Published in final edited form as:

Lang Acquis. 2015 ; 22(4): 355–383. doi:10.1080/10489223.2015.1028629.

Interpretation of Anaphoric Dependencies in Russian-speaking Children with and without Developmental Language Disorder

Natalia Rakhlin,

Wayne State University, Communication Sciences and Disorders, Detroit, United States

Sergey A. Kornilov,

Yale University, New Haven, United States. Moscow State University, Moscow, Russian Federation

Dr. Jodi Reich, and

Temple University, Philadelphia, United States

Dr. Elena L. Grigorenko

Yale University, Child Study Center, Department of Psychology, Department of Epidemiology & Public Health, New Haven, United States

Natalia Rakhlin: natalia.rakhlin@wayne.edu; Sergey A. Kornilov: sergey.kornilov@yale.edu; Jodi Reich: jodi.reich@temple.edu; Elena L. Grigorenko: elena.grigorenko@yale.edu

Abstract

We examined anaphora resolution in children with and without Developmental Language Disorder (DLD) to clarify whether 1) DLD is best understood as missing knowledge of certain linguistic operations/elements or as unreliable performance and 2) if comprehension of sentences with anaphoric expressions as objects and exceptionally case marked (ECM) subjects supports a particular theoretical account of anaphora. Fifty-four native-Russian-speaking children (age $M = 7;6$, $SD = 1;9$) were tested on a picture selection task. Children with DLD ($n=18$) underperformed overall, but displayed similar patterns to the typically developing (TD) group with respect to the extra difficulty of the ECM relative to the transitive and ECM pronouns relative to all other conditions. However, whereas pronouns were more difficult than reflexives for the TD children, this effect was not significant for the DLD group, whose reduced accuracy on reflexives washed out the effect of pronouns in that group. These results are consistent with performance-level vulnerability in DLD, arguably related to weaknesses in lexical processing and with the Reflexivity framework of Binding phenomena.

1. Introduction

The goal of this study was to examine intrasentential reference assignment to pronouns and reflexives by Russian-speaking children with and without Developmental Language Disorder (DLD). DLD, also referred to in research literature as Specific Language Impairment (SLI)¹, is an impairment of language acquisition defined as “persistent difficulties in the acquisition and use of language across modalities ... due to deficits in

Correspondence to: Natalia Rakhlin, natalia.rakhlin@wayne.edu.

comprehension or production, ... [when] “the difficulties are not attributable to hearing or other sensory impairment, motor dysfunction, or another medical or neurological condition and are not better explained by intellectual disability or global developmental delay”, ... [diagnosed] “based on language abilities below age expectations in one or more language domains” (American Psychiatric Association, 2013, p. 42). It has been found that children with DLD, in addition to the widely reported deficits in the production of grammatical morphology (Bedore & Leonard, 2001; Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999; Leonard & Bortolini, 1998; Leonard & Eyer, 1996; Penke, 2009), exhibit problems with the comprehension and production of complex structures, particularly structures with phrasal movement, such as wh-questions (Friedmann & Novogrodsky, 2011; van der Lely, Jones, & Marshall, 2011), relative clauses (Friedmann & Novogrodsky, 2004, 2007; Stavrakaki, 2001), and passives (Friedmann & Novogrodsky, 2007; Marshall, Marinis, & van der Lely, 2007). In addition, there is some evidence that it is not only the interpretive dependencies between moved phrases and their traces that present difficulty for this group of children, but also anaphoric dependencies between anaphoric expressions and their antecedents (Montgomery & Evans, 2009; van der Lely & Stollwerck, 1997). In the present study, we investigated this phenomenon by comparing the pattern, in which Russian-speaking children with and without DLD establish anaphoric relations between pronouns/reflexives as objects of a matrix verb or as exceptionally case marked (ECM) subjects of small clauses. The study had two goals: 1) to help understand the core deficit(s) underlying syntactic difficulties of children with DLD by illuminating whether the deficit is best understood as missing or incomplete linguistic knowledge of certain linguistic operations or abstract elements, or rather as unreliable performance, due to reduced processing resources or deficits in surface morphology; 2) to examine whether the patterns of accuracy in comprehension of sentences with pronouns and reflexives in children with and without DLD can provide evidence in favor of (or against) a particular theoretical account of anaphora.

With respect to the issue of the core deficit(s) in DLD, one view (henceforth, the syntactic view) attributes the weakness in production and comprehension of complex syntactic structures to a selective deficit in children’s syntactic system, such as a deficit in computing/representing complex syntactic dependencies (van der Lely, 2005) or a difference in the featural composition/specification of certain functional heads (Clahsen, Bartke, & Gollner, 1997; Jakubowicz & Nash, 2001; Mastropavlou & Tsimpli, 2011). Another view (henceforth, the processing view) posits processing limitations as the core causal deficit, such as a deficit in phonological short-term memory or phonological processing (Gathercole & Baddeley, 1990; Joanisse & Seidenberg, 1998; Montgomery & Evans, 2009). A third approach explains grammatical difficulties of children with DLD as a consequence of their weakness in extracting patterns and making generalizations with respect to surface morphology (henceforth, the morphological learning account), possibly linked to their deficit in the declarative and procedural memory systems (Conti-Ramsden & Jones, 1997; Lum, Gelgic, & Conti-Ramsden, 2010; Ullman & Pierpont, 2005). Under one version of this

¹We chose not to use the term SLI because, even though it is widely used in research, there is no consensus among language scientists on the validity of categorically distinguishing language impairment in children with IQ above and below a certain arbitrary cutoff. For this reason, SLI is not included as a specifier of a language disorder in the DSM-5 (American Psychiatric Association, 2013), following the recommendations of ASHA (<http://www.asha.org/slp/dsm-5/>). However, all of the inclusionary and exclusionary criteria typically applied for the category SLI are applicable to the DLD group in our sample.

hypothesis (Conti-Ramsden & Jones, 1997), difficulties with grammatical morphology in children with DLD are related to their difficulties with early word learning, such that their lexicons are relatively small compared to age-peers and therefore, do not provide enough raw data (do not reach a “critical mass”) for relevant generalizations to be made. At the same time, under this view, children with DLD may require a higher number of tokens (i.e., larger mental dictionaries) in order to reach a “critical verbal mass”, compared to typically developing (TD) children. If lexical learning is a core difficulty, the consequences may include underspecified, “fuzzy”, or unstable lexical representations of lexical items, particularly function words and morphological exponents and a reduced capacity for morphological processing. This “competence-level” deficit would be associated with inconsistency in children’s performance with respect of the same element or category, a hallmark of DLD (Holm, Crosbie, & Dodd, 2007). This would also interfere with sentence comprehension, particularly, in languages with rich morphological systems and in structures, in which interpretive heuristics cannot be relied on (i.e., interpretation cannot rely solely on lexical knowledge of word meanings and world knowledge), and children must make use of structural cues for morphologically driven word category identification and argument interpretation.

Computing referential dependencies between anaphoric expressions and antecedents involves multiple linguistic domains, including 1) principles of binding, assumed to be universal and, in their classical conceptualization, involving various structural notions, such as c-command and locality, 2) the lexicon, where anaphoric elements are identified as either pronominal or reflexive and matched with their respective declensional paradigms; 3) a procedure that assigns a semantic interpretation to the syntactic structure; and 4) pragmatics, which contextually licenses the “bound” (i.e., co-referring with a local c-commanding antecedent) or “free” (i.e., co-referring with an antecedent outside of the local domain) interpretation of anaphoric expressions outside of syntax (Reuland, 2003). Examining which of these components presents difficulty for children with DLD has the potential to shed light on the linguistic domain(s), in which children with DLD experience disproportionate difficulty (i.e., syntactic, morphological, lexical semantic or pragmatic). Furthermore, as has been previously argued in the literature on acquired language disorders, the patterns of performance by individuals with a disorder can inform linguistic theory by showing which theoretical account of the linguistic phenomenon in question is best compatible with the pattern of breakdown in a language disorder (breakdown-compatibility assumption; Grodzinsky, 1986). A theory that posits a structure of a given linguistic domain compatible with the patterns of impairment should be chosen over a theory incompatible with it. Thus, a system that presupposes certain componential sub-domains should be preferred to a uniform system for the entire domain if the observed breakdown pattern suggests selective deficits along the lines of the subdomains in the former system. This approach has been taken in cross-linguistic comprehension studies of pronouns as ECM subjects of small clauses compared to those as matrix clause objects in typical language development and acquired language disorders (Baauw & Cuetos, 2003; Baauw, Escobar, & Philip, 1997; Edwards & Varlokosta, 2007; Philip & Coopmans, 1996; Varlokosta, 2002).

These studies evaluated different versions of the Binding theory, namely the “classic” Binding (Chomsky, 1981) and the Reflexivity frameworks (Reinhart, 1983), which make

different predictions for language development and breakdown (for a review see Hamann, 2011). Our study followed this approach comparing the pattern of performance exhibited by children with DLD and their TD age peers on a task measuring comprehension of sentences with pronouns and reflexives in matrix clauses and ECM constructions. In the remainder of the article, we will review the core findings with regards to interpretations of anaphoric dependencies in TD children and adult patients with acquired agrammatism, as well as studies of this phenomenon in children with DLD, briefly outline the relevant notions from the classical Binding and the Reflexivity frameworks, and present the results of our experiment. We will argue that our results are most compatible with the Reflexivity framework and suggest that the difficulties children with DLD exhibited with respect to the interpretation of anaphoric dependencies in our study can be best explained by their vulnerability with respect to morphologically driven processing.

1.1. Anaphoric dependencies in typically developing children

Intrasentential anaphoric dependencies are well studied in TD children in a number of languages. The core observation concerns the contrast between children's interpretive accuracy with reflexives and pronouns. In particular, it has been well established that children perform with a high level of accuracy on sentences with reflexive objects, as in (1a), but not with the same types of sentences with pronominal objects, as in (1b). Thus, young children typically correctly reject sentences analogous to (1a) in contexts where Peter loves someone other than himself with the accuracy level of $\approx 80\% - 90\%$. On the other hand, children under the age of 6 in experimental settings erroneously accept the sentence in (1b) as true in contexts in which Peter loves himself $\approx 50\%$ of the time (Hamann, 2011). The findings with Russian-speaking children largely replicate this pattern. Thus, Avrutin and Wexler (1992) found that Russian-speaking children between the ages 4 and 7 incorrectly accepted a local antecedent for a pronoun 52% of the time on a Truth-Value Judgment task when the sentence had a referential NP as the subject; however, their performance was nearly error-free with reflexives (97% correct).

