Applied Psycholinguistics **38** (2017), 1173–1199
doi:10.1017/S0142716417000091

# The role of cognitive control in anaphor resolution in children with specific language impairment

ENIKŐ LADÁNYI
*Université Paris Descartes and Budapest University of Technology and Economics*

BENCE KAS
*Hungarian Academy of Sciences and Eötvös Loránd University*

ÁGNES LUKÁCS
*Budapest University of Technology and Economics*

Received: February 16, 2016 Accepted for publication: February 20, 2017

ADDRESS FOR CORRESPONDENCE
Enikő Ladányi, Laboratoire Psychologie de la Perception, Université Paris Descartes, Paris, France.
E-mail: eniko.ladanyi@parisdescartes.fr

ABSTRACT
We studied anaphor resolution and its relationship with cognitive control abilities in children with specific language impairment (SLI) and typically developing (TD) children. In a sentence–picture verification task assessing anaphor interpretation, the SLI group was less successful than age-matched TD peers, and displayed similar performance patterns as younger TD children in previous studies. The SLI group showed weaknesses in nonlinguistic cognitive control tasks, which were associated with anaphor interpretation results. These findings are in contrast with the view that proposes a grammar-specific deficit behind anaphor resolution problems in SLI. We suggest that anaphor interpretation in this population is delayed but not atypical, and this delay can be partly explained by weaker cognitive control abilities.

Children with specific language impairment (SLI) are characterized by a marked deficit in language development despite apparently typical intellectual, sensory, and social abilities and neurological state. This deficit may affect both expressive and receptive language in both lexical and grammatical domains. One of the central areas of difficulty in children with SLI is sentence comprehension, particularly evident in complex sentences (e.g., Bishop, Bright, James, Bishop, & van der Lely, 2000; Norbury, Bishop, & Briscoe, 2002; van der Lely, 1996, 1998; van der Lely & Harris, 1990; van der Lely & Stollwerck, 1997). Some argue that problems in sentence comprehension (alongside other language difficulties) are caused by a deficit in the innate representation of grammar (Rice, Wexler, & Redmond, 1999; van de Lely, 2005), while others claim that factors external to grammar such as weaknesses in working memory (Ellis Weismer & Thordardottir, 2002; Joannisse & Seidenberg, 1998; Montgomery, 1995, 2000, 2004; Montgomery & Evans, 2009; Montgomery, Magimairaj, & Finney, 2010; Norbury et al., 2002), auditive perception (Tallal & Piercy, 1973), or the general slowing of processing (Kail, 1994; Leonard et al., 2007; Miller, Kail, Leonard, & Tomblin, 2001; Windsor & Hwang, 1999) can account for the observed problems.

Sentences containing an anaphoric expression like a reflexive or a pronoun are among complex sentence types that might be problematic for children with SLI, and they are also good candidates for investigating different sources of language difficulties. The correct interpretation of anaphoric expressions depends upon another expression in context. It requires knowledge of clausal structure and syntactic dependencies, but it also builds on processing distant constituents, burdening working memory. In many cases, it requires selection among several candidates, relying on cognitive control. Thus, this domain of grammatical processing is a good focus of study for examining the contribution of linguistic and extralinguistic cognitive factors to language processing.

Although these factors make sentences with anaphoric expressions a good testing ground for disentangling different accounts of language impairment, as far as we are aware, there are only three studies that investigate anaphor interpretation in children with SLI (Montgomery & Evans, 2009; Novogrodsky & Friedman, 2010; van der Lely & Stollwerck, 1997). Van der Lely and Stollwerck (1997) investigated the interpretation of pronouns and reflexives to find out whether intrasentential reference assignment is impaired in children with SLI, and what the underlying causes of the impairment are. Using a sentence–picture verification task to test 12 English-speaking children with SLI aged between 9 years, 3 months (9;3) and 12;10 years and three language control groups, they observed significantly lower levels of performance in children with SLI than in any of the TD groups. They concluded that children with SLI show a modular language deficit concerning dependent structural relationships between syntactic constituents. Montgomery and Evans (2009)'s main aim was the investigation of the relationship between working memory and sentence comprehension. In their sentence comprehension task, they used a shorter version of van der Lely and Stollwerck (1997)'s design involving pronouns and reflexives, and they also tested comprehension of passive sentences in a sentence–picture verification task. They tested 24 English-speaking children with SLI aged between 6 and 12 years and two TD groups. They did not report results for sentences with anaphors separately, but the total scores were significantly lower in children with SLI than in age-matched TD children and similar to scores in the control group of language and memory matched TD children (see more details about the study below in the Sentence Comprehension and Working Memory in Children With SLI section).

In Novogrodsky and Friedman (2010)'s study, interpretation of pronouns and reflexives was tested together with comprehension of complex sentences involving *wh*-movement in a picture selection task in 12 Hebrew-speaking children with SLI aged between 9;3 and 13;10 and control children. Their objective was to find out whether different kinds of dependencies can be selectively impaired or the deficit involves all of them. While children in the SLI group showed impairments in the interpretation of sentences involving *wh*-movement relative to TD children, no group differences were observed in anaphor interpretation, suggesting the possibility of selective impairments.

The apparent controversy in the above findings may result from methodological differences. Impairments in anaphor interpretation were observed in the two studies relying on sentence–picture verification (Montgomery & Evans, 2009; van der Lely & Stollwerck, 1997); this measure could be more sensitive to complex sentence comprehension differences than choosing the picture that matches the sentence out of two candidates.

The two studies that found significant differences between the SLI and TD groups took different theoretical approaches and thus drew different conclusions concerning the source of difficulties in anaphor interpretation in SLI. Van der Lely and Stollwerck (1997) argue that difficulties with anaphor interpretation are a consequence of a grammar-specific deficit (in the binding module of grammar; see below). In contrast, Montgomery and Evans (2009) propose that anaphor comprehension deficiencies are secondary to reduced complex working memory capacity in children with SLI.

Our goal in the present study was to explore whether anaphor comprehension is problematic in Hungarian children with SLI, and to identify possible sources of the potential deficit in antecedent assignment. In what follows, we will review the generative linguistic account of anaphor interpretation, psycholinguistic models of anaphor comprehension based on online sentence comprehension studies, and various accounts of difficulties with anaphor processing. We will also discuss previous work on the associations between working memory and sentence comprehension in SLI. We then clarify the concept of cognitive control, and go on to discuss the potential role of cognitive control in anaphor interpretation in children with SLI.

## SYNTACTIC FACTORS IN ANAPHOR INTERPRETATION

Within generative grammar, the interpretation of anaphors is accounted for by binding theory (Chomsky, 1981, 1986). In binding theory, there are two main syntactic principles (Principles A and B) that constrain potential antecedents for reflexives and pronouns. Binding Principle A states that a reflexive must be bound in its governing category, which means that the reflexive has to have a c-commanding antecedent within the same governing category. In contrast, pronouns cannot be bound within the same clause according to Binding Principle B but should corefer with a suitable discourse element in the context. Thus, the antecedent of the reflexive *herself* must be *Mary* in (1a), whereas the antecedent of the pronoun *her* cannot be *Mary* in (1b).

(1) a. *Jane says Mary washed herself.*
b. *Jane says Mary washed her.*

From the perspective of processing, these principles express the syntactic locality conditions of binding that presuppose the recognition of the different types of anaphors.

 https://doi.org/10.1017/S0142716417000091

Chien and Wexler (1990) and Grodzinsky and Reinhart (1993) argue that sentences with a pronoun have two logical interpretations: when the pronoun (e.g., her in (1b)) is coindexed with the subject of the clause (Mary in (1b)) and when it is not. They propose that the correct interpretation has to be computed with the help of a pragmatic principle (see more details in the next section) that makes pronouns more difficult to interpret than reflexives. However, according to Chien and Wexler (1990) if the antecedent is a quantifier, only one interpretation is possible, because coindexation with a quantifier is only possible with a variable bound by (i.e., being in the same clause with) that quantifier. For this reason, in these cases the processing of the pronoun will not be more difficult than the processing of reflexives.

