $p < .001$; 5-year-olds: *late cue*: $E. = -1.3$, $S.E. = 0.37$, $z = -3.5$, $p < .001$; reference level set to early cues). Pairwise comparisons with Bonferroni correction revealed that for the four-year-old children, the *late cue* condition had lower accuracy than the *double* and *early cue* condition, which also had lower accuracy than the double cue condition (all significant differences at $p < .001$ level). For the five-year-old TLD children, there was no significant difference between the *early* and the

double cues (p = .34) and both conditions had higher accuracy than the *late cue* condition (p < .001 in both cases).

For the DLD children, *double cues* had higher accuracy than *early cues* ($E.$ = 0.43, $S.E.$ = 0.18, z = 2.42, p < .05) and *late cues* had lower accuracy than both (compared to *early cues*: $E.$ = -0.83, SE = 0.18, z =-4.68, p < .001). Object *wh*-questions also had lower accuracy than subject questions ($E.$ = -0.58, $S.E.$ = 0.15, z = -3.9, p < .001). No effects of age or other effects or interactions were observed.

Finally, we binomial testing with a 95% CI to examine whether children performed at chance level (p-value set to .33 as they needed to choose one out of three possible referents/animals). On the late cue conditions, all groups of children performed either at chance (4- and 5-year-old DLD children: p = .83; 4-year-old TD children: p =.41; 5-year-old children on subject questions: p = .78), or below chance (5-year-old children on subject questions: p < .05). On all other conditions, both groups of children performed above chance level.

Error analysis

Figure 3 displays error types for subject and object *wh*-questions within all cue conditions committed by the TLD and DLD children. Possible errors were *reversal* errors, when pointing to the opposite of the correct response, and *distractor* errors, when pointing to the middle animal in the picture. For ease of presentation the results are collapsed for age.

INSERT FIGURE 3 HERE

A multinomial logistic regression revealed various main effects and interactions. To unpack these interactions, we computed *paired* and *independent samples t*-tests with Bonferroni correction to account for the number of multiple comparisons. Overall, the DLD children produced more distractors errors across conditions compared to their TLD peers ($p < .01$). In the subject and object NP-cue conditions, children with DLD committed as many reversal as distractor errors for subject and object *wh*-questions ($p = .94$). In the double-cues and the *wh*-cue conditions, all children with DLD opted more often for the reverse than the distractor error in both the subject and the object *wh*-questions ($p < .001$ in both cases). The TLD children made significantly more reversal than distractor errors across cue conditions and question type ($p < .001$), and more errors on the NP-cue compared to the *wh*-cue condition ($p < .001$). The fewest errors were found on the double cue conditions ($p < .001$).

Predictors of accuracy on ambiguous object questions

As a last step in our analysis, we examined the factors that predicted performance on the comprehension of *wh*-questions in the TLD and DLD children. We focused on the NP-cue object condition, as this was the more challenging for both groups. The predictors that we focused on were knowledge of accusative case and the phonological loop as measured by the digit span task and the nonword repetition task. Results (Table 3) revealed that working memory were the most significant predictor for both children with DLD and their TLD counterparts. Nonword repetition explained more of the variance in the TLD (conditional $R^2 = .43$) compared to the DLD (conditional $R^2 = .18$) children (conditional R^2 contains variance from both the fixed and the

random effects. For the TLD children, the model with knowledge of case was also significant (*E.* = 0.14, *S.E.* = 0.06, $z = 2.33$, $p = .02$), but the model with nonword repetition had a better fit to the data (chi-square test: $p < .0001$) and was kept as the optimal model (the model with knowledge of case as a factor did not improve the optimal model).

INSERT TABLE 3 ABOUT HERE

Discussion

This study examined the comprehension of subject and object *welcher*-questions in monolingual German-speaking 4- to 5-year-old children with and without DLD. Our research questions concerned: (1) whether question type affects accuracy and error patterns in the comprehension of *welcher*-questions, (2) whether the presence and position of case-marking cues influences accuracy rates and error patterns, and (iii) how age, knowledge of case-marking and/or working memory influences children's ability to revise early assigned interpretations to ambiguous questions. We review the present findings and their significance for the different theories in the following sections.