1a Peter_i loves himself_{i/*j}.

1b Peter_i loves him_{*i/j}.

The observation that preschool children allow pronouns to co-refer with a local c-commanding antecedent was confirmed in a variety of languages; cf., however, some divergent findings in languages with clitic pronouns, such as Romance (McKee, 1992) or Greek (Varlokosta, 2002). This was initially taken as evidence that young children lack the knowledge of the syntactic principle regulating the establishment of the referential dependency between pronouns and their antecedents, namely, Binding Principle B (Chomsky, 1981); hence the original term used to describe this phenomenon as “Delay of Principle B”. Principle B requires pronouns and the NP, with which they are allowed to co-refer, not to be structurally “too close” to each other. The structural notion of the locality domain (i.e., what level of structure constitutes “too close”) had been the focus of numerous investigations in adult syntax and proposed to be a subject of parametric variation (Manzini & Wexler, 1987)².

Subsequent acquisition findings, however, suggested that the delay of Principle B hypothesis might be wrong. In particular, studies found that while young children have difficulty ruling out inappropriate reference assignment for pronouns with potential referential antecedents (as demonstrated earlier), they nevertheless perform at a high level of accuracy with pronouns that have a quantificational antecedent, as those in (2), correctly rejecting the bound variable interpretation (Chien & Wexler, 1990).³ This observation that young children consistently reject the bound variable interpretation for sentences with pronouns with universally quantified subjects suggests that they do not allow pronouns to be locally bound, as the errors with referential antecedents may suggest, and sheds doubt on the conclusion that young children lack syntactic knowledge regulating referential relations between pronouns and antecedents.

2 Every cat_i scratched him_{j/*i}.

This observation led to a reevaluation of the earlier acquisition findings (Grodzinsky & Reinhart, 1993) based on the theoretical work by Reinhart (1983), who established a distinction between two mechanisms of reference assignment: variable binding and co-reference; see Hamann (2011) for a review. According to this approach, grammar universally allows for a relationship between an anaphoric expression and the antecedent to be established via a syntactic mechanism (i.e., binding), or via a discourse-level operation (i.e., co-reference). Unlike in the former, in the latter, the anaphoric expression is not syntactically bound by the antecedent, i.e., via co-indexation in a c-command configuration, but refers to the same individual “accidentally”, through contextual licensing. In sentences with referential antecedents, both mechanisms are available, although accidental co-reference is strictly limited to certain contexts (Grodzinsky & Reinhart, 1993; Heim, 1993). For quantificational antecedents, on the other hand, accidental co-reference is not possible, and variable binding, the only potential option for the bound reading, is ruled out by Principle B. Since children consistently reject bound readings for sentences with quantificational antecedents, this implies that they have knowledge of Principle B.

In light of this evidence, the non-adult-like bound readings for pronouns with referential antecedents have been attributed to young children being too permissive with co-reference for non-co-indexed NPs due either to incomplete acquisition of a relevant syntax-discourse interface principle (Avrutin & Wexler, 1992; Chien & Wexler, 1990) or processing limitations limiting use of such a principle (Grodzinsky & Reinhart, 1993). Consequently, the earlier “Delay of Principle B” label was replaced with the “Pronoun Interpretation Problem” (PIP) and new explanations for this phenomenon have been put forward, including a version of a processing limitations account utilizing Reuland’s (2001) Primitive’s of Binding approach (Baauw, Zuckerman, Ruigendijk, & Avrutin, 2011). Others questioned the quantificational-referential asymmetry (Elbourne, 2005) and attributed PIP to a flawed

²The two other principles in the classic Binding theory were Principle A, requiring anaphors to be locally bound, and Principle C, requiring R-expressions to be free.

³Russian-speaking children’s performance with pronouns and quantificational NPs as antecedents was shown to be low (41% of responses with incorrectly accepted local universally quantified antecedent (Avrutin & Wexler, 1992). The explanation proposed for the contrast between Russian and English acquisition findings in this respect was related to differences between the English universal quantifier *every* and its Russian equivalent *kazhdyj*, and are beyond the scope of this article.

experimental design, in which sentences are presented in pragmatically infelicitous contexts (Conroy, Takahashi, Lidz, & Phillips, 2009).

Studies in languages with more diverse types of anaphoric expressions, such as Dutch and Greek, offered a new perspective on the PIP phenomenon suggesting that accounts of anaphora acquisition must involve not only the knowledge of structural and pragmatic principles governing binding and co-reference, but also the lexical features of various anaphoric expressions (Philip & Coopmans, 1996; Sigurjónsdóttir & Coopmans, 1996). These and other studies (Ruigendijk et al., 2011; Varlokosta, 2002) reported the ECM-effect, i.e., the PIP phenomenon being much more pronounced for ECM pronominal subjects, as in (3) compared to pronominal objects, as in (1b), for TD children as late as 8-years of age. In fact, in clitic languages, the ECM structure appears to be the only context that give rise to PIP (Baauw & Cuetos, 2003), a finding echoed in a number of studies of adult patients with agrammatic aphasia, discussed below. The accounts proposed for this pattern of performance included theoretical notions posited by the Reflexivity account (Reinhart & Reuland, 1993) also discussed below.

3 *The girl_i sees her_i dance.

1.2. Referential dependencies in acquired language disorders

Pronoun interpretation received substantial attention in the literature on adult agrammatism (Edwards & Varlokosta, 2007; Y Grodzinsky, Wexler, Chien, Marakovitz, & Solomon, 1993) in English; Ruigendijk, Vasic, and Avrutin (2006) and Vasi (2006) in Dutch; Baauw and Cuetos (2003) in Spanish. It was suggested that looking at the pattern of errors in establishing anaphoric relations is not only instructive with respect to identifying impaired components of the linguistic system, but it can also reveal how the grammar of anaphora resolution is organized in general (as part of the unimpaired human language faculty) and serve as an important way to evaluate competing linguistic theories. These studies investigated the contrast in interpretive accuracy between sentences with pronouns and reflexives as matrix objects and those with anaphoric expressions as ECM subjects of small clauses (Baauw & Cuetos, 2003; Edwards & Varlokosta, 2007; Ruigendijk et al., 2006). They argued that the observed lower accuracy with pronouns as ECM subjects in comparison with pronouns as matrix objects provides support against the classic Binding theory (Chomsky, 1981) in favor of the Reflexivity (Reinhart, 1983; Reinhart & Reuland, 1993) framework.⁴ This is because the latter decoupled structural from semantic aspects of anaphora resolution. It reformulated the Binding principles as non-structural conditions on predicates' ability to receive reflexive interpretation, related to their reflexive marking and applicable only to co-arguments of a given predicate. It separated these principles from purely configurational effects. The latter were re-formulated as a condition on chain formation, applicable to all A-chains, i.e., any sequence of co-indexed elements with the head of the chain in an argument position, including anaphoric expressions and their local antecedents. This theory provided a conceptual framework able to accommodate the latest cross-linguistic findings from the acquisition and acquired language disorders literature better than the classic Binding theory, which leaves the ECM effect unexplained.

⁴As well as later developments in the theory, such as Primitives of Binding (Reuland, 2001).

1.3. Reflexivity

The Reflexivity framework (Reinhart & Reuland, 1993) was developed to account for the observations from cross-linguistic studies of anaphora in adult grammar, which revealed variations in the systems of anaphoric expressions that go beyond the clear-cut English-type dichotomy between pronouns and reflexives. The precise linguistic motivation for this framework is beyond the scope of this article (see Reuland, 2003, for a thorough review). What is relevant for the issue at hand is that the new system of rules proposed to regulate referential dependencies between an expanded class of anaphoric expressions and antecedents has two components replacing the classic unitary Binding Principles, namely the newly formulated Binding conditions, which license the reflexivity of predicates, and the Chain Condition, which regulates permissible sequences of co-indexed elements. The relevant theoretical Binding Condition-related definitions are listed in (4) and Chain Condition-related definitions are listed in (4):

- 4 Binding Condition A: a reflexive-marked syntactic predicate must be interpreted reflexively.

Binding Condition B: a reflexive semantic predicate must be reflexive-marked.

A predicate is reflexively marked iff it is lexically reflexive or one of its arguments is a SELF-anaphor.

A predicate is reflexive iff two of its arguments are co-indexed.

Thus, Condition A requires a syntactic predicate to be interpreted reflexively, i.e., to receive an interpretation in which two of its arguments are co-indexed, if it is either lexically reflexive or has a so-called SELF-anaphor (the type of reflexive pronouns equivalent to the English reflexives like “himself”) as an argument. Condition B, conversely, requires for any reflexively interpreted predicate to be reflexively marked, i.e., to either be lexically reflexive or have a SELF-anaphor as its argument. In this system, Condition A licenses the grammaticality of the sentence in (5), while no other condition rules it out. The ungrammatical sentence in (6) is correctly ruled out by Condition B because the verb is neither lexically reflexive nor has a self-anaphor as its argument.

- 5 Peter_i loves himself_i.

- 6 *Peter_i loves him_i.

Within the Reflexivity framework, the relationship between the pronoun and its potential antecedent in sentences like (5–6) and (7) is regulated by different constraints because they differ with respect to the relations between the pronoun and its potential antecedent, the matrix subject.

- 7 *Peter_i heard him_i cry.