## PROCESSING ANAPHORS IN SENTENCES

An increasing number of studies investigate when and how the abovementioned syntactic constraints take effect during anaphor processing, relying on online sentence comprehension methods (priming, self-paced reading, eye-tracking during reading, and eye-tracking during listening), mainly in adults. Some of the findings suggest that the adult parser only considers syntactically appropriate antecedents (those which are in accordance with Principle A and B) but not structurally inappropriate ones (e.g., Nicol & Swinney, 1989). This view is usually called the *early filter hypothesis*.

Other online sentence comprehension studies however argue for the *multiple constraints approach* (e.g., Badecker & Straub, 2002), proposing that syntactic and discourse factors affect anaphor processing simultaneously. According to the latter view during the course of sentence processing, several potential antecedents are activated in the discourse context beyond the syntactically appropriate antecedent, and the inappropriate one(s) is/are ruled out later during processing. In the multiple constraints approach, this process controls the interpretation of both pronouns and reflexives, but empirical evidence on the activation of multiple potential antecedents is only available for pronouns in the majority of studies (e.g., Runner, Sussman, & Tanenhaus, 2003). Results about reflexives are less consistent. Some studies have found no evidence for the activation of syntactically incorrect potential antecedents in the case of reflexives (cross-modal priming: Nicol, 1988; Nicol & Swinney, 1989; eye tracking during listening: Clackson, Felser, & Clahsen, 2011; event-related potential: Xiang, Dillon, & Phillips, 2009), while other results support the activation of multiple potential antecedents for reflexives as well (self-paced reading: Badecker & Straub, 2002; eye-tracking during reading: Cunnings & Felser, 2013; eye-tracking during listening: Clackson & Heyer, 2014). Although Clackson and Heyer (2014) also emphasize that the effect is stronger in the case of pronouns, they argue that the lack of evidence for multiple potential antecedents during reflexive processing in other studies is due to a methodological flaw. Overall, online sentence comprehension studies suggest that when there is a prominent potential antecedent in the discourse context during the processing of reflexives and pronouns, the processing system does not rule it out immediately even if it is inappropriate based on binding principles. This effect seems to be stronger in the case of pronouns, especially when the other potential antecedent(s) is/are not quantified noun(s).

Most of the above online studies tested adults; offline studies suggest changes in anaphor comprehension during development. The results show that while children are already quite successful at interpreting reflexives at age 3, they make mistakes with pronouns until age 6 (Chien & Wexler, 1990; Guasti, 2002; Perovic, Modyanova, & Wexler, 2013; Rákosi & Tóth, 2016), but if the competing antecedent is a quantified noun, the difference between pronouns and reflexives decreases (Chien & Wexler, 1990). Clackson et al. (2011) investigated developmental changes in anaphor resolution with an eye-movement monitoring during listening paradigm in children (6–9 years) and in adults. Participants were auditorily presented with two-sentence paragraphs with either reflexives or pronouns at the end of the second sentences. Both the first and the second sentence contained an animate character. In the double match condition, the gender of the two characters matched each other and the gender of the anaphor as well (2a). In the single match condition, the anaphor only matched its antecedent (the proper noun of the first sentence in the case of pronouns and the proper noun of the second sentence in the case of reflexives) in gender (2b).

(2) a. Double-match
*Mr. Jones was listening very hard. He knew that Peter was playing some classical music to himself/him on the new piano.*
b. Single-match
*Susan was listening very hard. She knew that Peter was playing some classical music to himself/her on the new piano.*

While participants were hearing the sentences, a visual display with four pictures was presented. Two pictures depicted the two animate characters mentioned in the sentences: one depicted an inanimate object that was also mentioned in the second sentence, and the fourth picture depicted a distractor inanimate object not mentioned in the sentences. Looking times for the pictures were measured during the processing window of the anaphoric expression. Both children and adults were temporarily distracted by the competitor antecedent in the double match condition in the case of pronouns shown by longer looking times to the reference of the competing antecedent in the double than in the single match condition, but the effect was significantly higher in children. During the interpretation of reflexives, only children looked significantly longer at the competing antecedent in the double than in the single match condition (but see Clackson & Heyer, 2014). This suggests that if a noun phrase is not a potential antecedent based on binding principles, but is supported by other cues (gender match and recency/primacy in this case), it distracts children by creating competition between antecedents supported by different cues in the case of both pronouns and reflexives. These results show that although the final interpretation of reflexives is adultlike in children, their processing is different when binding principles and discourse prominence provide conflicting cues.

Taken together, the above results show that multiple potential antecedents have additional processing costs both in children and in adults, but it is not yet clear what mechanism plays a role in selecting the correct one and why young children have more difficulties than adults. Earlier studies emphasized the contribution of a pragmatic principle (Chien & Wexler, 1990; Grodzinsky & Reinhart, 1993) or working memory factors (Grodzinsky & Reinhart, 1993; Montgomery & Evans, 2009), while Clackson et al. (2011) attributed an important role to cognitive control. Chien and Wexler (1990) argue that a pragmatic principle (Principle P) is necessary for ruling out coreference between the pronoun and an NP in its local domain, and thus for identifying the antecedent of a pronoun. They claim that while syntactic principles like Principle A and B are innate, pragmatic principles have to be learned during the course of development. This proposal is in line with the model of syntactic acquisition in the principles and parameters theory (Guasti, 2002). Grodzinsky and Reinhart (1993) also argue that the operation of, in their case, an innate-pragmatic principle, which they call Rule I, is necessary for ruling out inappropriate coreference in the case of pronouns. They emphasize the importance of general processing abilities, more specifically of working memory, in the use of the pragmatic principle. Both theories were primarily formulated to account for developmental differences between the interpretation of pronouns and reflexives, but their conclusions are considered to be true for anaphor interpretation in general. Recent studies investigating the role of working memory with targeted working memory (Montgomery & Evans, 2009) and cognitive control (Clackson et al., 2011; Clackson & Heyer, 2014) tests also support the role of these abilities in anaphor interpretation. In the following section, we review studies that investigated the relationship between working memory and sentence comprehension in SLI, including Montgomery and Evans's (2009) study, which aimed to test the relationship specifically with sentences with anaphors. Before discussing the association with cognitive control, we would like to make some clarifications about the use of the concept and its relationship with working memory in the Working Memory, Executive Functions, and Cognitive Control Section.

## SENTENCE COMPREHENSION AND WORKING MEMORY IN SLI

The importance of working memory, that is, our ability to store and manipulate information simultaneously, in sentence comprehension has been widely documented in adults (Just & Carpenter, 1992; Miyake, Carpenter, & Just, 1994; Waters & Caplan, 1996). Although the limitations of working memory, especially in the verbal domain, are among the proposed core deficits in children with SLI (e.g., Archibald & Gathercole, 2006; Ellis Weismer, Evans, & Hesketh, 1999; Gathercole & Baddeley, 1990; Hesketh & Conti-Ramsden, 2013; Marton, Kelmenson, & Pinkhasova, 2007; Marton & Schwartz, 2003), we found only a limited number of studies specifically investigating the relationship between sentence comprehension and working memory in language impairment. Montgomery (1995, 2004) tested the relationship between phonological working memory measured by a nonword repetition task and sentence comprehension in 8;2 (1995) and 8;9 (2004) aged children with SLI and in their TD peers. Sentence comprehension was tested with auditory sentences in a picture selection task. A positive correlation was found between nonword repetition and sentence comprehension performance in the whole group of children in Montgomery (1995), but there was no significant correlation between the two measures in Montgomery (2004). More relevant to our study, Montgomery and Evans (2009) investigated the role of working memory in the comprehension of complex sentences with a passive structure, and with reflexives and with pronouns. Beside nonword repetition span, attentional resource capacity/allocation was also measured by the competing language processing task (CLPT; Gaulin & Campbell, 1994). The CLPT is a listening span task, where participants listen to increasingly larger sets of sentences. They have to decide if the sentence is true or not after each sentence, and after each set they have to recall the last word of each of the sentences in the actual set. In the SLI group, performance on the CLPT task was associated with performance on the complex sentence comprehension task, while in age-matched TD peers, the correlation was not significant.