Subject-object asymmetry

A robust finding of the present study was that both the typically developing and the language-impaired children exhibited better performance on subject compared to object *wh*-questions, in line with previous studies in the literature (e.g., Friedmann and Novogrodsky, 2004, 2007). For

the TLD 5-year-old children, performance on subject questions reached highest accuracy rates irrespective of cues, and it was above chance on both subject and object *welcher*-questions. An exception were *welcher*-questions with NP-cues, where all groups showed chance or below chance performance. When children did not choose the target referent, they either opted for distractor errors, in the case of subject questions or reversal errors, as was the case of object questions. Interestingly, in the NP-cue condition, all children adopted more reversal errors than distractor errors regardless of question type.

Which theoretical accounts of DLD can explain these results? According to the *Canonicity Hypothesis* (Friedmann & Novogrodsky, 2004, 2006, 2011) children's with DLD low performance on object *wh*-questions is due to the mismatch between thematic role assignment and linear word order. If children parse object *wh*-questions linearly to assign thematic roles to arguments, they are expected to be led down the 'garden-path' and adopt more reversal errors than distractor errors in this condition, while also exhibiting chance performance. Similar predictions regarding optional performance for the children with DLD on object questions but due to different theoretical assumptions (optionality in movement operations) are also put forward by the *RDDR* (van der Lely, 2005). The results of the present study partly confirm these predictions, as all children (regardless of impairment status) performed above chance on all object *wh*-questions carrying either 'double cues' or '*wh*-cues' and performed at chance only on object 'NP-cue' *wh*-questions. This suggests that question type but also cue ambiguity and position influenced children's performance. We discuss this in the following section.

Cue ambiguity and position affected performance

In our study, we found that performance on *welcher*-questions in both the TLD and the DLD children was influenced by the position of the cue. That is, performance was higher when the *wh*-question contained unambiguous, facilitatory case-marking cues on both the *wh*-phrase and the second NP, compared to when the cue was only in a sentence-initial position ('*early*-cue' condition). Performance dropped significantly in both groups when the disambiguating cue was placed in a sentence-final position ('*late-cue'* condition). These results suggest that both groups regardless of impairment status were sensitive to the presence and nature of case-marking cues. However, only the TLD children were able to make full use of the unambiguous cues and reach highest accuracy rates performance on the '*double-cue'* condition, whereas both groups performed at chance or below chance on the '*late-cue'* condition. Taken together, these results point to children's sensitivity to the nature and position of cue regardless of impairment status, as it has been suggested in other psycholinguistic studies with children and adults (Choi and Trueswell, 2010; Omaki et al., 2014). It seems that early occurring cues facilitate the comprehension of complex *wh*-questions, and that this felicitous interpretation is reinforced even further when both the *wh*-phrase and NP2 contain disambiguating cues, at least in this offline comprehension task, which required children to listen to the entire sentence before pointing. In contrast, when the disambiguating cues occur in a sentence-final position, it is difficult for children of this age, regardless of impairment status, to revise their initially assigned interpretation and reach the target sentence interpretation. For object questions in the '*NP-cue'* condition, children's interpretation seems to involve assigning the agent role to the first NP they encounter,

as indicated by their predominant error pattern of reversal errors. That means, that after listening to the entire sentence and encountering NP2 in nominative case in this condition, they opted for the agent in the sentence. The presence of an accusative marked NP2 in the subject '*NP-cue'* condition gave rise to a different error types, with children opting for both reversal but also more distractor errors. These error patterns indicate that children were sensitive to the case morphology on NP2 but were not able to felicitously integrate this information to reach the target sentence interpretation despite the offline nature of the task. Importantly, this pattern of results was observed in the four-year-old TLD children and the four- and five-year-old children with DLD, with only the five-year-old TLD children performed above chance on the subject '*NP-cue'* condition.

The importance of having clear unambiguous cues was also evidenced in the higher accuracy children exhibited in the '*double-cue'* condition had in relation to the *early cue* (*wh*-cue) condition, although in both conditions, the disambiguating cue appeared in a sentence-initial position. From a linguistic perspective, this finding shows that when the two NPs are unambiguously marked for case as in the 'double cues' condition, '*intervention effects'* can be overridden (Friedmann & Novogrodsky, 2011). From a more psychological point of view, this difference may also be related to the offline nature of our task, where children had time to reflect at the end of the sentence and were able to confirm their initial interpretation of the question against the sentence-final case-marking cues as well. The ability to reflect on the sentence presupposes that the sentence can be retained in working memory in the first place, a process not as straightforward for the children with DLD, as discussed in the next section.

Age, case-marking and working memory as predictors for children's performance on ambiguous object questions.