In sentences like (6), with a pronoun as the object of a matrix verb, both the pronoun and its potential antecedent, subject NP, are arguments of the same predicate. Thus, the verb forms a potentially reflexive semantic predicate, ruled out by Binding Principle B, as the reflexive interpretation requires reflexive marking of the predicate, absent in such sentences. In ECM sentences like (7), on the other hand, the matrix subject and the pronoun are not co-

arguments, as the internal argument of the matrix verb is the entire small clause containing the pronoun as the subject (i.e., the external argument) of the lower verb. Thus, the pronoun is thematically unrelated to the matrix verb (even though it is in a Case relation with it, i.e., is assigned Accusative Case by the matrix verb). Here, the matrix verb does not form a potentially reflexive predicate; Principle B does not apply and cannot explain the ungrammaticality of (7). The principle proposed to correctly rule out the bound interpretation in (7) is the A-Chain Condition given in (8), which applies in both matrix clauses and ECM sentences, although it is superfluous in the former. The A-Chain Condition applies to the sequences in which the tail of the chain is a trace, co-indexed with the head, i.e., phrase moved into an A-position, and also to those that arise through co-indexation for interpretive purposes.

- 8** A maximal A-chain ($\alpha_1 \dots \alpha_n$) has exactly one link - α_1 - which is both [+R] and marked for structural Case.

An element is [+R] iff it carries a full specification for phi-features and is referentially independent.

Pronouns are [+R], anaphors [-R].

The A-chain condition states that the head of the chain must be occupied by a fully referential element (i.e., a full NP or a pronoun), while the tail of the chain may only be occupied by a referentially deficient element, i.e. not fully specified for morphosyntactic phi-features (person, number and gender). Thus, the A-chain condition rules out the example in (7) because there a chain is created, in which the tail is a referentially independent [+R] element (i.e., a pronoun).

To be able to correctly rule out bound interpretations for pronouns, the child needs to have knowledge of the lexical properties of anaphoric elements and predicates, Binding Conditions (operating at the syntax-semantics interface) and the purely syntactic A-chain condition (plus the discourse-level mechanism licensing accidental co-reference). Looking at children's bound interpretation of pronouns as matrix objects as opposed to ECM subjects allows us to isolate syntax from other domains while pursuing the question of what domain is the source of difficulty in typical and atypical language development, the question we addressed in our study.

1.4. Referential Dependencies in DLD

Only a few studies of anaphoric dependencies have been conducted with children with DLD. One study (Montgomery & Evans, 2009) looked at the comprehension of complex constructions by children with DLD ages 9 and above and their TD peers. This study did not focus on specific questions related to children's capacity to compute referential relations between two elements, but used biclausal sentences with a second conjunct containing anaphoric expressions (either pronouns or reflexives) as a type of complex construction to study the degree to which children with DLD underperform on syntactic structures with embedded clauses and to establish whether this deficit is related to deficits in attention and memory. This study did not report accuracy results separately for pronouns and reflexives, and found that children with DLD underperformed on bi-clausal sentences with anaphoric

expressions relative to monoclausal sentences without such expressions and relative to TD children's performance. They also found that children's performance on the biclausal sentences was related to children's attentional/resource capacity, supporting the processing view of DLD.

Another study with 9–12-year-old children with DLD focused on the pronoun-reflexive contrast in sentences with referential and quantificational antecedents in simple transitive constructions and sentences with the target structure embedded as a subordinated clause (as in examples 9–10) using a picture verification task (van der Lely & Stollwerck, 1997). The results with simple transitive constructions showed that children with DLD performed as well as TD children on quantifier-pronoun and name-reflexive sentences, but significantly underperformed on the name-pronoun and quantifier-reflexive sentences.

On the bi-clausal sentences, children with DLD again underperformed on the name-pronoun sentences and performed poorly on both name-reflexive and quantifier-reflexive conditions, as those in (9) and (10). For example, children with DLD frequently accepted (9) as true in the context in which Mowgli was tickling himself, demonstrating that when lexical knowledge of reflexives (i.e., knowing that a reflexive expresses a self-directed action) is insufficient, and syntactic knowledge (i.e., that it must be co-referential with a local c-commanding antecedent) is needed to rule out improper co-reference, their accuracy was not significantly different from chance. Their accuracy was also at chance-level with quantifier-reflexive sentences, as (10), which they accepted as true in contexts in which the monkeys were tickling Mowgli and also in which the monkeys were tickling each other.

9 Mowgli says *Baloo Bear is tickling himself*.

10 Mowgli says *every monkey is tickling himself*.

Thus, the performance of children with DLD mirrored that of younger TD children on sentences with pronouns and referential antecedents, while they also exhibited difficulties with reflexives not previously attested in TD children. This finding was interpreted as evidence that children with DLD exhibit not just a delay in the development of this aspect of grammar, but have a qualitatively different grammar. In particular, since children with DLD performed at chance in situations in which lexical knowledge was insufficient for ruling out inappropriate co-reference, it was argued that their performance indicates a lack of syntactic knowledge of the Binding Principles (as formulated in the classic Binding theory) and syntactic representations underspecified with respect to co-indexation. These deficits were conceptualized as a manifestation of a broad representational deficit in building complex syntactic dependencies.

In a recent study of Greek (Varlokosta & Nerantzini, 2013), children with DLD were tested on sentences with pronominal clitics and reflexives in simple and complex (ECM and secondary predicate) contexts. These results indicated that for both DLD and TD groups, the complex contexts were more difficult than simple ones, particularly in the clitic conditions. However, for the DLD group, in simple contexts, reflexives were substantially more difficult than clitics (with both referential and quantificational antecedents), a contrast not found with their TD age peers.

Thus, studies demonstrated that for children with DLD 1) error rates in interpretation of anaphoric expressions increased with the syntactically complexity of the target sentence; 2) reflexives, not just pronominals, presented difficulty, as would be expected based on the data from TD children. Cross-linguistic studies, however, produced additional data revealing a more complex pattern of results with respect to the relative difficulty of anaphoric expressions⁵. Thus, in the English study (van der Lely & Stollwerck, 1997), strong pronouns with referential antecedents were difficult for the DLD group in both simple and complex contexts, while reflexives with referential antecedents presented difficulty only in complex contexts. However, in the Greek study (Varlokosta & Nerantzini, 2013; cf. also Varlokosta, 2002), pronominal clitics were easy in simple contexts (irrespective of the antecedent type), but reflexives were difficult (regardless of the antecedent type) in both simple and complex contexts, albeit syntactic complexity exacerbated their difficulty.

Thus, the outstanding research questions are 1) whether the ECM effect applies to strong pronouns for children with DLD, as previous studies of DLD looking at ECM structure involved clitics; 2) whether the difficulty with reflexives reported previously with children with DLD can be replicated in languages in which reflexives possess properties different from English; and 3) whether the difficulty with reflexives is limited to structures of greater syntactically complexity, such as ECM constructions, or is a more general phenomenon.

1.5 Anaphors and Pronouns in Russian

The system of pronouns and reflexives in Russian resembles that of English in some basic ways. For example, in both languages, there is a clear-cut distinction between reflexives, which can, and pronouns, which cannot, have a local antecedent, illustrated in (11).

- 11** Peter_i lubit ego_{a_i/j} /sebj_{a_i/*j}
 Peter_{MASC-NOM-SING} love_{3rd-SING} he_{ACC} /self_{ACC}
 "Peter loves him/himself."

Russian reflexives, however, have properties distinct from English. For a detailed description of anaphor binding in Russian, see Rappaport (1986); Avrutin and Wexler (1992); (Bailyn, 2002). Below, we will summarize those properties of the Russian system of anaphoric expressions that are relevant to the current investigation.

First, unlike polymorphemic self-anaphors in English, the Russian self-anaphor (*sebia*) is morphologically simplex and not morphologically derived from the personal pronouns. Second, unlike in English, in which both pronouns and reflexives have gender/number marking (e.g., *him – himself, they-themselves, etc.*) rendering the morphological marking ambiguous with respect to the referentiality status of the element, in Russian, only the pronouns have gender/number marking (in concord with their antecedents), while reflexives lack person, gender, number marking and are insensitive to phi-features of the antecedent. Thus, the same reflexive pronoun *sebia* is used with antecedents that are feminine, masculine, neuter, and plural, as in example (12–13), while the corresponding sentences

⁵The data are also mixed with respect to the relative difficulty of strong versus weak pronouns for children with DLD in clitic languages. See Varlokosta and Nerantzini (2013) for a summary of these findings.

of the complementizer “chtoby”, which, he argued, at early acquisition stages selects for a tenseless clause and the syntactic features [+AGR, –TNS], with the past tense morphology being analyzed by children as such. The issue of long-distance binding is beyond the scope of the present study. However, this phenomenon may lead to a general delay in the acquisition of reflexives in children with DLD manifesting in a variety of syntactic contexts, including simple and ECM constructions, the focus of our investigation,

2.1 Predictions

Based on the Reflexivity account, we can put forward a set of not mutually exclusive, but distinct predictions with regards to expected relative accuracy across various sentence and anaphoric element types depending on the core DLD deficit we postulate. Thus, if children with DLD lack the knowledge of the Binding Conditions, this should affect their performance on simple transitive sentences, but not on ECM constructions, as these apply only to the former and not the latter. If the source of difficulty involves a broad deficit in establishing discontinuous syntactic dependencies linked to syntactic complexity (van der Lely et al., 2011), children with DLD should underperform on the more syntactically complex ECM conditions, including the reflexive, relative to the simple transitive sentences. If the weakness is related to applying the A-chain condition, this should selectively affect children’s performance on the ECM pronouns, and not ECM reflexives, as this condition is what rules out the ungrammatical bound readings for ECM pronouns. If the key deficit is related to lexical processing, e.g., a weakness in retrieving featurally correct representations of inflectional elements (with both phi-features and case for pronouns or case features only for reflexives), performance on the ECM pronouns should, again, be affected. However, this weakness may also affect children’s performance in simple transitives, as applying the Binding conditions involves determining the reflexivity status of the arguments. Children may fail to identify a reflexive element as reflexive and assign the sentence a free reading or incorrectly identify a pronoun as a reflexive assigning the sentence a bound reading. Thus, the latter deficit is compatible with the least selective pattern of errors. Finally, children with DLD may show the pattern of performance demonstrated for younger TD children, both in Russian and other non-clitic languages, namely pronouns being generally harder than reflexives. We tested these predictions in the experiment that we will present and discuss next. These predictions are summarized in Table 2.