## WORKING MEMORY, EXECUTIVE FUNCTIONS, AND COGNITIVE CONTROL

The CLPT task and its adult version, the listening span task, are frequently used measures of the processing part of working memory. Most working memory models differentiate between a passive storage component and an active processing component (central executive in the multicomponent model, Baddeley & Hitch, 1974; focus of attention in the embedded process model, Cowan, 1995, 1999; and executive attention in the executive attention view, Engle, 2002). These components are suggested to be used when the contents of short-term memory have to be manipulated; they maintain goal-relevant information in a highly active accessible state under conditions of interference. Similar abilities began to be recognized as an important function of the prefrontal cortex by neuropsychological and brain imaging studies; these are usually referred to as executive functions or cognitive control. Although the role of cognitive control or executive functions in various areas of cognition is the focus of many studies, there is no consensus about the definitions of these concepts and the relationships between them in the literature. In this paper, we will use the term cognitive control to refer to these controlling functions in line with the view presented in Novick, Trueswell, and Thompson-Schill (2005). Novick et al. (2005) consider cognitive control as a process responsible for the resolution of conflict or interference between contradicting representations (based on Miller & Cohen, 2001). Conflict often arises from the presence of an automatic response/stimulus characterization, which is irrelevant in the actual situation and has to be overwritten by a goal-relevant response/stimulus characterization. A typical experimental design for generation of conflict and thus for measuring cognitive control is the Stroop task (Stroop, 1935) in which color names are presented written in different ink colors (e.g., the word *green* printed in blue) and participants have to name the ink color, which required that they override the automatic response generated by the word meaning.

## COGNITIVE CONTROL, ANAPHOR RESOLUTION, AND SLI

Novick et al. (2005, 2010) suggest that cognitive control is a core process in language as well, helping resolve linguistic conflict in cases of complex sentences with structural ambiguity, in homonym processing, or in word retrieval. Cognitive control was found to be important in the comprehension of complex sentences when competing syntactic analyses are present for several structures in adults (del Río et al., 2011; January, Trueswell, & Thompson-Schill, 2009; Novick et al., 2005; 2010; Ye & Zhou, 2009), but we have not found any studies targeting the relationship between sentence comprehension and cognitive control in children with SLI. There are, however, several studies investigating cognitive control abilities in SLI motivated by the assumption that the impairment of cognitive control might contribute to language problems. As we have discussed previously, concepts referring to abilities controlling and coordinating our thoughts and actions vary within the literature, and the lack of clear conceptualization of these phenomena and their relationships with each other makes it difficult to generalize findings of specific studies. Keeping this in mind, a growing body of evidence shows problems in the abilities responsible for the resolution of conflict between competing representations by selecting relevant and inhibiting irrelevant information in children with SLI (e.g., Finneran, Francis, & Leonard 2009; Henry, Messer, & Nash, 2012; Im-Bolter, Johnson, & Pascual-Leone, 2006; Marton, 2008; Marton, Campanelli, Eichorn, Scheuer, & Yoon, 2014; Spaulding, 2010), although other studies show no difference between SLI and TD groups (Lukács, Ladányi, Fazekas, & Kemény, 2016; Noterdaeme, Amorosa, Mildenberger, Sitter, & Minow, 2001).

In the current study, our aim was to investigate the hypothesis that cognitive control is involved in anaphor resolution, and thus an impairment of cognitive control in SLI contributes to sentence processing difficulties in the case of sentences containing anaphors. This hypotheses was motivated by Clackson et al. (2011)'s study described above, and their suggestion that as cognitive control is responsible for selecting among competing representations in general, it might also be necessary for inhibiting syntactically inappropriate antecedents in anaphor resolution. They also predict asymmetries in the processing of pronouns and reflexives. As the authors argue, competition might be higher in the case of pronouns because unlike Principle A, Principle B does not determine a unique referent and its interpretation requires recourse to, and integration of, additional information sources. Therefore, it is more difficult to inhibit a semantically or pragmatically prominent but syntactically inappropriate antecedent for pronouns than it is for reflexives. This difference should be eliminated if the competing antecedent is a quantified noun. This difference explains why processing difficulties apparent in online measures are more expressed in the case of pronouns in adults. Furthermore, it can also account for developmental changes in anaphor comprehension: cognitive control develops until adolescence that explains the lack of adultlike processing of pronouns (shown both by online and offline studies) and reflexives (appearing only if investigated by online methods) in children.

## THE CURRENT STUDY

The current study investigates whether anaphor resolution is impaired in Hungarian-speaking children with SLI and whether individual differences in cognitive control contribute to differences in anaphor resolution performance in SLI and in typical development. Our questions were motivated by several lines of research. First, Montgomery and Evans (2009) found a correlation between complex working memory and anaphor interpretation in children with SLI, and recent studies suggest that cognitive control might be necessary for performing complex working memory tasks (such as listening span tasks). Second, Clackson et al. (2011) argued that cognitive control abilities are crucial for anaphor processing. Third, several studies found reduced efficiency of cognitive control in children with SLI (e.g., Finneran et al., 2009; Henry et al., 2012; Im-Bolter et al., 2006; Marton, 2008; Marton et al., 2014; Spaulding, 2010).

Motivated by the above lines of research, our first aim was to investigate whether Hungarian primary school children with SLI show differences in anaphor comprehension compared to their typically developing peers. Our second aim was to find out whether children with SLI show impairments in cognitive control tasks compared to TD children. The third focus of our research was on the relationship between anaphor comprehension and cognitive control abilities. Our fourth aim was to investigate the differences between the relationship of cognitive control and pronoun processing and cognitive control and reflexive processing.

Considering cross-linguistic factors, it seems that the syntactic properties of anaphoric structures in Hungarian are very similar to those in English. Reflexives and pronouns have distinct, well-identifiable lexical paradigms, and their patterns of dependency follow Principle A and Principle B, respectively. Initial studies on the acquisition of pronoun interpretation in Hungarian revealed an asymmetry between reflexive and pronominal comprehension in preschool children similar to the above-cited findings in English (Rákosi & Tóth, 2016). Thus, expecting the same asymmetry in the comprehension of reflexives and pronouns in school-age children with SLI might be considered as a null hypothesis based on our earlier results on SLI in Hungarian. These studies showed that the pattern of language impairment in Hungarian SLI can be characterized as a general language weakness with a significant decrease in complex structures and vocabulary. Hungarian children with SLI show only a few additional problems in the area of grammar such as the processing of atypical word order patterns, verbal suffixes, and lexical case marking. However, these weaknesses have been explained by factors external to grammar, such as verbal working memory effects, low frequency of occurrence, and phonological complexity (Kas, Lukács, & Szentkuti-Kiss, 2016; Lukács, Kas, & Leonard, 2013; Lukács, Leonard, Kas, & Pléh, 2009). Therefore, we expect school-age Hungarian children with SLI to perform similarly to typically developing preschool children in the interpretation of reflexives and pronouns, showing an asymmetry preferring the former.