One of the main questions in our study was whether children's comprehension of *welcher*-questions would improve as a function of age. Results showed that this was indeed the case for the TLD children but not for the children with DLD. The performance of the 5-year-old TLD children improved considerably compared to their younger TLD peers, managing to reach (almost) highest performance in the 'double cue' and '*early*-cue' conditions, and high performance on the subject '*late-cue'* condition. This shows that with age, five-year-old TLD children can revise their initial interpretation, especially when the word order of the structure is canonical (subject *wh*-questions), and these results are in line with Roesch and Chondrogianni (2015, 2016). In contrast, accuracy or error rates in the DLD children showed no differences between the 4- and the 5-year-old groups in their response strategies. Instead, we found a persistent 'guessing strategy' (chance level at 33%), which suggests a stagnating language development for complex structures even by the age of five years. It should be noted that even the five-year-old TLD children cannot revise their initial interpretation in the context of object questions, contrary to what has been reported for seven- to ten-year-olds in the Schouwenaars et al., (2018) study, who had almost 90% accuracy on a similar condition.

Given that in order to felicitously understand *welcher*-questions in German, one needs to be able to know case-marking and retain verbal information in their working memory, what is the contribution of case-marking knowledge and working memory to the comprehension of questions

at this age? To answer this question, we focused on the ambiguous '*late-cue'* condition because this was the most challenging one, as it provides late occurring cues and requires verbal material to be retained in working memory until the end of the sentence is reached.

The logistic regression analysis with case-marking and working memory as predictors revealed that knowledge of case-marking played little role for the children with DLD, and that working memory as measured by performance on the nonword repetition task was an important predictor for both groups of children. For both groups, the contribution of the nonword repetition task on performance on sentences with late arriving cues confirmed that for both groups the comprehension of these sentences is taxing, and that children, regardless of impairment status, need to make great use of working memory resources to comprehend these sentences. This finding is consistent with the *Linguistic Processing Theory* (Deevy & Leonard, 2004; Leonard, 2014) that postulates that DLD children's limited working memory capacity limits their ability to perform sentence interpretation efficiently. This finding is also consistent with the *Verbal Processing Constraints Account* (Choi & Trueswell, 2010) that predicts that TLD children of a preschool age will fail to revise initially assigned interpretations to sentences with late arriving cues, and that this is due to the TLD children's limited working memory capacity. It is worth noting that the nonword repetition task explained more of the variance in the comprehension skills of the TLD (approximately 43)% compared to the DLD (approximately 18%) children. This is consistent with studies showing that simple WM tasks can be predictive of young TLD children's comprehension of complex syntactic structures, whereas in children with DLD, the ability to process complex sentences is more related to the verbal central executive component of WM (Frizelle & Fletcher,

2015; but see also Delage & Frauenfelder, 2019 for a link between simple span tasks and expressive skills in older children with DLD). As such, the lack of a digit span effect may be due to an insufficient range of performance on this task in both groups, and once individual variability is taken into account with our mixed-effect logistic regression analysis. Finally, for the TLD children, the nonword repetition task explained the variance in performance on the object late-cues condition more strongly than children's performance on case-marking. This suggests that TLD children's knowledge of case marking helps them to interpret ambiguous sentences, but that the felicitous interpretation of such structures is primarily contingent upon children's working memory abilities, at least at this age.

Conclusions

This study investigated how German-speaking children with DLD and TLD comprehend *welcher* subject and object questions with (un)ambiguous cues at different positions in the sentence. Results confirmed the subject-object asymmetry reported in previous studies, but also showed that this is modulated by cue position and ambiguity, with questions with case-marking cues on both the *wh*-phrase and the second NP outperforming questions, with a sentence initial cue; questions with sentence-final cues were the hardest ones to comprehend and elicited chance performance. For the TLD children, performance improved in the 5-year-old TLD children; however, this was not the case for the DLD children, who are challenged by these structures at both the ages of four and five years. Finally, working memory accounted for the inability of children with DLD to revise their initial interpretation of ambiguous questions.

References

Adams, A. M., & Gathercole, S. E. (2000). Limitations in working memory: implications for language development. *International Journal of Language and Communication Disorders*, 35(1), 95–116. https://doi.org/10.1080/136828200247278

Adani, F., van der Lely, H. K. J., Forgiarini, M. & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: a comprehension study with Italian children. *Lingua 120*(9), 2418–66. https://doi.org/10.1016/j.lingua.2010.03.018

Adani, F., Forgiarini, M., Guasti, M.T. & van der Lély, H.K. J. (2013). Number dissimilarities facilitate the comprehension of relative clauses in children with (Grammatical) Specific Language Impairment. *Journal of Child Language 41*(4), 1-31. https://doi.org/10.1017/S0305000913000184

Arosio, F., Yatsushiro, K., Forgiarini, M., & Guasti, M.T. (2012). Morphological Information and Memory Resources in Children's Processing of Relative Clauses in German. *Language learning and development 8*(4), 340-364. http://dx.doi.org/10.1080/15475441.2011.634691

Baayen, R.H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. DOI:10.1016/J.JML.2007.12.005

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4.