3. Method

3.1 Participants

To investigate children’s interpretation of sentences with intrasentential anaphoric dependencies, we recruited 54 native-Russian-speaking children between the ages of 5;1 and 10;11. Of these, 18 children (9 female) between the ages of 5;3 and 10;4 with a mean age of 7;4 ($SD = 1;9$) had been identified as having DLD. As a comparison group, we recruited 36 TD children (18 female), between the ages of 5;1 and 10;11 with a mean age of 7;8 ($SD = 1;9$). The groups were matched on age, $t(52) = .68, p = .50$.

All of the children were native speakers of Russian, with both groups recruited from similar rural populations matched on geographic, dialectal, and socio-economic indicators. The

DLD classification was based on measures of expressive language using a standardized narrative task adapted for diagnosing DLD in Russian (see below). The recruitment of children was carried out through the local kindergartens, elementary and secondary schools. The normative sample of TD children was recruited following teachers' referrals, who had nominated children with no language, reading or academic difficulties for participation in our study. Informed consents were collected from all study participants' parents and oral assents from each child prior to each testing session. The study was approved by the proper Institutional Review Boards in both USA and Russia.

All of the children included in the DLD group displayed a nonverbal IQ score above the cut-off for intellectual disability (≥ 75 ; $M = 97.00$, $SD = 14.64$) as measured by either the Cattell's Culture-Fair Intelligence Test Scale 2 (CFIT; Cattell & Cattell, 1973) or the Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, 1998). When scores from both tests were available, the higher score was used. Both tests are aimed at assessing nonverbal intelligence and were administered individually. None of the children were diagnosed with autism, Down syndrome, hearing impairment or any other known genomic and/or developmental disorders as indicated by their medical records.

3.2. Language assessment

Since no published standardized language development assessments were available for Russian at the time of the study, we used a narrative task for ascertaining children's language development status, which had been validated as a measure of language development in TD children and clinical populations (Botting, 2002; Norbury & Bishop, 2003). To elicit a narrative, we used wordless storybooks *Frog, Where Are You?* (Mayer, 1969) and *One Frog too Many* (Mayer & Mayer, 1975). The audio and the transcripts of the narratives were analyzed by two native-Russian linguists and rated on: 1) Phonetic and Prosodic Characteristics (i.e., phonological simplifications and omissions, phonological substitutions, and prosodic abnormality); 2) Wellformedness (frequency of ill-formed structures adjusted for the length of the narrative); 3) Syntactic Complexity (frequency of complex structures: embedded and conjoined clauses, passives, and participial constructions); 4) mean length of utterance in words (MLU_w); 5) Lexical Richness (i.e., a ratio of distinct lexemes to the number of words); and 6) frequency of Semantic and Pragmatic Errors.

For the purposes of this study, children were included in the DLD group if they exhibited performance at the level below $-1 SD$ from the mean of an external normative comparison group used as a standardization sample (anonymous, 2013) on at least two of the six measures listed above using age-adjusted robust Z-scores. All children displayed marked deficits in grammatical and/or syntactic development: Wellformedness ($n=7$), Syntactic Complexity ($n=1$), MLU ($n=4$), or any combination of these measures ($n=6$). Thus, all of the children included in the DLD group exhibited a deficit in expressive grammar, which in some, but not all, of the children co-occurred with deficits in other language domains.

3.3. Materials

To test children's comprehension of sentences with anaphoric dependencies, we used a picture selection task. It consisted of four experimental and one filler condition with 6 items per condition. The experimental conditions followed a 2 (pronoun, reflexive) X 2 (transitive, ECM) design. In the pronoun sentences, both the subject NP and the object/ECM subject pronoun were of the same gender to allow the noun to be a potential antecedent for the pronoun. Reflexives in Russian do not have gender/number features, therefore, the same reflexive is used with nouns of any gender. Below are examples of the four experimental conditions:

- 16** Babushka poshchekotala jejo.
grandmother_{FEM-NOM-SING} tickle_{PAST-FEM-SING} her_{FEM-ACC-SING}
"The grandmother tickled her."
- 17** Babushka poshchekotala sebia
grandmother_{FEM-NOM-SING} tickle_{PAST-FEM-SING} self_{ACC}
"The grandmother tickled herself."
- 18** Malchik uvidel jego prygajushchim
boy_{MASC-NOM-SING} see_{PAST-MASC-SING} him_{MASC-ACC-SING} jumping_{MASC-INSTR-SING}
"The boy saw him jumping."
- 19** Malchik uvidel sebia prygajushchim
Boy_{MASC-NOM-SING} see_{PAST-MASC-SING} self_{ACC} jumping_{MASC-INSTR-SING}
"The boy saw himself jumping."

Each item consisted of a sentence presented orally by a native speaker of Russian and three pictures, from which the child had to select the one that was the best depiction of the meaning of the sentence. The response options for each item included the picture matching the correct interpretation, a picture consistent with an incorrect anaphoric interpretation (e.g., a bound reading for pronouns or free reading for reflexives), and a distractor, which had the same actors as the sentence, but an unrelated non-reflexive and non-transitive action (i.e., one of the actors performing an action directed towards an inanimate object rather than towards him- or herself or the second actor; see Appendix A for sample items). In order to make the ECM reflexive sentences pragmatically plausible, in both ECM conditions, the relevant actor was depicted looking in the mirror and seeing either him/herself or the other character performing the action mentioned in the sentence). The items were balanced between masculine and feminine subjects.

The filler condition consisted of sentences with a tensed subordinate clause with a pronoun as the subject of the embedded clause, as in (20). In this sentence, the pronoun is allowed to co-refer with the matrix subject; thus, the sentence is fully ambiguous between the bound and free reading. Thus, two of the picture options constituted a correct response and only the distractor was an incorrect response option. The complete set of the verbal stimuli are available in Appendix B.

- 20** Dedushka_i думаet chto on_{i/j} plachet
grandpa_{MASC-NOM-SING} think_{3rd-SING} that he_{MASC-NOM-SING} cry_{3rd-SING}

“The grandfather thinks that he is crying.”

Before being presented the experimental items, the participants were given two training items to familiarize them with the procedure, which all of the participants passed.

The children were tested at their schools, individually, in a quiet room. They were instructed to listen to each sentence and then select the picture that best matched the sentence.

4. Results

The mean accuracy across the four experimental conditions for each group is shown in Figure 1.

Following Jaeger (2008), a set of mixed 2 (anaphoric type) x 2 (sentence type) x 2 (group) mixed logit models was fitted at the item-level data with crossed random effects for items and subjects⁶. Mixed effects logit models are a type of generalized linear mixed model that employs a logit link function for dichotomous dependent variables. In these models, the outcome is analyzed as the linear combination of fixed (corresponding to the effects of independent variables) and random effects (e.g., participant- and item-specific effects that model the variability at the level of individual participants and items, respectively). This approach is particularly suitable for analyzing item-level binary accuracy data and has important advantages (Dixon, 2008; Jaeger, 2008; Quené & van den Bergh, 2008) over traditional ANOVA performed on average accuracy scores: e.g., mixed models do not make the assumption of homogeneity of variances, and have overall higher statistical power combined with better control of Type I error.

The modeling of conditional response probabilities was performed in R using the *lmer* function from the *lme4* library (Bates & Maechler, 2010) with Laplace approximation.⁷ To avoid collinearity between the main effects and interactions and to ease the interpretation of regression coefficients (Est.), all factors were coded using a set of weighted contrasts; so that the intercept represented weighted grand average log(odds) of correct response and the unstandardized coefficients for factor variables represented their fixed effects.

The model with fixed effects of group, anaphoric type, and sentence type, as well as random effects of participants and items on the intercept, displayed better fit than the baseline model with just random effects, $\chi^2(7) = 38.21, p < .00001$. We found a statistically significant effect of group, with the DLD group underperforming overall (Est. = $-.94$, $SE = .27$, $p = .0003$). In addition, there was a significant effect of sentence type, with lower accuracy in the ECM sentences (Est. = -1.80 , $SE = .43$, $p < .0001$), but no effect of anaphoric type (Est. = $.68$, $SE = .43$, $p = .1074$). However, we found a significant anaphoric type by sentence type interaction (Est. = 2.40 , $SE = .86$, $p = .0050$). The follow-up analysis revealed a significant effect of anaphoric type for ECM constructions (Est. = 1.90 , $SE = .69$, $p = .0061$) but not for

⁶We also examined the possibility of including random effects (slopes) of sentence type and pronominal type; however, this did not improve the fit of the model.

⁷An item analysis revealed an outlier item in the reflexive condition, which had a significantly higher error rate compared to other items in this condition in the TD group (under 70% correct compared to the nearly 90% mean accuracy in the condition). After inspecting the pictures for this item, we concluded that this unusually high error rate may be attributable to visual confusability of the picture options and excluded this item from further analyses.

transitive sentences ($Est. = -.55$, $SE = .49$, $p = .2625$), indicating that in the former but not the latter conditions, sentences with pronouns were more difficult than those with reflexives. Finally, a group by anaphoric type interaction was also significant ($Est. = -.82$, $SE = .33$, $p = .0126$). The follow-up analyses showed that while for the children with DLD, the effect of anaphoric type was not significant and close to zero ($Est = .12$, $SE = .43$, $p = .7746$), it was larger and significant for the TD children ($Est = .91$, $SE = .46$, $p = .0475$; higher accuracy in reflexives conditions).

To better understand the nature of the interpretive errors made by the children, we examined the picture choices made by the children on the error trials to see whether their incorrect responses stemmed predominately from allowing improper anaphoric dependencies or were equally likely to be the distractor option. For the purpose of this analysis, all incorrect responses were categorically coded as distractors (i.e., a picture with the same actors as the target but with an unrelated action), or anaphoric errors (i.e., a picture with the same actors and action, but representing inappropriate anaphoric resolution, namely a free reading in a reflexive condition or bound reading in a pronoun condition).