Anaphor comprehension was tested by an adaptation of van der Lely and Stollwerck (1997)'s sentence comprehension task. Cognitive control was tested by three tasks: a backward digit span task, which is frequently used to measure complex working memory; an *n*-back task, which requires the updating of relevant information in working memory and is often used for measuring cognitive control; and a modified version of the Stroop task (Stroop, 1935), which is one of the most prevalent cognitive control tasks in the literature and measures cognitive control without the storage component. The *n*-back task shows similarities with anaphor comprehension (a previously seen/heard element has to be selected for processing the actual element), but does not involve the sentence-processing confound of the reading span/listening span/CLPT task used in previous studies as a cognitive control measure. Both the *n*-back and the backward digit span tasks contain storage and processing components. Since storage is also necessary for sentence comprehension, the storage component of a complex task can in itself be responsible for associations between the cognitive control and sentence comprehension measures. As our main focus was on the role of the cognitive control component, we assessed storage abilities with a nonword repetition task to be able to control for the effect of storage capacity during the analysis.

Based on previous findings we expected the following:

1. Children with SLI would have problems with anaphor interpretation, and this difficulty would be especially prominent in the case of pronouns and sentences without a quantifier.
2. Children with SLI would show cognitive control impairments in all of the three cognitive control tasks.
3. Difficulties with anaphor comprehension would be associated with deficits in cognitive control in children with SLI shown by significant correlations with the Stroop task, and with the *n*-back and backward digit span tasks even when storage capacity is controlled for.
4. Performance on pronouns would show a stronger correlation with cognitive control measures than performance on reflexives.

To test these hypotheses, we examined anaphor interpretation abilities and cognitive control functions in children with SLI and age-matched typically developing peers.

## METHODS

### *Participants*

Sixty children participated in our study.[1] The SLI group consisted of 30 Hungarian-speaking children (8 girls, 22 boys) who were selected from two special schools for children with language impairments. Their mean age was 8.93 years with a standard deviation of 1.18 years. Only children with normal hearing and no history of neurological impairments were included. All participants' IQ was in the normal range (above 85 scores on Raven's Coloured Progressive Matrices; Raven, Court, & Raven, 1987). Children meeting the above criteria were screened further for inclusion in the SLI group based on criteria that are commonly used in SLI research (see, e.g., Leonard 2014/1998, Tager-Flusberg & Cooper, 1999). Linguistic abilities were assessed with four tests, and children who performed at least 1.5 *SD* below age norms on at least two out of the four tests were included in the SLI group. These four tests included two receptive and two expressive tests. The receptive tests were the Hungarian versions of the Peabody Picture Vocabulary Test (Peabody Képes Szókincsteszt; Csányi, 1974) and the Test for Reception of Grammar (Bishop, 1983; Nyelvtani Szerkezetek Megértése Teszt; Lukács, Győri, & Rózsa, 2012). The expressive tests were the Hungarian Sentence

Repetition Test (Magyar Mondatutánmondási Teszt; Kas & Lukács, 2016), and a nonword repetition test (Álszóismétlési teszt; Racsmány, Lukács, Németh, & Pléh, 2005).

Typically developing children were matched individually to children in the SLI group on chronological age and sex, and they were matched groupwise on nonverbal IQ (Raven et al., 1987). Demographic and screening data for the two groups are shown in Table 1. All children were tested with the informed consent of their parents, in accordance with the principles set out in the Declaration of Helsinki and the stipulations of the local institutional review board.

### *Design and procedure*

*Anaphor interpretation.* Our anaphor interpretation task was based on van der Lely and Stollwerck's study (1997), more specifically on their Experiment 2. A yes/no sentence–picture judgment task was used in which children saw a picture while they heard a sentence presented by the experimenter. The picture either matched the sentence (sentence: *Róbert Gida azt mondja, hogy Nyuszi lelocsolja őt.* "Christopher Robin says that Rabbit sprinkles him"; picture: Rabbit sprinkles Christopher Robin) or was a mismatch (same sentence, picture: Christopher Robin sprinkles Rabbit). There were four types of experimental sentences and one control sentence type. Every sentence had the same structure with a main clause of the form *X azt mondja, hogy* . . . "X says that . . ." followed by a subordinate clause of the form *Y Z-olja W-t* "Y is Z-ing W," where X, Y, and W were cartoon characters familiar to the children, and Z was always a transitive verb. In the subordinate clauses, following van der Lely and Stollwerck (1997), half of the experimental conditions had a referential definite NP and half had a quantified NP in subject position.[2] The object NPs were reflexives in half of the sentences and pronouns in the other half. These factors add up to four experimental sentence types: no quantifier–pronoun; quantifier–pronoun; no quantifier–reflexive; and quantifier–reflexive (see Examples (3a–d)). We also had a control condition in which both the subject and the object were definite referential NPs, but we did not include them in the analysis.

(3) a. No quantifier, pronoun

| | | | | |
|---|---|---|---|---|
| *Róbert Gida* | *azt* | *mondja,* | *hogy* | *Nyuszi* |
| Christopher Robin | that-ACC | say-PRS.3SG | that | Rabbit |
| *lelocsolja* | *őt.* | | | |
| sprinkle-PRS.3SG | s/he-ACC. | | | |

"Christopher Robin says that Rabbit sprinkles him."

b. Quantifier, pronoun

| | | | | |
|---|---|---|---|---|
| *Róbert Gida* | *azt* | *mondja,* | *hogy* | *minden* |
| Christopher Robin | that-ACC | say-PRS.3SG | that | every |
| *Nyuszi* | *lelocsolja* | *őt.* | | |
| rabbit | sprinkle-PRS.3SG | s/he-ACC. | | |

"Christopher Robin says that every rabbit sprinkles him."

c. No quantifier, reflexive

| | | | | |
|---|---|---|---|---|
| *Róbert Gida* | *azt* | *mondja,* | *hogy* | *Nyuszi* |
| Christopher Robin | that-ACC | say-PRS.3SG | that | Rabbit |

Table 1. *Demographic and screening data of the SLI and TD groups*

| | TD | | | | SLI | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | *SD* | Min | Max | Mean | *SD* | Min | Max | *F* | Signif. | $\eta_p^2$ |
| Age | 8.93 | 1.18 | 7.08 | 11.33 | 8.95 | 1.18 | 7.08 | 11.25 | 0.008 | .928 | 0.000 |
| Stand. scores | | | | | | | | | | | |
| Raven IQ | 106.27 | 9.94 | 85 | 125 | 101.9 | 10.09 | 85 | 130 | 2.851 | .097 | 0.47 |
| Raw scores | | | | | | | | | | | |
| Nonword repetition | 6.33 | 0.96 | 4 | 8 | 3.43 | 1.17 | 0 | 5 | 110.803 | .000 | 0.656 |
| PPVT | 124.033 | 12.21 | 102 | 146 | 96.73 | 18.91 | 66 | 132 | 44.131 | .000 | 0.432 |
| TROG blocks | 17.77 | 1.7 | 13 | 20 | 13.33 | 2.26 | 9 | 18 | 73.694 | .000 | 0.560 |
| Sentence repetition | 36.83 | 4.64 | 16 | 40 | 20.57 | 8.24 | 0 | 36 | 88.761 | .000 | 0.605 |

*Note:* TD, Typically developing; SLI, specific language impairments; PPVT, Peabody Picture Vocabulary Test; TROG, Test of Reception of Grammar.