Belletti, A., Friedmann, N., Brunato, D. & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew

and Italian, *Lingua*, *122*(10), 1053-1069. https://doi:10.1016/j.lingua.2012.02.007

Brandt, S., Nitschke, S. & Kidd, E. (2017). Priming the Comprehension of German Object Relative Clauses, Language Learning and Development, 13:3, 241-261, DOI: 10.1080/15475441.2016.1235500

Brandt, S., Lieven, E. & Tomasello, M. (2016). German Children's Use of Word Order and Case Marking to Interpret Simple and Complex Sentences: Testing Differences Between Constructions and Lexical Items, Language Learning and Development, 12:2, 156-182, DOI: 10.1080/15475441.2015.1052448

Choi, Y. & Trueswell, J.C. (2010). Children's (in)ability to recover from garden-paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology, 106*(1), 41-61.

Chomsky, N. & Lasnik, H. (1995). Principles and Parameters. (pp. 13-128). In N. Chomsky (Eds.), *A Minimalist Programme*. MIT Press.

Clahsen, H. (1982). *Spracherwerb in der Kindheit*. Narr.

Clahsen, H. (1988). *Normale und gestörte Kindersprache. Linguistische Untersuchungen zum Erwerb von Syntax und Morphologie*. Benjamins.

Clahsen, H., Kursawe, C. & Penke, M. (1995). Introducing CP: Wh-Questions and Subordinate Clauses in German Child Language. *Essex Research Reports in Linguistics 7*, 1-28.

Clahsen, H. & Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-Corpus. In J. Meisel (Eds.), *The Acquisition of Verb Placement: Functional Categories and V2*

Phenomena in Language Acquisition (pp. 181-223). Dordrecht.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German Children's Comprehension of Word Order and Case Marking in Causative Sentences. *Child Development*, 79(4): 1152 – 1167, DOI: 10.1111/j.1467-8624.2008.01181.x

De Vincenzi, M., Arduino, L.S., Ciccarelli, L. & Job, R., (1999). Parsing strategies in children comprehension of interrogative sentences. In S. Bagnara (Ed.), *Proceedings of the European Conference on Cognitive Science* (pp. 301-308). Istituto di Psicologia del CNR.

Deevy, P. & Leonard, L. B. (2004). The comprehension of Wh-questions in children with specific language impairment. *Journal of Speech, Language, and Hearing Research 47*, 802–815. https://doi.org/10.1044/1092-4388(2004/060)

Deevy, P. & Leonard, L.B. (2006). Cognitive and Linguistic Issues in the study of children with specific language impairment. In M. J. Taxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd ed., pp. 1143-1171). Elsevier Academic Press. https://doi.org/10.1016/B978-012369374-7/50031-6

Delage, H., & Frauenfelder, U. (2019). Relationship between working memory and complex syntax in children with Developmental Language Disorder. *Journal of Child Language*, 1-33. https://doi.org/10.1017/S0305000919000722

Ebbels, S. & van der Lely, H. (2001). Meta-syntactic therapy using visual coding for children with severe persistent SLI. *International Journal of Language & Communication Disorders 36*, 345-350. https://doi.org/10.3109/13682820109177909

Ervin-Tripp, S. M. (1970). Discourse agreement: How children answer questions. In R. Hayes

(Ed.), *Cognition and language learning* (pp. 79-107). New York.

Everitt, A., Hannaford, P., & Conti-Ramsden, G. (2013). Markers for persistent specific expressive language delay in 3-4-year-olds. *International Journal of Language and Communication Disorders 48*, 534–553. https://doi.org/10.1111/1460-6984.12028

Frazier, L. & Fodor, J.D. (1978). The sausage machine: A new two-stage parsing model. *Cognition 6*, 291-325. https://doi.org/10.1016/0010-0277(78)90002-1

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Erlbaum.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language 31*, 661–681. https://doi.org/10.1017/s0305000904006269

Friedmann, N. & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language 101*, 50-63. https://doi.org/10.1016/j.bandl.2006.09.006

Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua 121*, 367-382. https://doi.org/10.1016/j.lingua.2010.10.004.