The reasoning behind this analysis was that if the child was equally likely to make distractor or anaphoric errors, these errors were likely general performance errors, i.e., not directly driven by a failure to apply the relevant principle regulating anaphora resolution, but by a more general failure to interpret the sentence (e.g., due to its processing complexity exceeding their processing resources or as a result of a momentary attention lapse during stimulus presentation) and, consequently, selecting the picture arbitrarily. Such pattern of errors is not an indicator of a deficit in the knowledge of the constraints governing the system of anaphoric relations or the capacity to apply them in real time. If, on the other hand, the child consistently ignored the distractors, her errors were more likely to be an indication of a liability involving an incorrect application of some component of the grammar regulating anaphora resolution (due to a lexical error, a representational deficit involving the knowledge of certain constraint(s) related to establishing anaphoric relations or a limitation in applying them in real time).

To this end, we, first, compared the two groups on the relative frequency of anaphoric vs. distractor errors in each condition by conducting a set of non-parametric Mann-Whitney U tests. No group differences were found for any condition (all p 's $> .26$). Secondly, we used the Wilcoxon Signed Ranks Test to compare the proportions of each type of error for each condition separately in the two groups. The results of this error analysis are presented in Table 3. Note, that the following observations applied only to those children who have committed at least 1 error, and are therefore based on a slightly reduced sample.

For the TD group, we found that in three of the conditions, both error types had similar frequencies: transitive pronoun ($p = .416$), transitive reflexive ($p = .104$), and ECM reflexive ($p = 1.00$). Thus, the occasional errors in these conditions (in all of which the accuracy in this group was high - 80–90%) were likely superficial performance-level errors that did not reflect any “deep”, non-adult characteristics of children’s grammar or sentence processing. However, in the ECM pronoun condition (in which accuracy level for this group was only about 50%, with the chance level performance being 33% correct in this 3-option picture-

choice design), the TD children made significantly more anaphoric rather than distractor errors ($p < .001$). Thus, in this one condition only, the TD children were most likely to make an error by assigning inappropriate local reference to the ECM subject pronouns.

The DLD group displayed a different error pattern. This group had similar rates of distractor and anaphoric errors only in one condition, the ECM reflexives ($p = 1.00$). In three other conditions, i.e., the transitive reflexives, transitive pronouns, and ECM pronouns, significantly more errors were of the anaphoric rather than distractor type (p 's = .024; .013; and .028, respectively). This suggests that only in the ECM reflexive condition were DLD children's errors of a general processing nature; the errors in all other conditions made by children with DLD were likely driven by deficiencies in their knowledge of the grammar of anaphora resolution or the capacity for using it in real time. Thus, unlike their TD counterparts, children with DLD systematically accepted bound readings for pronouns and free readings for reflexives in transitive sentences, in addition to frequently accepting bound readings for ECM pronouns, with only the latter being the case for the TD group as well. Although the DLD group also rather frequently made errors with ECM reflexives ($\approx 35\%$ of the time), these errors likely stemmed from the general syntactic complexity of ECM sentences, instead of the anaphoric element per se.

5. Discussion

Our experimental study revealed a number of noteworthy results. First, we found that children with DLD showed an overall lower performance compared to the TD children, confirming that this population has a deficit in interpreting anaphoric dependencies, in addition to the well-documented deficits with dependencies between moved phrases and their traces and other discontinuous syntactic dependencies (Friedmann & Novogrodsky, 2004, 2011; Montgomery & Evans, 2009; Stavrakaki, 2001; van der Lely et al., 2011).

Secondly, we established certain similarities in the two groups' patterns of performance. Thus, the effect of sentence type generalized to both groups of children, with lower accuracy on the ECM sentences in comparison to the simple transitive sentences. The greater relative difficulty of the ECM sentences relative to the simple transitive counterparts, independently from the anaphoric type, suggests that their interpretive complexity stems not solely from the presence of an anaphoric dependency, but, at least in part, from a greater syntactic complexity of these sentences, which contain a subordinate (small) clause. ECM constructions in Russian have an additional complexity factor, namely the so-called "predicate instrumental", i.e., the predicate participle is case marked discordantly from the ECM subject. The latter is marked with accusative case, but the former with instrumental case (which in Russian marks secondary predicates, in addition to instruments and passive agents), while the two agree in gender/number.

21 Malchik uvidel jejo prygajushchej
 boy_{MASC-NOM-SING} see_{PAST-MASC-SING} her_{FEM-ACC-SING} jumping_{FEM-INSTR-SING}
 "The boy saw her jumping."

If the ECM subject is reflexive and therefore unmarked for gender/number (and must obligatorily co-refer with the matrix subject), the participle agrees in gender/number with the matrix subject:

22 Malchik uvidel sebja prygajushchim
 boy_{MASC-NOM-SING} see_{PAST-MASC-SING} self_{ACC} jumping_{MASC-INTR-SING}
 “The boy saw himself jumping.”

These complex morphosyntactic patterns coupled with an additional predicate phrase (Bailyn, 2002) would explain extra syntactic complexity of ECM sentences. The extra complexity of the ECM sentences, however, did not disproportionately affect children with DLD (as indicated by a lack of a significant group-by-sentence-type interaction), as may be expected under the theories that posit a broad deficit in representing complex structural syntactic relations as a core feature of DLD (van der Lely, 2005; van der Lely & Stollwerck, 1997) or those theories that posit a general processing capacity deficit in DLD that imposes extra vulnerability to errors in complex syntactic structures (Montgomery & Evans, 2009).

With respect to the contrast between pronouns and reflexives, we again found certain similarities in the performance pattern between the two groups. Namely, for both groups, the ECM pronoun condition was more difficult than the ECM reflexive condition, which manifested in a significant sentence type by anaphoric type interaction. In fact, the ECM pronoun was by far the most difficult condition for both groups, resulting in the highest error rate, as well as showing improper local co-reference as the more frequent error type, indicating improper anaphoric resolution rather than general interpretive failure as the cause for the errors. As argued above, being able to systematically ignore the distractors but still frequently choosing improper co-reference is indicative of having assigned the sentence a semantic interpretation (i.e., syntactically parsing the sentence, interpreting the predicate, assigning theta roles to its arguments and establishing anaphoric relations), but failing to rule out improper anaphoric relations. Thus, the high error rate with ECM pronouns in both groups seems likely to stem from the demands directly associated with anaphoric resolution in this configuration. Furthermore, these high demands, as was the case with the factor of sentence complexity, did not disproportionately affect children with DLD. In contrast, on the ECM reflexives, the errors for both groups were as likely to be distractors as improper anaphoric relations, suggesting general processing failure due to syntactic complexity of these sentences, rather than improper anaphoric resolution.

The observation of the greater selective difficulty of ECM pronoun sentences is in line with the previous findings of studies with children as well as adults with agrammatic aphasia (Edwards & Varlokosta, 2007; Ruigendijk et al., 2011; Varlokosta & Nerantzini, 2013). Furthermore, as argued in these studies, adopting the Reflexivity account (T Reinhart & Reuland, 1993) and its theoretical descendants (Reuland, 2001), which decouple Binding Principles (reconceptualized as semantic conditions on predicates) from configurational conditions on co-reference, provides a straightforward explanation for the observed discrepancy in accuracy between ECM pronouns and all other conditions. The interpretation of pronouns in ECM constructions under this approach is regulated by the Chain Condition and not by Binding Conditions, which apply only to co-arguments of the same predicate. Thus, the latter can be used in transitive clauses to rule out improper co-reference even if the

Chain Condition fails to be applied. In ECM constructions, the Chain Condition (and the constructs associated with it) can be pinpointed as a source of difficulty for both groups. A selective difficulty with ECM pronoun sentences, therefore, is consistent with a weakness in ruling out improper chains (i.e., allowing tails to be fully referential), and not a deficit with respect to Binding conditions, as those do not apply in ECM constructions. Furthermore, because the classic Binding theory provides no differential mechanism of anaphoric resolution in ECM as opposed to transitive sentences, it does not explain the selective difficulty of the ECM pronouns, violating the *breakdown-compatibility* assumption (Grodzinsky, 1986).

As far as the more precise explanation of why children often fail to rule out improper chains (i.e., those with fully referential tails), lexical knowledge of pronouns has been previously proposed as a possible cause. It was hypothesized that young children frequently interpret pronouns as [-R] as a consequence of not having fully acquired their featural composition (Baauw, 2002; Philip & Coopmans, 1996). Sentences, in which pronouns are unspecified for phi-features and interpreted as co-indexed with a local c-commanding antecedent would not violate the Chain Condition, explaining why children permit local co-reference for ECM pronominal subjects even when they correctly rule it out in simple transitive sentences using the Binding Conditions.

The underspecification of pronouns, however, seems a less likely explanation in our study. As described above, as part of the language assessment, elicited narrative samples from all of the participants had been analyzed for wellformedness for the purposes of ascertaining children's group status. Unfortunately, no specific measure of pronoun accuracy was derived, as the Wellformedness index we used included all lexical, syntactic and morphosyntactic errors. However, an informal analysis revealed that although both groups frequently used pronouns anaphorically and deictically (the children with DLD relying on pronouns for maintaining reference somewhat more frequently than TD children), there was a contrast in accuracy. While the TD children exhibited no evidence of pronouns being underspecified for phi-features, as they demonstrated a high degree of accuracy in person/gender/number agreement, among the errors documented in the DLD group, there were agreement errors in pronoun use (both in the antecedent-pronoun gender/number agreement and in subject-verb agreement between pronouns and verbs). If this contrast indicates a difference in the lexical knowledge of pronominal phi-features between the children in the TD and DLD groups and if errors on ECM pronouns stem from a failure to recognize pronouns as fully specified for phi-features, it is difficult to explain why it did not translate into discernably different patterns of performance on the ECM pronoun condition for the two study groups. Instead, both groups performed slightly above chance (about 40% and 48% accuracy in the DLD and TD groups, respectively).