*lelocsolja* *magát.*
sprinkle-PRS.3SG self-1SG-ACC.
"Christopher Robin says that Rabbit sprinkles himself."

d. Quantifier, reflexive
*Róbert Gida* *azt* *mondja,* *hogy* *minden*
that-ACC Christopher Robin say-PRS.3SG that every
*nyuszi* *lelocsolja* *magát.*
Rabbit sprinkle-PRS.3SG self-1SG-ACC
"Christopher Robin says that every rabbit sprinkles himself."

Each experimental sentence was presented three times, once with a matching picture (match condition) and twice with nonmatching pictures (mismatch conditions). There were two types of mismatch conditions, one involving an antecedent error and another involving an agent error. In the *antecedent mismatch* condition, the agent in the picture matched the agent (subject) of the subordinate clause, but the patient in the picture was different from the patient (object) in the sentence. If the object was a pronoun, the patient in the picture was the subject character of the subordinate clause (as would follow from an incorrect reflexive interpretation of the pronoun). Similarly, if the object was a reflexive, the patient in the picture was the main clause subject character (in line with an incorrect pronoun interpretation).

In the *agent mismatch* condition, the agent in the picture was incorrectly the subject of the main clause (instead of the subject of the subordinate clause). There was also a patient mismatch: if the object was a pronoun, the patient in the picture was the character expressed by the subject of the subordinate clause (incorrect reflexive interpretation). Similarly, if the object was a reflexive, the patient in the picture was the main clause subject (incorrect pronoun interpretation; Table 2 illustrates the different sentence types and conditions). There were 12 (4 sentence types $\times$ 3 picture match types) experimental conditions and 2 control conditions (1 sentence type with a match and a mismatch condition), yielding 14 conditions altogether. There were 6 sentences with 6 different action verbs in each condition, yielding $14 \times 6 = 84$ test sentences altogether.

Children were informed that they were going to see pictures and hear sentences, and their task was to decide if the character is telling the truth about the picture or not. Before the experiment children were asked to name the characters to be presented during the experiment, and they were reminded of the names they did not know. This task as well as the following cognitive control tasks were programmed and presented with the E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2012). Sentences were read by the experimenter after the picture appeared. The answers were coded by the experimenter by pressing a button on the keyboard based on the answer of the child. Pictures were presented in two blocks with a little break in the middle. Testing took approximately 20–25 min per child.

*Backward digit span task.* In the backward digit span task, children were presented with sequences of numbers auditorily and they had to repeat them in a reversed order. Sequences of different lengths were presented, and each length was associated with four items. Sequences were presented in increasing length,

Table 2. *Examples of sentence types and scenarios presented in pictures in different conditions*

| | | Match | Antecedent Mismatch | Agent Mismatch |
|---|---|---|---|---|
| Pronoun, no quantifier | ***Róbert Gida azt mondja, hogy Nyuszi lelocsolja őt.*** "Christopher Robin says that Rabbit sprinkles him." | Rabbit sprinkles Christopher Robin. | Rabbit sprinkles himself. | Christopher Robin sprinkles Rabbit. |
| Pronoun, quantifier | ***Róbert Gida azt mondja, hogy minden nyuszi lelocsolja őt.*** "Christopher Robin says that every rabbit sprinkles him." | Every rabbit sprinkles Christopher Robin | Every rabbit sprinkles himself. | Christopher Robin sprinkles every rabbit. |
| Reflexive, no quantifier | ***Róbert Gida azt mondja, hogy Nyuszi lelocsolja magát.*** "Christopher Robin says that Rabbit sprinkles himself." | Rabbit sprinkles himself. | Rabbit sprinkles Christopher Robin. | Christopher Robin sprinkles himself. |
| Reflexive, quantifier | ***Róbert Gida azt mondja, hogy minden nyuszi lelocsolja magát.*** "Christopher Robin says that every rabbit sprinkles himself." | Every rabbit sprinkles himself. | Every rabbit sprinkles Christopher Robin. | Christopher Robin sprinkles himself. |

and the child had to repeat at least two out of four items to proceed to the next level of length. If the participant made three errors in one block, testing was terminated, and the span of the participant was established as sequence-length of the block before the last, that is, the maximum length that was completed. The span established this way was the measure we used in the analysis of results. Testing started with two items and the longest possible sequence contained nine items. The task lasted for 3–4 min.

*N-back task.* During the *n*-back task, participants were presented with a sequence of letters on the computer screen, and their task was to indicate (by pressing "ENTER") when the current letter matched the one presented "*n*" steps earlier.

We used one and two back conditions, in two blocks with about a minute break between them. Each block consisted of 60 trials, from which 10 were *n*-back trials (i.e., stimuli that match the ones presented "*n*" before), which appeared pseudorandomly within the blocks. We calculated the discrimination index ($P_r$) in the two-back condition for each child, which is the difference between the number of hits (when the participant correctly pressed "ENTER" on an "*n*-back trial," i.e., when the current item was identical to the target item, with a maximum of 10 hits per block) and the number of false alarms (the participant pressed "ENTER" on a not "*n*-back trial," i.e., the actual stimulus was not identical to the one presented "*n*" before) and used this score in the analyses. The task took about 10 min to administer.

*Stroop task.* During the Stroop task the child was sitting in front of the computer screen wearing headphones. Pictures of animals appeared on the screen with simultaneous animal names presented auditorily through the headphones. Stimuli consisted of four animal pictures (cow, cat, cock, and horse) and their names. In the congruent condition the auditory name matched the picture (e.g., a cow appears and the word *cow* is heard); in the incongruent condition they did not match (e.g., a cow appears and the word *horse* is heard); and in the control condition only the picture appeared without an auditory stimulus. Conditions appeared in blocks of 60 trials. Blocks as well as stimuli within the blocks appeared in a randomized order. The first block was preceded by an instruction and a short practice session. Children were asked to press a button corresponding to the picture they see on a special keyboard with pictures of the animals. Reaction times necessary for pressing a button were collected. We calculated the difference between reaction times in the incongruent and the control conditions as the measure of cognitive control.

*Nonword repetition task.* During the nonword repetition task children were asked to repeat auditorily presented nonsense words that followed the phonotactic rules of Hungarian. We used the Hungarian nonword repetition task (Racsmány et al., 2005), which contains words with one to nine syllables, each length associated with four items. Children had to repeat increasingly longer items, and they could proceed to the next level only if they could repeat at least two out of the four items. The child's span, the last length level of which s/he could repeat at least two items, was used during analysis. The task lasted for about 5–6 min.

### *Data analysis*

We analyzed the results of the anaphor comprehension task in a repeated-measures analysis of variance (ANOVA). Match and mismatch sentences were analyzed separately to yield more easily interpretable results. For match sentences anaphor type (pronoun vs. reflexive) and quantifier (quantifier vs. no quantifier) were used as two-level within-subject factors and group (SLI vs. TD) as a two-level between-subject factor. In the case of the mismatch sentences, the effects of mismatch type (antecedent vs. agent mismatch), anaphor type (pronoun vs. reflexive), and quantifier (quantifier vs. no quantifier) were investigated as two-level within-subject factors and group (SLI vs. TD) as a two-level between-subject factor. Significant interactions were analyzed further with paired-sampled *t* tests.

Group differences in cognitive control tasks between the SLI and TD groups were investigated with a one-way ANOVA. Relationships between anaphor comprehension performance and cognitive control tasks were examined by correlation analyses. To eliminate the effect of short-term memory, we included nonword repetition scores as a covariant in partial correlation analysis. For investigating the relative contribution of various cognitive control and short-term memory abilities to anaphor comprehension, we conducted a stepwise multiple linear regression analysis.