Frizelle, P., & Fletcher, P. (2015). The role of memory in processing relative clauses in children with specific language impairment. *American Journal of Speech-Language Pathology*, 24, 47–59. https://doi.org/10.1044/2014_AJSLP-13-0153

Gillam, R. B., Montgomery, J. W., Gillam, S. L., & Evans, J. L. (2017). Working memory in child

language disorders. In R. G. Schwartz (Ed.), Handbook of child language disorders (2nd ed.) (pp. 213–37). New York: Routledge.

Guasti, M.T., Branchini, C., Arosio, F., & Vernice, M. (2012a). A developmental study of subject and object relatives in Italian. *Revue Roumaine de Linguistique 78*(2), 105-116.

Guasti, M.T., Branchini, C., & Arosio, F. (2012b). Interference in the production of Italian subject and object *wh*-questions. *Applied Psycholinguistics 33*, 185-223. https://doi.org/10.1017/S014271641100034

Grewendorf, G., Hamm, F. & Sternefeld, W. (1987). *Sprachliches Wissen* (3rd ed.). Suhrkamp Verlag.

Grewendorf, G. (2002). *Minimalistische Syntax*. Francke Verlag.

Grimm, H., Aktas, M. & Frevert, S. (2010). SETK 3-5 ('*Sprachentwicklungstest für drei- bis fünfjährige Kinder*'. Hogrefe.

Haider, H. (2010). *The Syntax of German*. *Cambridge Syntax Guides*. (1st ed.) Cambridge University Press.

Haiden, M., Ferré, S. Prévost, P., Scheidnes, M. & Tuller, L. (2009). Production and Comprehension of *wh*-questions in the acquisition of French: Comparing L2 children and children with SLI. *Conference Deutsche Gesellschaft für Sprachwissenschaft* (*DGfS),* Osnabrück.

Hansson, K., & Nettelbladt, U. (2006). *Wh*-questions in Swedish children with SLI. *International Journal of Speech-Language Pathology 8*, 376–383. https://doi.org/10.1080/14417040600880722

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2001). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research 29*(6), 597–626.

Jensen de López, K., Sundashi, O.L. & Chondrogianni, V. (2014). Annoying Danish relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of Child Language 41*, 51-83. https://doi.org/10.1017/S0305000912000517

Jeuk, S. (2008). "Der Katze sieht den Vogel" Aspekte des Genuserwerbs im Grundschulalter. In B. Ahrenholz (Ed.), *Zweitspracherwerb: Diagnosen, Verläufe, Voraussetzungen* (pp. 135-150). Filibach.

Kidd, E. & Bavin, E. L. (2002). English-speaking children's comprehension of relative clauses: evidence for general-cognitive and language-specific constraints on development. *Journal of Psycholinguistic Research 31*, 599–617. https://doi.org/10.1023/A:1021265021141

Köpcke, K.-M. (2003). Grammatische Komplexität und die Beherrschung der Kasusmorphologie durch Grundschulkinder. *Didaktik Deutsch 14*(3), 55-68.

Lindner, K. (2003). The development of sentence interpretation strategies in monolingual German-learning children with and without language impairment. *Special issue on specific language impairment. Linguistics* .41: 213 –254.

Lee, R.F. & Ashmore, L.L. (1983). Receptive and expressive *wh*-question performance by language-delayed children. *Journal of Communication and Disorders 16*(2), 99-109. https://doi.org/10.1016/0021-9924(83)90041-2

Leonard, L. (2014). *Children with specific language impairment*. (2nd ed.) MIT Press.

Marinis, T., & van der Lely, H. K. J. (2007). On-line processing of *wh*-questions in children with G-SLI and typically developing children. *International Journal of Language & Communication Disorders 42*, 557-582. https://doi.org/10.1080/13682820601058190

Omaki, A., Davidson-White, I., Goro, T., Lidz, J., Phillips, C. (2014). No fear of commitment: children's incremental interpretation in English and Japanese *wh*-questions. *Language Learning and Development 10*, 206-233.

O'Grady, W. (2005). *How Children learn Language*. Cambridge University Press.

Penner, Z. (1994). Asking Questions without CPs? On the Acquisition of Wh Questions in Bernese Swiss German and Standard German. In T. Hoekstra & B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 177 – 214). John Benjamins.

Penner, Z. & Kölliker Funk, M. (1998). *Therapie und Diagnose von Grammatikerwerbsstörungen*. HPS–Reihe 14.