Another possible explanation for the difficulty that children with DLD have with ECM pronouns attributes it to insufficient processing resources (Baauw et al., 2011; Ruigendijk et al., 2011). A specific account used to explain the relationship between processing capacity and ECM pronoun errors utilizes the economy hierarchy from the Primitives of Binding theory (Reuland, 2001), which dictates that various ways of establishing dependencies vary in processing costs. In particular, syntactic operations (i.e., feature checking, which in this

framework is used to establish an A-chain) are less costly for the language processor than extra-syntactic ones (e.g., semantic – variable binding, or pragmatic – co-reference) and therefore are always given precedence. If a less costly operation is available, costlier operations are blocked. Thus, if feature-checking is available (and if variable binding and co-reference options would yield indistinguishable from the former interpretation), the latter two are disallowed, and feature checking remains the only permitted option. When the more economical syntactic dependency is blocked by an independent constraint, as is the case with pronouns, which form an A-chain with their antecedents in narrow syntax and thus violate the Chain Condition because of their fully referential status, the “bound reading” is ruled out. In sum, this framework assumes a set of complex computations, including reference set computation (i.e., constructing a global comparison set to evaluate alternative derivations (Reinhart, 1999), which by itself comes with a processing cost. Children, who have lower processing resources than typical adults, cannot cope with the high processing costs involved in this computation and consequently permit an extra-syntactic referential dependency established through discourse even if a “cheaper” option is available, resulting in being too “liberal” with local co-reference.

The results of our study are consistent with an explanation attributing the ECM pronoun errors to an interaction between linguistic computation and processing limitations affecting both groups of children to a similar extent. Such an explanation can account for the above chance (albeit low) accuracy rate in both groups, more consistent with a probabilistic rather than categorical loss of accuracy, and is in line with the growing body of research showing that even TD children above the ages 5–6, generally able to establish complex syntactic dependencies in low demand tasks, are influenced in their performance by limitations in processing capacity during experiments testing on- and off-line sentence comprehension (Clahsen & Felser, 2006).

Finally, in addition to the qualitative similarities between the two groups, we observed one difference in the pattern of performance manifested as an anaphoric type by group interaction. In particular, we found that in the TD group, sentences with pronouns generally resulted in somewhat lower accuracy in comparison to sentences with reflexives, consistent with the processing account outlined above and in concordance with the previous findings across a number of languages that support the theories attributing these errors to a processing (or pragmatic) deficit (Avrutin & Wexler, 1992; Chien & Wexler, 1990; Grodzinsky & Reinhart, 1993) rather than an all-or-nothing representational syntactic knowledge deficit. In contrast, for the DLD group, the effect of anaphoric type was small and not significant, minimized by the relatively low accuracy of the reflexive conditions. To answer conclusively what is behind this lack of the anaphoric type effect in the DLD group requires further experimental studies. However, there is some evidence that it may be related to the application of the Binding principles. Since on error trials children with DLD (unlike their TD counterparts) chose the improper co-reference response significantly more frequently than the distractor response in both pronoun and reflexive conditions (except for ECM reflexives), this suggests that the interpretive failure in all of these conditions in this group was symptomatic of a specific weakness in establishing anaphoric dependencies rather than a general failure to interpret the sentence, as we argued above. Our findings echo the pattern of results from a study of Greek children with DLD (Varlokosta & Nerantzini,

2013), which used clitics instead of strong pronouns, suggesting a common deficit being at play in both studies despite of the differences in the grammatical systems governing the use of pronouns and reflexives in the two languages.

It is unlikely that the deficit in question involves a categorical lack of linguistic knowledge, however, given the relatively high accuracy of the DLD group on the transitive pronouns (85%) and reflexives (70%) in our study. Such relative accuracy reduction (of varying degree on different conditions) is better compatible with a deficit in applying the Binding Conditions during sentence interpretation, rather than with a categorical lack of the Binding Conditions from children's grammars.

We can conjecture that this performance level deficit stems from the children's weakness in lexical processing. During sentence comprehension, the child has to retrieve the appropriate element from their mental lexicon, including its phi-features and reflexivity status. According to many psycholinguistic models of word recognition, the process of lexical access involves early (prior to lexical access) morphological decomposition (Lehtonen et al., 2007; Lewis, Solomyak, & Marantz, 2011; Zweig & Pylkkänen, 2009). These models assume that, before lexical lookup, words undergo mandatory affix stripping (there are different versions of the decomposition model with regards to what types of words are subject to this operation, discussion of which is beyond the scope of the present study). As has been argued with respect to Finnish (Lehtonen et al., 2007), in morphologically rich languages, in which each word comes in a large number of paradigmatic core forms, it is plausible that inflected forms are decomposed during recognition, which would be more economical for memory than storing each inflected form separately. Following this process, the lexical entry for the resulting stem is looked up in the mental lexicon, the stem is recombined with the affixes, and grammatical and lexical knowledge is consulted to evaluate the resulting combination (Taft & Forster, 1975). This makes processing morphologically complex words slower and more error prone than processing simplex stems (Lehtonen et al., 2007) in general, and has been related to procedural memory, a neurocognitive system sub-serving structure-building and concatenating linguistic operations, including morphological processing (Pinker & Ullman, 2002).

A deficit in procedural memory (Ullman, 2001) and/or "fuzzy" or unstable lexical representations (Holm et al., 2007) of anaphoric elements as well as other closed class items (perhaps due to limitations in other memory systems, such as declarative or phonological short-term memory) would result in an increase in errors (and reduction in speed) when processing morphologically complex words in children with DLD relative to their typically developing peers. This would affect children's accuracy on sentence interpretation tasks that require using morphological and other structural cues. Our results, namely, the lack of the anaphoric type effect for the DLD group (both pronouns and reflexives being difficult to a similar degree, unlike the TD group, for whom pronouns presented relatively greater difficulty than the reflexives) and the overall lower accuracy of the DLD group (without this group being more strongly affected by the sentence complexity) are compatible with a deficit in processing surface morphology or lexical access than with a selective or global syntactic or general processing deficit.

6. Conclusion

Our results have confirmed the previous reports that children with DLD exhibit a deficit in interpreting sentences involving anaphor resolution with both pronouns and reflexives. We also found that the ECM effect and sentence complexity affected the performance of children with DLD to an extent similar to which they affected the performance of TD children. On the other hand, the type of the anaphoric element did not have the same effect on the performance of the two groups of children, with pronouns and reflexives being relatively similar in difficulty for the DLD, but not the TD group.

This study contributes to the existing literature in two ways. First, with respect to the theoretical understanding of developmental language disorder, which has important implications for the study of language development in general, we believe we can clarify certain important points concerning the key juxtaposition, in which the debate on the underlying cognitive deficit of DLD is typically couched. Much of the existing literature on DLD (or SLI) conceptualizes the disorder as impairment in either domain-specific or domain-general cognitive systems. The first alternative views DLD/SLI as a deficit at the level of linguistic competence, which may selectively affect computational components of language (grammar), as opposed to its lexical aspects. According to this view, the typically observed overlap between lexical and grammatical deficits in children with DLD are taken as examples of a disorder subtype distinct from what these studies call grammatical SLI or G-SLI (van der Lely, 2005). It is, however, clear from the clinical literature on DLD that there is a relationship between reduced vocabularies and deficits in grammar both concurrently and longitudinally in both typically developing children and children with DLD (Dionne, Dale, Boivin, & Plomin, 2003; Moyle, Ellis Weismer, Evans, & Lindstrom, 2007; Thordardottir, Weismer, & Evans, 2002), an observation that needs to be taken seriously by those developing theoretical accounts of the disorder.

The opposing view argues for a deficit at the level of the child's domain-general capacity to process input, such as a deficit in auditory processing (Tallal, Miller, & Fitch, 1993) or a deficit in working memory (Gathercole & Baddeley, 1990), assumed to affect acquisition of the lexicon and grammar indirectly, as its downstream consequence. The literature that attempted to document auditory processing deficits in children with DLD is mixed (Bishop, Carlyon, Deeks, & Bishop, 1999; Nitttrouer, 1999) and at least for a substantial number of cases, linguistic deficits are not attributable to deficits in auditory processing (Rosen, 2003). Although working memory deficits in this population of children are well documented (Archibald & Gathercole, 2006; Montgomery & Evans, 2009; Weismer, Evans, & Hesketh, 1999), the precise mechanisms behind the associations between reduced working memory (and other memory mechanisms) and the wide range of language deficits in DLD remain rather imprecise and require clarification.

We have argued that our results are inconsistent with both views - categorical all-or-nothing lack of syntactic knowledge deficit sparing the lexicon or a generalized processing deficit related to limited resource capacity. We argued that what we uncovered was a performance-level deficit likely rooted in children's weakness in lexical acquisition and processing, which has direct consequences for accuracy of grammatical processing.. Thus, we found continuity

between children with DLD and their TD counterparts with respect to mental representations and computations in their grammar of anaphor resolution, as well as with respect to their sentence processing systems. The gap between the linguistic systems of the two groups of children seems to lie in the domains of lexical representations and/or lexical processing. If this finding can be corroborated and generalized to DLD in other languages, this would have important implications, reminiscent of the idea that parametric differences between languages are limited to the lexical differences of functional heads (the so-called Borer-Chomsky Conjecture; Baker, 2008), suggesting that DLD impairs only the acquisition of the linguistic features that are idiosyncratic and subject to cross-linguistic variation, but not the universal aspects of language.

With respect to our second goal, namely to examine whether we can provide new empirical support for (or against) the Reflexivity account of Binding (T Reinhart & Reuland, 1993), we indeed found strong evidence converging with previous research that this account provides the best explanation for the developmental data from Russian-speaking children with and without DLD. We believe that supplying new independent empirical corroboration of a theoretical account of a syntactic phenomenon is an important source of evidence for syntactic theory.

Acknowledgments

We thank Roman Kopusov from Northern State Medical Academy (Arkhangelsk, Russia) for his help with data collection. We also gratefully acknowledge the contributions to this research by the late Dr. Maria Babyonyshev.

FUNDING

This research was supported by NIH grant R01 DC007665 to Elena L. Grigorenko. Grantees undertaking such projects are encouraged to express freely their professional judgment. This article, therefore, does not necessarily reflect the position or policies of the National Institutes of Health, and no official endorsement should be inferred.