## RESULTS

### *Anaphor interpretation*

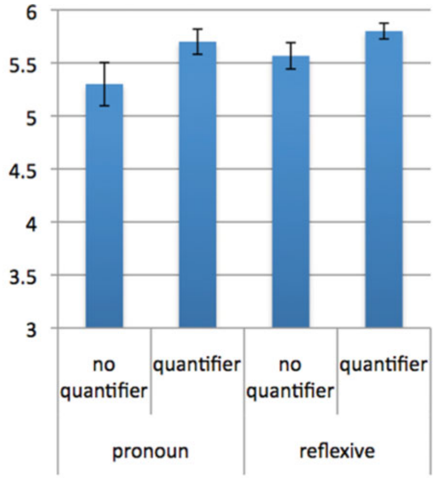
*Match sentences.* For match sentences (results are presented in Figure 1), the ANOVA showed a significant main effect of quantifier, $F$ (1, 58) = 6.231, $p$ = .015, $\eta^2 = 0.097$, demonstrating significantly better performance for sentences with a quantifier than without a quantifier and the Quantifier × Group interaction was also significant, $F$ (1, 58) = 9.509, $p$ = .003, $\eta^2 = 0.141$. No other main effects or interactions were significant (all $p$s > .05).

To further investigate the interaction we tested the effect of the quantifier factor separately in the two groups: the SLI group showed significantly better performance on sentences with a quantifier, $t$ (29) = 4.462, $p$ < 001, and in the TD group the difference was not significant, $t$ (29) = 0.720, *ns*.

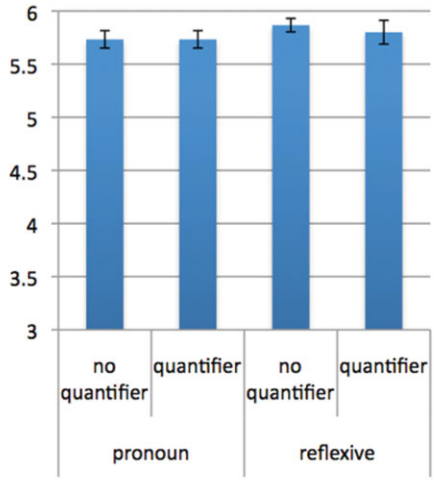
*Mismatch sentences.* For mismatch sentences (results are presented in Figure 2), the main effect of anaphor type, $F$ (1, 58) = 26.871, $p$ < .001, $\eta^2 = 0.317$, quantifier, $F$ (1, 58) = 11.325, $p$ = .001, $\eta^2 = 0.163$, and group, $F$ = 21.955, $p$ < .001, $\eta^2 = 0.275$, were significant. Two-way interactions between anaphor type and group, $F$ (1, 58) = 9.214, $p$ = .004, $\eta^2 = 0.137$, quantifier and group, $F$ (1, 58) = 4.939, $p$ = .030, $\eta^2 = 0.078$, mismatch type and anaphor type, $F$ (1, 58) = 19.750, $p$ < .001, $\eta^2 = 0.254$, mismatch type and quantifier, $F$ (1, 58) = 25.693, $p$ < .001, $\eta^2 = 0.307$, and anaphor type and quantifier, $F$ (1, 58) = 13.218, $p$ = .001, $\eta^2 = 0.186$, were also significant. The Mismatch Type × Anaphor Type × Group, $F$ (1, 58) = 5.637, $p$ = .021, $\eta^2 = 0.089$, as well as the Mismatch Type × Quantifier × Group, $F$ (1, 58) = 6.207, $p$ = .016, $\eta^2 = 0.097$, interactions also reached significance.

Interactions and main effects were analyzed further by conducting paired sampled $t$ tests. To correct for multiple comparisons, the level of alpha was divided by the number of $t$ tests conducted ($\alpha = 0.05/10 = 0.005$). We started the analysis with three-way interactions and investigated further lower order significant interactions and main effects only if they were not qualified by higher order interactions.

We broke down the Mismatch Type × Anaphor Type × Group interaction. We tested the effects of the mismatch type and anaphor type factors separately in the two groups with a 2 × 2 ANOVA. The two-way interaction was significant in both groups although the effect was smaller in the TD group, SLI: $F$ (1, 29) = 14.335, $p$ = .001, $\eta^2 = 0.331$; TD: $F$ (1, 29) = 5.659, $p$ = .024, $\eta^2 = 0.163$. To




(a)




(b)

Figure 1. (Color online) Average number of correct responses in the (a) specific language impairment group and (b) typically developing group for “match” sentences. Note that the range for correct responses was between 0 and 6 for each sentence type; but because all means were above 3, the results are presented in a range between 3 and 6 for the sake of exposition.

further analyze the interaction, we compared the effect of the anaphor type in the case of the two mismatch types separately in the two groups. In the SLI group a significantly better performance appeared on reflexives than on pronouns in the agent mismatch condition, $t\,(29) = 4.894$, $p < .001$, while there was no difference

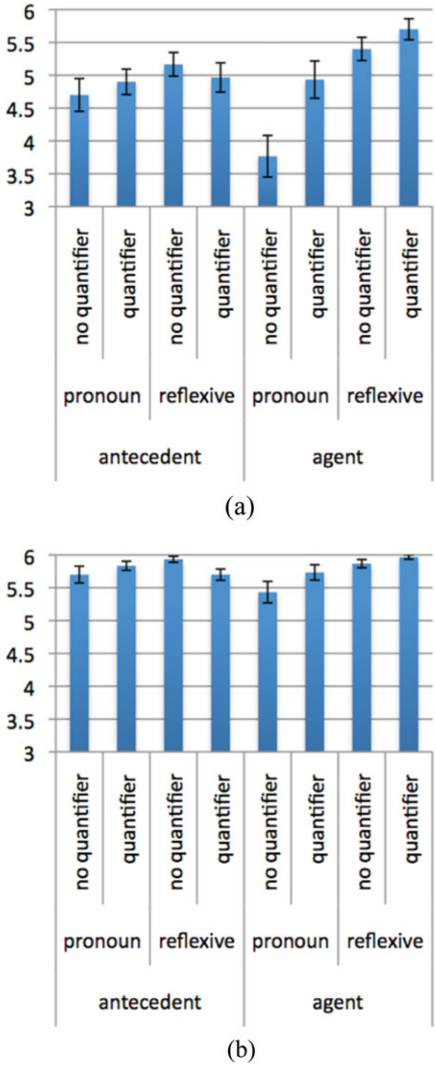



Figure 2. (Color online) Average number of correct responses in the (a) specific language impairment group and (b) typically developing group for "mismatch" sentences.

in the antecedent mismatch condition, $t$ (29) = 1.887, *ns*. In the TD group we found the same pattern, antecedent: $t$ (29) = 0.551, *ns*; agent: $t$ (29) = 2.942, $p = .006$.

For breaking down the Mismatch Type × Quantifier × Group interaction, we conducted a 2 × 2 ANOVA with mismatch type and quantifier factors in the two groups separately. The interaction was significant in both groups, but the effect was stronger in the SLI group, SLI: $F$ (1, 29) = 18.815, $p < .001$, $\eta^2 = 0.393$ TD: $F$ (1, 29) = 6.905, $p = .014$, $\eta^2 = 0.192$. To further analyze the interaction, we investigated the effect of quantifier in sentences with antecedent and agent mismatches separately in the two groups. In the SLI group there was no difference between performance on sentences with and without a quantifier when an antecedent mismatch was present, $t$ (29) = 0.000, *ns*, but a better result appeared on sentences with a quantifier in the case of agent mismatches, $t$ (29) = 5.047, $p < .001$. In the TD group the difference was not significant in either condition, antecedent: $t$ (29) = 0.551, *ns*; agent: $t$ (29) = 1.934, *ns*.