Roesch, A.D. & Chondrogianni, V. (2015). The use of cues in the comprehension of *wh*-questions by German-speaking children with or without SLI. In C. Hamann & E. Ruigendijk (Eds.), *Proceedings of GALA 2013* (pp. 379-402). Cambridge Scholars. http://www.cascadilla.com/bucld38toc.html

Roesch, A.D. & Chrondrogianni, V. (2016). "Which mouse kissed the frog?": Comprehension of German *Wh*-Questions in Simultaneous and Early Sequential Bilingual Children, *Journal of Child Language 43*(3), 635-661. https://doi.org/10.1017/S0305000916000015

Schouwenaars, A., Hendriks, P. & Ruigendijk, E. (2018). German children's processing of morphosyntactic cues in *wh*-questions. *Applied Psycholinguistics 39*(6), 1279-1318.

https://doi.org/ 10.1017/S0142716418000334

Schrey-Dern, D. (2006). *Sprachentwicklungsstörungen: Logopädische Diagnostik und Therapieplanung*. (1st ed.) Thieme.

Schulz, P. (2007). Erstspracherwerb Deutsch: Sprachliche Fähigkeiten von eins bis zehn. In U. Graf & E. Moser Opitz (Eds.), *Diagnostik am Schulanfang* (Band 4, pp. 67–86). Schneider Verlag.

Schulz, P. & Wenzel, R. (2007). Comprehension and production of *wh*-questions in early L2 learners of German: Like or unlike SLI? *Vortrag anl. International Symposium on Bilingualism* (isb6) Hamburg.

Schulz, P., Tracy, R. & Wenzel, R. (2008). Linguistische Sprachstandserhebung – Deutsch als Zweitsprache (LiSe-Daz): Theoretische Grundlagen und erste Ergebnisse. In B. Ahrenholz (Ed.), *Zweitspracherwerb: Diagnosen, Verläufe, Voraussetzungen* (1st ed., pp. 17-41). Filibach.

Schulz, P. & Tracy, R. (2011). LiSe-DaZ ('*Linguistische Sprachstandserhebung – Deutsch als Zweitsprache*'). Hogrefe.

Siegmüller, J., Herzog, C. & Herrman, H. (2005). Syntaktische und lexikalische Aspekte beim Verstehen von Informationsfragen. *LOGOS Interdisziplinär, 13*(1), 29–35.

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in Specifically Language Impaired and Normally Developing Greek children. *Brain and Language 77*, 419-431. https://doi.org/10.1006/brln.2000.2412

Stavrakaki, S. (2006). Developmental perspectives on Specific Language Impairment: Evidence

from the production of *wh*-questions by Greek SLI children over time. *Advances in Speech-Language Pathology 8*(4), 384-396. https://doi.org/10.1080/14417040600880714

Tracy, R. (1994). Raising questions: Formal and functional aspects of the acquisition of *wh*-questions in German. In R. Tracy & E. Lattey (Eds.), *How Tolerant is Universal Grammar: Essays on Language Learnability and Language Variation* (pp. 1-34). Niemeyer.

Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition 73*, 89–134.

Tyack, D., Ingram, D., (1977). Children's production and comprehension of questions. *Journal of Child Language 4*, 211-224. https://doi.org/10.1017/S0305000900001616

Unsworth, S. (2005). *Child L2, Adult L2, Child L1: Differences and Similarities. A study on the acquisition of direct object scrambling in Dutch*. LOT Dissertation Series.

Van der Lely, H. K. J. & Battell, J. (2003). *Wh*-movement in children with grammatical SLI: a test of the RDDR Hypothesis. *Language 79*, 153-181.

Vugs, B., Hendriks, M., Cuperus, J., Knoors, H., Verhoeven, L. (2017). Developmental Associations Between Working Memory and Language in Children With Specific Language Impairment: A Longitudinal Study. *J Speech Lang Hear Res. 60*(11), 3284-3294. https://doi:10.1044/2017_JSLHR-L-17-0042

Wechsler, D. (2011). *Wechsler Preschool and Primary Scale of Intelligence™ – Third Edition (WPPSI-3)*. Pearson Assessment.

Weighall, A.R. (2008). The Kindergarten Path Effect Revisited: Children's Use of Context in Processing Structural Ambiguities. *Journal of Experimental Child Psychology 99*, 75-95.

https://doi.org/10.1016/j.jecp.2007.10.004

Wexler, K. (1994). Optional Infinitives, Head Movement, and the Economy Of Derivation in Child Language. In D. Lightfoot & N. Hornstein (Eds.), *Verb Movement* (pp. 305-350). Cambridge University Press.