References

- Acarlar F, Johnston JR. Acquisition of Turkish grammatical morphology by children with developmental disorders. *International Journal of Language & Communication Disorders*. 2011; 46(6):728–738. [PubMed: 22026573]
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5. Arlington, VA: American Psychiatric Publishing; 2013.
- Archibald LMD, Gathercole SE. Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders*. 2006; 41(6):675–693. 10.1080/13682820500442602 [PubMed: 17079222]
- Avrutin S, Wexler K. Development of principle B in Russian: Coindexation at LF and coreference. *Language Acquisition*. 1992; 2(4):259–306.
- Baauw, S. *Grammatical Features and the Acquisition of Reference: A Comparative Study of Dutch and Spanish*. Routledge; 2002.
- Baauw S, Cueto F. The interpretation of pronouns in Spanish language acquisition and breakdown: Evidence for the ‘Delayed Principle B Effect’ as a non-unitary phenomenon. *Language Acquisition*. 2003; 11:219–275.
- Baauw, S.; Escobar, L.; Philip, W. A delay of principle B effect in Spanish speaking children: The role of lexical feature acquisition. Paper presented at the GALA ‘97; Human Communication Centre; 1997.
- Baauw, S.; Zuckerman, S.; Ruigendijk, E.; Avrutin, S. Principle B delays as a processing problem: Evidence from task effects. In: Grimm, A.; Muller, A.; Hamann, C.; Ruigendijk, E., editors.

- Production-Comprehension Asymmetries in Child Language. Belin: Mouton De Gruyter; 2011. p. 247-272.
- Bailyn J. Overt predictors. *Journal of Slavic linguistics*. 2002; 10(1-2):23-52.
- Baker M. The macroparameter in a microparametric world. *The limits of syntactic variation*. 2008; 132:351.
- Bates, D.; Maechler, D. lme4: Linear mixed-effects models using Eigen and Eigenfaces (Version 0.999375-37). 2010. Retrieved from <http://CRAN.R-project.org/package=lme4>
- Bedore LM, Leonard LB. Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research*. 1998; 41(5):1185-1192.
- Bedore LM, Leonard LB. Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech Language and Hearing Research*. 2001; 44(4):905-924.
- Bishop DVM, Carlyon RP, Deeks JM, Bishop SJ. Auditory Temporal Processing Impairment Neither Necessary nor Sufficient for Causing Language Impairment in Children. *Journal of Speech, Language, and Hearing Research*. 1999; 42(6):1295-1310.10.1044/jslhr.4206.1295
- Bortolini U, Caselli MC, Deevy P, Leonard LB. Specific language impairment in Italian: The first steps in the search for a clinical marker. *International Journal of Language & Communication Disorders*. 2002; 37(2):77-93. [PubMed: 12012618]
- Botting N. Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching & Therapy*. 2002; 18(1):1-21.
- Bracken, BA.; McCallum, RS. Examiner's Manual: Universal Nonverbal Intelligence Test (UNIT). Itasca, IL: Riverside Publishing; 1998.
- Cattell, RB.; Cattell, HEP. Measuring intelligence with the culture fair tests. Champaign, IL: Institute for Personality and Ability Testing; 1973.
- Chien YC, Wexler K. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*. 1990; 1(3):225-295.
- Chomsky, N. *Lectures on Government and Binding*. Holland: Foris; 1981.
- Clahsen H, Bartke S, Gollner S. Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics*. 1997; 10(2-3):151-171.
- Clahsen H, Felser C. Grammatical processing in language learners. *Applied Psycholinguistics*. 2006; 27(1):3-42.10.1017/S0142716406060024
- Conroy AE, Takahashi E, Lidz J, Phillips C. Equal treatment for all antecedents: How children succeed with principle B. *Linguistic Inquiry*. 2009; 40:446-486.
- Conti-Ramsden G, Jones M. Verb use in specific language impairment. *Journal of Speech Language and Hearing Research*. 1997; 40(6):1298-1313.
- Dionne G, Dale PS, Boivin M, Plomin R. Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development*. 2003; 74(2):394-412.10.1111/1467-8624.7402005 [PubMed: 12705562]
- Dixon P. Models of accuracy in repeated-measures designs. *Journal of Memory and Language*. 2008; 59:447-456.
- Dromi E, Leonard LB, Adam G, Zadunaisky-Ehrlich S. Verb Agreement Morphology in Hebrew-Speaking Children With Specific Language Impairment. *Journal of Speech, Language & Hearing Research*. 1999; 42(6):1414.
- Edwards S, Varlokosta S. Pronominal and anaphoric reference in agrammatism. *Journal of Neurolinguistics*. 2007; 20:423-444.
- Elbourne P. On the acquisition of principle B. *Linguistic Inquiry*. 2005; 36:333-365.
- Friedmann N, Novogrodsky R. The acquisition of relative clause comprehension in Hebrew: a study of SLI and normal development. *J Child Lang*. 2004; 31(3):661-681. [PubMed: 15612394]
- Friedmann N, Novogrodsky R. Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain Lang*. 2007; 101(1):50-63. S0093-934X(06)00392-0 [pii]. 10.1016/j.bandl.2006.09.006 [PubMed: 17084444]
- Friedmann N, Novogrodsky R. Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua*. 2011; 121(3):367-382.10.1016/J.Lingua.2010.10.004

- Gathercole SE, Baddeley AD. Phonological Memory Deficits in Language Disordered Children - Is There a Causal Connection. *Journal of Memory and Language*. 1990; 29(3):336–360.
- Grodzinsky Y. Language deficits and the theory of syntax. *Brain and Language*. 1986; 27:135–159. [PubMed: 3947938]
- Grodzinsky Y, Reinhart T. The Innateness of Binding and Coreference. *Linguistic Inquiry*. 1993; 24(1):69–101.
- Grodzinsky Y, Wexler K, Chien YC, Marakovitz S, Solomon J. The breakdown of binding relations. *Brain and Language*. 1993; 45:396–422. [PubMed: 8269332]
- Hamann. Binding and coreference: Views from Child Language. In: de Villiers, JG.; Roeper, T., editors. *Handbook of Generative Approaches to Language Acquisition*. New York: Springer; 2011.
- Heim, I. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. MIT; 1993.
- Holm A, Crosbie S, Dodd B. Differentiating normal variability from inconsistency in children's speech: normative data. *Int J Lang Commun Disord*. 2007; 42(4):467–486.10.1080/13682820600988967 [PubMed: 17613100]
- Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*. 2008; 59:434–446. [PubMed: 19884961]
- Jakubowicz C, Nash L. Functional categories and syntactic operations in (ab)normal language acquisition. *Brain and Language*. 2001; 77(3):321–339. [PubMed: 11386700]
- Joanisse MF, Seidenberg MS. Specific language impairment: a deficit in grammar or processing? *Trends in Cognitive Sciences*. 1998; 2(7):240–247. [PubMed: 21244922]
- Lehtonen M, Cunillera T, Rodríguez-Fornells A, Hultén A, Tuomainen J, Laine M. Recognition of morphologically complex words in Finnish: Evidence from event-related potentials. *Brain Research*. 2007; 1148(0):123–137. doi: <http://dx.doi.org/10.1016/j.brainres.2007.02.026>. [PubMed: 17382308]
- Leonard LB, Bortolini U. Grammatical morphology and the role of weak syllables in the speech of Italian-speaking children with specific language impairment. *J Speech Lang Hear Res*. 1998; 41(6):1363–1374. [PubMed: 9859891]
- Leonard LB, Eyer JA. Deficits of grammatical morphology in children with specific language impairment and their implications for notions of bootstrapping. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. 1996:233–247.
- Lewis G, Solomyak O, Marantz A. The neural basis of obligatory decomposition of suffixed words. *Brain and Language*. 2011; 118(3):118–127. doi: <http://dx.doi.org/10.1016/j.bandl.2011.04.004>. [PubMed: 21620455]
- Lum JAG, Gelgic C, Conti-Ramsden G. Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders*. 2010; 45(1):96–107.10.3109/13682820902752285 [PubMed: 19900077]
- Manzini MR, Wexler K. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry*. 1987; 18(3):413–444.
- Marshall C, Marinis T, van der Lely HKJ. Passive verb morphology: The effect of phonotactics on passive comprehension in typically developing and Grammatical-SLI children. *Lingua*. 2007; 117(8):1434–1447.
- Mastropavlou M, Tsimpli IM. Complementizers and subordination in typical language acquisition and SLI. *Lingua*. 2011; 121(3):442–462.10.1016/J.Lingua.2010.10.009
- Mayer, M. *Frog, Where are you?*. New York: Dial Press; 1969.
- Mayer, M.; Mayer, M. *One Frog Too Many*. New York: Dial Press; 1975.
- McKee C. A comparison of pronouns and anaphors in Italian and English acquisition. *Language Acquisition*. 1992; 2(1):21–54.
- Montgomery JW, Evans JL. Complex sentence comprehension and working memory in children with specific language impairment. *J Speech Lang Hear Res*. 2009; 52(2):269–288. 1092-4388_2008_07-0116 [pii]. 10.1044/1092-4388(2008/07-0116) [PubMed: 18723601]