We tested the effect of the quantifier factor separately in the case of pronouns and reflexives to unpack the Anaphor Type × Quantifier interaction. Sentences with quantifiers were comprehended more successfully in the case of pronouns, $t$ (59) = 3.976, $p < .001$ *ns*, but there was no significant difference in the case of reflexives, $t$ (59) = 0.125, *ns*.

### *Cognitive control tasks*

In the backward digit span task, a significant difference was observed between the performance of the SLI (mean span = 2.41, *SE* = 0.12) and TD (mean span = 3.29, *SE* = 0.15) groups, $F$ (1, 56) = 19.372 $p < .001$. Similarly, on the *n*-backs task, children with SLI (mean score = 3.14, *SE* = 0.49) performed significantly below the TD group (mean score = 5.47, *SE* = 0.49) as shown by the one-way ANOVA, $F$ (1, 58) = 11.093, $p = .002$. There were no significant differences between the two groups on the size of the Stroop effect, $F$ (1, 59) = 2.554, *ns*; mean Stroop effect in the SLI group: 210.57 ms, *SE* = 43.79, mean Stroop effect in the TD group: 124.58 ms, *SE* = 31.26.

### *Correlation analyses*

To learn whether cognitive control abilities are associated with anaphor resolution, we tested correlations between the number of correct answers on the sentence comprehension task and performance measures on the cognitive control tasks with Pearson's bivariate correlation analysis. We found a significant correlation between backward digit span and anaphor interpretation in the SLI group ($r = .39$, $p = .038$). To control for the contribution of short-term memory span in the SLI group, short-term memory capacity was partialed out by including the nonword repetition span (which was part of the screening battery; see Table 1) as a covariant and correlation analysis was rerun. The correlation remained significant (and actually became stronger; $r = .407$, $p = .035$). In the TD group the correlation was not significant ($r = .18$, *ns*), even after controlling for short-term memory span ($r = .172$, *ns*).

To investigate potential differences between the relationship of the backward digit span with interpretation of pronouns and reflexives, we ran the correlation analysis separately with performance on the two anaphor types in the SLI group. The correlation was not significant with performance on sentences with pronouns ($r = .261$, *ns*) but reached significance in the case of reflexives ($r = .467$, $p = .011$). The pattern remained the same when a partial correlation was conducted with nonword repetition as a covariant (pronouns: $r = .280$, *ns*; reflexives: $r = .463$, $p = .015$).

Performance on the *n*-back task was strongly associated with anaphor interpretation scores in the SLI group ($r = .67$, $p < .001$) and stayed significant even after controlling for the effect of short-term memory by including nonword repetition as a covariant ($r = .714$, $p < .001$). Neither of these correlations was significant in the TD group (bivariate: $r = .36$, *ns*; partial: $r = .36$, *ns*).

Running the correlation separately with pronouns and reflexives in the SLI group showed significant correlations in both cases (pronouns: $r = .606$, $p < .001$; reflexives: $r = .622$, $p < .001$) even after partialing out short-term memory (pronouns: $r = .630$, $p < .001$; reflexives: $r = .631$, $p < .001$).

A significant negative correlation was observed between performance on the Stroop task and anaphor comprehension scores in the SLI group ($r = -.399$, $p = .029$). Higher scores on the Stroop task are associated with weaker cognitive control abilities, while higher scores on the sentence comprehension task reflect better anaphor comprehension abilities. Therefore, if we assume that weaker cognitive control leads to more difficulties in anaphor comprehension, we expect a negative correlation between the two measures, which means that the negative direction of the correlation is in line with our expectations. The correlation in the SLI group remained significant after partialing out nonword repetition scores ($r = -.423$, $p = .022$). The correlations were not significant in the TD group (bivariate: $r = -.022$, *ns*; partial: –.051, *ns*).

Testing the correlations separately in the case of pronouns and reflexives in the SLI group showed a significant correlation in the case of reflexives ($r = -.580$, $p = .001$) even after controlling for short-term memory ($r = -.585$, $p = .001$) but not in the case of pronouns ($r = -.197$, *ns*).

To determine which factors are the best predictors of performance on the anaphor comprehension task in the SLI group, we included backward digit span, *n*-back score, nonword repetition span, and Stroop scores in a multiple stepwise linear regression analysis conducted on the results of the SLI group. We used the number of correct answers on the anaphor comprehension task as a dependent variable. Nonword repetition, backward digit span, and Stroop scores were excluded by the analysis, and *n*-back score alone proved to be the best model of anaphor resolution performance explaining 46% of the variance in performance on the binding task ($R^2 = .460$, $p < .001$, standardized $b = 0.678$).

## DISCUSSION

We investigated anaphor resolution and cognitive control abilities of Hungarian primary school children with SLI and with TD. Our first aim was to find out whether Hungarian children with SLI show problems with anaphor resolution relative to their typically developing peers. Difficulties were expected in the SLI group, especially in sentences with a pronoun and without a quantifier as a result of their higher cognitive control requirements. Overall, we found that children with SLI are less successful in anaphor interpretation than TD children: while TD children performed at ceiling on almost all sentence types, significant differences appeared between conditions in the SLI group. However, most children with SLI displayed evidence of relatively good knowledge of anaphor interpretation: they performed better than chance level as a group, and even at an individual level, chance level performance was observed only in 6 children out of 30 in one or more conditions. Below we will only discuss patterns in mismatch responses in detail, since results were less diverse in the match condition due to high performance in both groups. Furthermore, as there were no significant differences between conditions in most cases in the TD group, we mainly discuss SLI patterns below, and only mention TD results where there was a difference between sentence types.

In accordance with our expectations, we found better performance on reflexives than on pronouns and on sentences with a quantifier than without a quantifier in the SLI group. This pattern is more similar to previous results of younger TD children (e.g., Chien &Wexler, 1990; Guasti, 2002; Perovic et al., 2013) than to performance patterns of children with SLI in van der Lely and Stollwerck (1997)'s study. The positive effect of quantifiers on the resolution of pronouns together with the lack of such an effect on reflexives shown by the entire group in our study was also observed in younger TD children in previous studies (e.g., Chien & Wexler, 1990). They are in contrast, however, with van der Lely and Stollwerck (1997)'s findings; in their study, the SLI group showed relatively low (chance level) performance in the quantifier-reflexive and non-quantifier-pronoun conditions and relatively good performance in nonquantifier-reflexive and quantifier-pronoun conditions leading to a very different pattern of results from ours. Furthermore, performance levels in the SLI group were generally lower than in our study. These differences can be the result of children participating in the van der Lely and Stollwerck study having a more severe impairment or the use of different selection criteria in the two studies.

The above-mentioned pattern in the SLI group (better performance on sentences with reflexives vs. pronouns and on sentences with a quantifier vs. nonquantified name) was present only in the case of the agent mismatches but not when a picture with antecedent mismatch was presented. Better performance on reflexives than on pronouns in the case of agent mismatches appeared in the TD group as well while their performance did not differ based on the presence or absence of a quantifier. A possible explanation for this is that children might revert to a simpler strategy based on linear order to interpret a sentence when it is difficult to analyze. Agent mismatches represent a scenario that matches an interpretation based on the linear strategy (in the picture for the sentence "Christopher Robin says that Rabbit sprinkles him," Christopher Robin, the first noun in the sentence, acts as an agent and sprinkles Rabbit, the second noun in the sentence). For such items, interpretations based on linear order lead to incorrectly accepting sentences with agent mismatches. Children seemed to be better at analyzing sentences with a reflexive or with a quantifier, as shown by higher rates of correct rejections of agent mismatches. Antecedent mismatches in contrast depict scenarios that are false both based on syntactic analysis and on the linear strategy, which could explain a lack of difference between easier and more difficult sentences.