Table 1. Raw scores of the LiSeDaZ (Schulz & Tracy, 2011) and the SETK 3-5 (Grimm, Aktas & Frevert, 2010) for the TLD children and the children with SLI.

		4-yrs TLD (N=17)			**4-yrs DLD (N= 16)**			***t*-tests**
		Mean	**Range**	**SD**	**Mean**	**Range**	**SD**	
Verbs (max. score of 12 points)		9.88	7 – 12	1.87	11.13	6 – 12	2.03	t (31) = 1.83 ; p = .07
Negation (max. score of 12 points)		10.06	6 – 12	1.71	8.12	0 -12	4.21	t (31) = 1.75 ; p = .09
Case-marking[1] (max. score of 5 points for ACC and 4 points for DAT)	ACC	2.29	2 – 5	2.99	.25	0 – 2	.58	t (31) = 6.56 ; p < .001
	DAT	.8	0 – 3	1.3	0	0 – 0	0	t (31) = 4.19 ; p < .001
Word Span task (max. score of 6 points)		4	2 – 5	2.15	3	0 – 4	.87	t (31) = .5 ; p = .62
Nonword repetition (max. score of 18 points)		10.65	9 – 17	5.78	3.5	0 – 12	3.3	t (31) = 4.34 ; p < .001
		5-yrs TLD (N= 15)			**5-yrs DLD (N= 15)**			***t*-tests**
		Mean	**Range**	**SD**	**Mean**	**Range**	**SD**	
Verbs (max. score of 12 points)		11.2	7 - 12	1.69	10.53	8 - 12	1.81	t (28) = 1.04 ; p = .29
Negation (max. score of 12 points)		11.53	10 - 12	.74	.8	3 - 12	2.2	t (28) = 5.88 ; p < .005
Case-marking[2] (max. score of 5 points for ACC and 4 points for DAT)	ACC	3.37	1 - 5	1.78	1.2	0 - 5	1.7	t (28) = 13.36 ; p < .001
	DAT	1.2	1 - 3	.2	0	0 - 0	0	t (28) = 5.94 ; p < .001
Word Span task (max. score of 6 points)		5.27	4 - 5	.96	3.27	3 - 4	.6	t (28)= 8.87 ; p <.001
Nonword repetition (max. score of 18 points)		15.9	13 –18	4.5	4	2 – 9	2.2	t (28) = 17.05 ; p < .001

Table 2. Optimal models for the DLD and TLD children on the comprehension of subject and object *wh*-questions per cue and age group.

Predictors	**Est.**	**S.E.**	***Z***	***p***
(Intercept)	0.34	0.18	1.89	0.06
TLD	0.70	0.26	2.68	0.01
object *wh*-questions	-0.56	0.18	-3.04	0.00
late cues	-0.77	0.22	-3.46	0.00
double cues	0.56	0.22	2.51	0.01
5-year-olds	0.19	0.25	0.76	0.45
TLD: object *wh*-questions	-0.09	0.28	-0.33	0.74
TLD: late cues	-1.05	0.31	-3.39	0.00
TLD : double cues	2.00	0.45	4.48	0.00
TLD : 5-year-olds	1.26	0.43	2.96	0.00
object *wh*-questions: 5-year-olds	-0.05	0.25	-0.19	0.85
late cues: 5-year-olds	-0.12	0.31	-0.38	0.70
double cues: 5-year-olds	-0.27	0.30	-0.88	0.38
TLD : object *wh*-questions : 5-year-olds	-0.59	0.42	-1.40	0.16
TLD : late cue: 5-year-olds	-1.17	0.48	-2.44	0.01
TLD : double cues: 5-year-olds	-1.68	0.60	-2.80	0.01

Table 3. Predictors of performance on the *object late cue* condition for the DLD and the TLD children.

DLD children					**TLD children**				
Predictors	***Odds Ratios***	***CI***	***p***	***df***	***Predictors***	***Odds Ratios***	***CI***	***P***	***Df***
(Intercept)	1.16	0.87 – 1.55	0.321	1880.00	(Intercept)	2.12	1.24-3.63	0.006	1892
Nonword repetition	1.13	1.02 – 1.25	**0.018**	1880.00	Nonword repetition	1.22	1.09-1.36	0.001	1892
Random Effects					**Random Effects**				
σ^2	3.29				σ^2	3.29			
$\tau_{00\ Participant}$	0.00				$\tau_{00\ Participant}$	0.02			
$\tau_{00\ Item}$	0.69				$\tau_{00\ Item}$	2.441			
ICC	0.17				ICC	0.43			
$N_{Participant}$	63				$N_{Participant}$	63			
N_{Item}	36				N_{Item}	36			
Observations	1884				Observations	1869			
Conditional R^2	0.178				Conditional R^2	0.431			

Note. Conditional R^2 reflects the variance from both the fixed and the random effects

Figures:

Figure 1. Example item depicting two animals of the same kind on right and left and another of a different kind in the middle.