- Moyle MJ, Ellis Weismer S, Evans JL, Lindstrom MJ. Longitudinal Relationships Between Lexical and Grammatical Development in Typical and Late-Talking Children. *Journal of Speech, Language, and Hearing Research*. 2007; 50(2):508–528.10.1044/1092-4388(2007/035)
- Nittrouer S. Do Temporal Processing Deficits Cause Phonological Processing Problems? *Journal of Speech, Language, and Hearing Research*. 1999; 42(4):925–942.10.1044/jslhr.4204.925
- Norbury CF, Bishop DVM. Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders*. 2003; 38(3):287–313. [PubMed: 12851080]
- Penke, M. Morphology and Language Disorder. In: Ball, MJ.; Perkins, MR.; Muller, N.; Howard, S., editors. *The Handbook of Clinical Linguistics*. Oxford: Wiley-Blackwell; 2009. p. 212-227.
- Philip, W.; Coopmans, P. The role of lexical feature acquisition in the development of pronominal anaphora. Paper presented at the Amsterdam Series on Child Language Development; Amsterdam. Instituut Algemene Taalwetenschap; 1996.
- Pinker S, Ullman MT. The past and future of the past tense. *Trends in Cognitive Sciences*. 2002; 6(11):456–463. doi: [http://dx.doi.org/10.1016/S1364-6613\(02\)01990-3](http://dx.doi.org/10.1016/S1364-6613(02)01990-3). [PubMed: 12457895]
- Quené H, van den Bergh H. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*. 2008; 59(4):413–425.
- Rappaport G. On Anaphor binding in Russian. *Natural Language & Linguistic Theory*. 1986; 4(1):97–120.
- Reinhart, T. Anaphora and semantic interpretation. London: Croom Helm; 1983.
- Reinhart, T. The processing cost of reference-set computation: Guess patterns in acquisition. Utrecht University: Utrecht Institute of Linguistics OTS; 1999.
- Reinhart T, Reuland E. Reflexivity. *Linguistic Inquiry*. 1993; 24(4):657–720.
- Reuland E. Primitives of binding. *Linguistic Inquiry*. 2001; 32(3):439–492.10.1162/002438901750372522
- Reuland E. Anaphoric dependencies: A window into the architecture of the language system. *Glott International*. 2003; 7(1/2):3–25.
- Rosen S. Auditory processing in dyslexia and specific language impairment: is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics*. 2003; 31(3–4):509–527. doi: [http://dx.doi.org/10.1016/S0095-4470\(03\)00046-9](http://dx.doi.org/10.1016/S0095-4470(03)00046-9).
- Ruigendijk, E.; Baauw, S.; Zuckerman, S.; Vasic, N.; de Lange, J.; Avrutin, S. A cross-linguistic study on the interpretation of pronouns by children and agrammatic speakers: Evidence from Dutch, Spanish, and Italian. In: Gibson, E.; Pearlmutter, NJ., editors. *Processing and Acquisition of Reference*. MIT Press; 2011. p. 133-155.
- Ruigendijk E, Vasic N, Avrutin S. Reference assignment: Using language breakdown to choose between theoretical approaches. *Brain and Language*. 2006; 96(3):302–317. [PubMed: 16085300]
- Sigurjónsdóttir, S.; Coopmans, P. Amsterdam Series on Child Language Development, ASCLD 5. Vol. 68. Amsterdam: Instituut Algemene Taalwetenschap; 1996. The acquisition of anaphoric relations in Dutch.
- Stavroulaki S. Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain Lang*. 2001; 77(3):419–431. S0093-934X(00)92412-X [pii]. 10.1006/brln.2000.2412 [PubMed: 11386707]
- Taft M, Forster KI. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*. 1975; 14(6):638–647. doi: [http://dx.doi.org/10.1016/S0022-5371\(75\)80051-X](http://dx.doi.org/10.1016/S0022-5371(75)80051-X).
- Tallal P, Miller S, Fitch RH. Neurobiological Basis of Speech: A Case for the Preeminence of Temporal Processing. *Annals of the New York Academy of Sciences*. 1993; 682(1):27–47.10.1111/j.1749-6632.1993.tb22957.x [PubMed: 7686725]
- Thordardottir ET, Weismer SE, Evans JL. Continuity in lexical and morphological development in Icelandic and English-speaking 2-year-olds. *First Language*. 2002; 22(1):3–28.10.1177/014272370202206401
- Ullman MT. A neurocognitive perspective on language: The declarative/procedural model. *Nat Rev Neurosci*. 2001; 2(10):717–726. [PubMed: 11584309]

- Ullman MT, Pierpont EI. Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*. 2005; 41(3):399–433.10.1016/S0010-9452(08)70276-4 [PubMed: 15871604]
- van der Lely HKJ. Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in Cognitive Sciences*. 2005; 9(2):53–59. doi: <http://dx.doi.org/10.1016/j.tics.2004.12.002>. [PubMed: 15668097]
- van der Lely HKJ, Jones M, Marshall C. Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua*. 2011; 121(3):408–422.10.1016/j.lingua.2010.10.007 [PubMed: 21318176]
- van der Lely HKJ, Stollwerck L. Binding theory and grammatical specific language impairment in children. *Cognition*. 1997; 62(3):245–290. S0010-0277(96)00783-4 [pii]. [PubMed: 9187060]
- Varlokosta, S. (A)symmetries in the acquisition of principle B in typically- developing and specifically language impaired (SLI) children. In: Lasser, I., editor. *The Process of Language Acquisition*. Berlin: Peter Lang Verlag Publishing; 2002. p. 81-98.
- Varlokosta, S.; Nerantzini, M. In: Gavrilidou, Z.; Efthymiou, A.; Thomadaki, E.; Kambakis-Vougiouklis, P., editors. *Pronominal and anaphoric reference in Greek specific language impairment; Selected Papers of the 10th International Conference on Greek Linguistics*; Komotini: Democritus University of Thrace; 2013. p. 584-591.
- Vasi, N. Doctoral Dissertation. Utrecht Institute for Linguistics, OTS, LOT Series; 2006. Pronoun comprehension in agrammatic aphasia. The structure and use of linguistic knowledge.
- Weismer SE, Evans J, Hesketh LJ. An Examination of Verbal Working Memory Capacity in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42(5):1249–1260.10.1044/jslhr.4205.1249
- Zweig E, Pylkkänen L. A visual M170 effect of morphological complexity. *Language and Cognitive Processes*. 2009; 24(3):412–439.10.1080/01690960802180420

Appendix 1. Sample items

Conditions 1 & 2: “Grandmother tickled her/herself.”

Conditions 3 & 4: “The girl sees her/herself jumping.”

Filler Condition: “The grandfather thinks that he is crying.”

Appendix 2. List of stimuli by condition

Condition	Stimuli	English Translation
Transitive Pronoun	Мальчик погладил его.	The boy patted him.
	Бабушка пощекотала её.	The grandmother tickled her.
	Девочка ущипнула её.	The girl pinched her.
	Дедушка потрогал его.	The grandfather touched him.
	Мальчик сфотографировал его.	The boy took a picture of him.
	Бабушка нарисовала её.	The grandmother drew her.
Transitive reflexive	Бабушка нарисовала себя.	The grandmother drew herself.
	Мальчик погладил себя.	The boy patted himself.
	Мальчик сфотографировал себя.	The boy took a picture of himself.
	Бабушка пощекотала себя.	The grandmother tickled herself.
	Дедушка потрогал себя.	The grandfather touched himself.
	Девочка ущипнула себя.	The girl pinched herself.
ECM pronoun	Мальчик видит его кривляющимся.	The boy sees him making faces.
	Дедушка видит его плачущим.	The grandfather sees him crying.
	Бабушка видит её смеющейся.	The grandmother sees her laughing.
	Девочка видит её прыгающей.	The girl sees her jumping.
	Бабушка видит её сидящей.	The grandmother sees her sitting down.
	Мальчик видит его танцующим.	The boy sees him dancing.
ECM reflexive	Мальчик видит себя танцующим.	The boy sees himself dancing.
	Бабушка видит себя сидящей.	The grandmother sees herself sitting down.
	Мальчик видит себя кривляющимся.	The boy sees himself making faces.
	Девочка видит себя прыгающей.	The girl sees herself jumping.
	Бабушка видит себя смеющейся.	The grandmother sees herself laughing.
	Дедушка видит себя плачущим.	The grandfather sees himself crying.
Fillers	Девочка думает что она прыгает.	The girl thinks that she is jumping.
	Бабушка думает, что она смеётся.	The grandmother thinks that she is laughing.
	Мальчик думает что он танцует.	The boy thinks that he is dancing.
	Дедушка думает, что он плачет.	The grandfather thinks that he is crying.
	Мальчик думает, что он кривляется.	The boy thinks that he is grimacing.
	Бабушка думает, что она садится.	The grandmother thinks that she is sitting down.

Figure 1. Mean accuracy (cell means with 95% confidence intervals) for the two groups of children. The dotted line represents chance performance.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Case paradigms for 3rd person pronouns and reflexive in Russian

	Nominative	Accusative	Genitive	Dative	Instrumental	Prepositional
3-rd person pronouns (m/f/pl)	on/ona/oni	jego/jejo/ix	jego/jejo/ix	jemu/jej/im	im/cej/imi	niom/cej/nix
Reflexive	—	sebia	sebia	sebie	soboj	sebie

Table 2

The pattern of results predicted by a selective deficit associated with each syntactic principle or element

	PIP	Binding Cond.	A-chain Condition	Syntactic complexity	Morpho-syntactic features
Simple transitive pronouns	-	-	-	-	-
Simple transitive reflexives	-	-	-	-	-
ECM pronouns	-	-	-	-	-
ECM reflexives	-	-	-	-	-

Note: “ - ” indicates the type of sentences accuracy in which is expected to be affected by a deficit in the corresponding syntactic principle. “ - ” indicates the type of sentences accuracy of which is not expected to be affected by the corresponding syntactic principle.

Table 3

Average proportions of anaphoric errors in two groups of children

Condition	TD			DLD		
	<i>n</i>	<i>M (SD)</i>	<i>Z, p</i>	<i>n</i>	<i>M (SD)</i>	<i>Z, p</i>
Transitive Pronominal	11	.79 (.40)	-1.94, .046	11	.95 (.15)	-3.16, .013
Transitive Reflexive	14	.82 (.37)	-2.50, .014	15	.84 (.30)	-2.92, .024
ECM Pronominal	34	.82 (.28)	-4.22, <.001	18	.78 (.28)	-2.92, .028
ECM Reflexive	26	.55 (.45)	-.50, 1.000	16	.62 (.32)	-1.46, 1.000

Note. *Z, p* represent the results of the Wilcoxon Signed Ranks Tests for the comparisons of the proportion of anaphoric vs. the proportion of distractor errors separately for each group and condition. The reported *p*-values were adjusted using the Bonferroni correction (*p*/*k*, i.e. *p*/8).