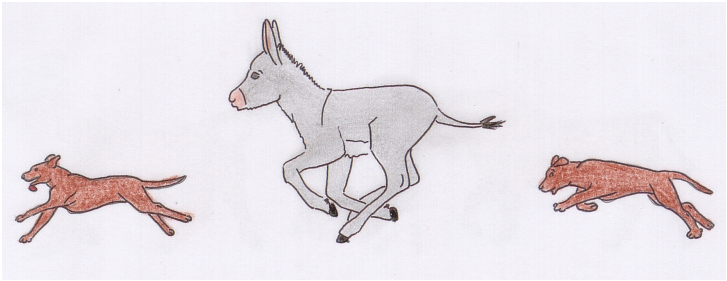


Figure 2. Proportion accuracy on the subject and object *wh*-questions by cue type for the DLD and the TLD four- and five-year-old children.

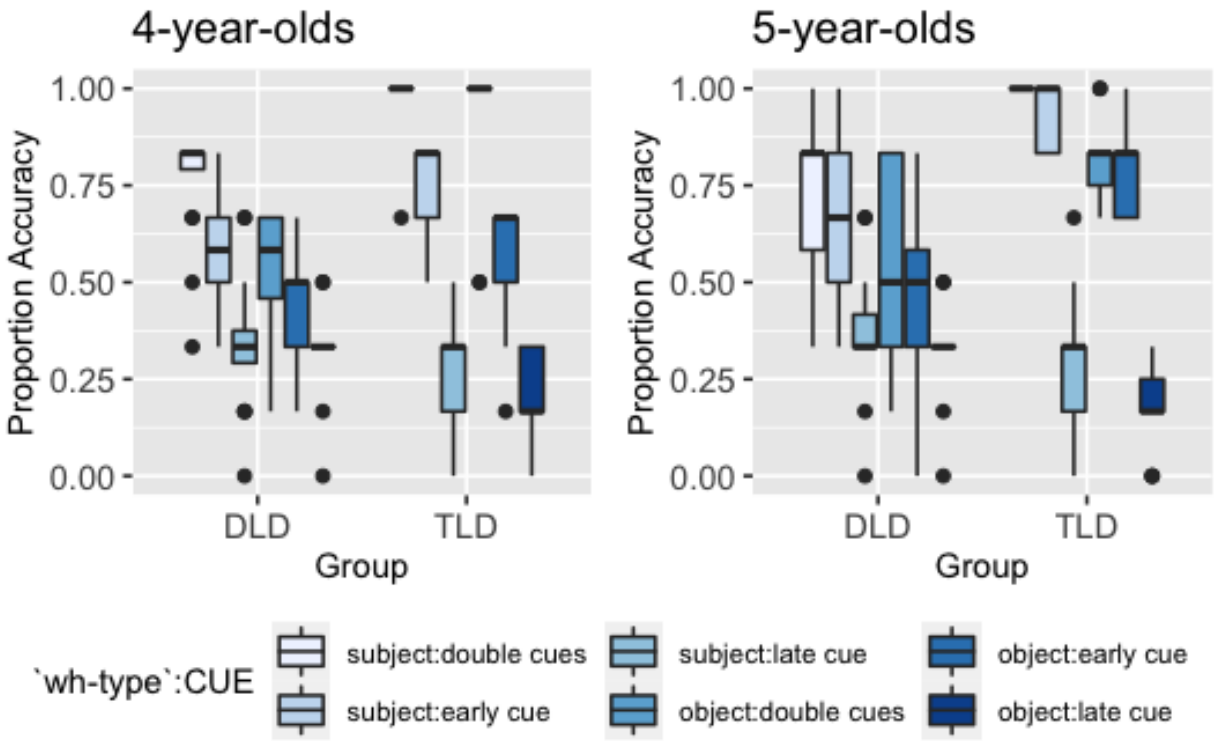


Figure 3. Proportion of responses (Correct, Distractor, Reverse) on subject and object *wh*-questions by cue type for the DLD and the TLD children.