A common explanation for the differences between pronouns and reflexives (weaker performance for pronouns in sentences without a quantifier but no evidence of a quantifier-based difference in the case of reflexives) is that nonsyntactic abilities (pragmatic principle, Chien & Wexler, 1990; working memory, Grodzinsky & Reinhart, 1993) also affect selecting the antecedent for pronouns when the sentence contains a nonquantified name, but processing cost is reduced if the other potential antecedent is a quantified noun. Following Clackson et al. (2011)'s explanation, we suggest that the demand for cognitive control might be an important factor among such costs: this cost is higher when several equally prominent antecedents are present, as in the case of pronouns and sentences with a nonquantified name and especially in sentences with pronouns and with no quantifier. Cognitive control develops slowly; therefore, it is not very efficient in young children (Novick et al., 2005). Furthermore, as reviewed above, some studies report deficits of cognitive control in children with SLI. Difficulties with sentences with a pronoun and without a quantifier in young children in previous studies and in children with SLI in our study can both be accounted for by participants' less efficient cognitive control causing problems in sentences with a higher cognitive control demand.

This account also found support in our findings on differences in cognitive control between children with SLI and TD. Based on previous research, we expected difficulties in SLI on all three of our cognitive control tasks, and our results showed lower performance in the SLI than in the TD group on two: children with SLI showed lower performance than the TD group on the *n*-back and on the backward digit span tasks, but not on the Stroop task.

To test the hypothesis that low cognitive control abilities contribute to difficulties with anaphor interpretation in children with SLI, we explored the relationship between cognitive control and anaphor resolution abilities. Performance on the backward digit span and the *n*-back task was strongly associated with anaphor interpretation performance in the SLI but not in the TD group. The correlations remained significant after controlling for the contribution of short-term memory span, which shows that short-term memory, which is a common component of the digit backword/*n*-back and the sentence comprehension tasks, did not entirely account for these associations, and that differences in cognitive control also contribute to differences in anaphor resolution.

Primarily, we were interested in associations between anaphor comprehension scores and those cognitive control scores in which the SLI group showed a weaker performance than the TD group. Nevertheless, we also observed a significant correlation between Stroop scores (on which the SLI and TD groups did not differ) and anaphor comprehension scores in the SLI group. This suggests that although the abilities involved in the Stroop task are not impaired in children with SLI, the ability of handling interference in the Stroop task is generally related to the comprehension of sentences with anaphors. This relationship was not observed in the TD group, probably due to the low variance in the scores.

Among the cognitive control tasks, *n*-back performance alone was the best predictor of anaphor interpretation performance (as shown by regression analysis), and backward digit span and nonword repetition scores (involving functions that are also at least partially subsumed under the *n*-back task) did not contribute further to explaining variance in performance. This suggests that updating is the most critical cognitive control function involved in anaphor interpretation.

We also aimed to investigate the differences in cognitive control involvement in the interpretation of pronouns versus reflexives. Clackson et al. (2011) suggested greater involvement of cognitive control in pronoun processing; this prediction was not borne out by our findings. In the case of the *n*-back task, a significant correlation appeared with both anaphor types. Backward digit span was associated with performance on reflexives but not on pronouns. As far as we know, this was the first attempt to test differential involvement of cognitive control in reflexive versus pronoun resolution, but we need further studies to investigate differences in cognitive control requirement of different anaphor types.

Taken together, similar patterns of anaphor interpretation performance in children with SLI in our study and in younger TD children in previous studies and associations between anaphor interpretation and cognitive control suggest that weaker performance on anaphor resolution in SLI is not a result of a specific grammatical deficit. The different patterns of results between groups in van der Lely and Stollwerck (1997), namely, that children with SLI show lower performance on sentences with a quantifier and a reflexive as well as with a nonquantified name and a pronoun, were a critical argument in their grammar-specific account for anaphor interpretation difficulties in children with SLI. There was no evidence of such a pattern in our findings, and our results are more in line with the idea that the development of anaphor resolution is very similar in children with SLI and in TD. Overall, results from our study and previous ones suggest that the observed delay in SLI is mainly due to limitations in processing resources, and can be partly attributed to the delay in the development of cognitive control abilities. However, as results on anaphor interpretation abilities in SLI are scarce, more studies are needed to establish the factors contributing to the deficit.

## ACKNOWLEDGMENTS

This research was supported by Research Grant OTKA K 83619 from the Hungarian National Science Foundation (to Á.L., Principal Investigator).

## NOTES

1. One child did not complete the *n*-back task, and three children's scores were missing for the backward digit span task.
2. Note that the subject of the subordinate clause was a proper noun in the nonquantified condition and a common noun in the quantified condition.

## REFERENCES

Archibald, L., & Gathercole, S. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders, 41*, 675–693.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York: Academic Press.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*, 748–769.

Bishop, D. V. M. (1983). *Test for Reception of Grammar*. Manchester: Medical Research Council.

Bishop, D. V. M., Bright, P., James, C., Bishop, S. J., & van der Lely, H. K. J. (2000). Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics, 21*, 159–181.

Chien, Y.-C. & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition, 1*, 225–295.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language, 65*, 128–144.

Clackson, K., & Heyer, V. (2014). Reflexive anaphor resolution in spoken language comprehension: Structural constraints and beyond. *Frontiers in Psychology: Language Sciences, 5*, 904.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 62–101). Cambridge: Cambridge University Press.

Csányi, I. (1974). *Peabody Szókincs-Teszt* [Peabody Receptive Vocabulary Test]. Budapest: Bárczi Gusztáv Gyógypedagógiai Főiskola.

Cunnings, I., & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes, 28*, 188–219.

del Rió, D., Maestu´, F., Lopéz-Higes, R., Moratti, S., Gutieŕrez, R., Maestu´, C., & del-Pozo, F. (2011). Conflict and cognitive control during sentence comprehension: Recruitment of a frontal network during the processing of Spanish object-first sentences. *Neuropsychologia, 49*, 382–391.

Ellis Weismer, S., Evans, J., & Hesketh, L. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1249–1260.

Ellis Weismer, S., & Thordardottir, E. T. (2002). Cognition and language. In P. Accardo, A. Capute, & B. Rogers (Eds.), *Disorders of language development* (pp. 21–37). Timonium, MD: York Press.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19–23.

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in preschool-age children with specific language impairment. *Journal of Speech, Language and Hearing Research, 52*, 915–929.

Gathercole, S., & Baddeley, A. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336–360.

Gaulin, C., & Campbell, T. (1994). Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual and Motor Skills, 79*, 55–64.

Grodzinsky, Y., & Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry, 24*, 69–101.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Henry, L. A., Messer, D., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry, 53*, 37–45.

Hesketh, A., & Conti-Ramsden, G. (2013). Memory and language in middle childhood for individuals with a history of specific language impairment. *PLOS ONE, 8*, e56314.

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006) Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77*, 1822–1841.

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience, 21*, 2434–2444.

Joanisse, M. F., & Seidenberg, M. S. (1998). Specific language impairment in children: An impairment in grammar or processing? *Trends in Cognitive Sciences, 2*, 240–246.

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech and Hearing Research, 37*, 418–421.

Kas, B., & Lukács, Á. (2016). *Magyar Mondatutánmondási Teszt* [Hungarian Sentence Repetition Test]. Unpublished manuscript.

Kas, B., Lukács, Á., & Szentkuti-Kiss, K. (2016) A szórend és az esetjelölés szerepe specifikus nyelvi zavart mutató gyerekek mondatfeldolgozásában. In B. Kas (Ed.), *Szavad ne feledd! Tanulmányok Bánréti Zoltán tiszteletére*. Budapest: MTA Nyelvtudományi Intézet.

Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press. (Original work published 1998)

Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50*, 408–428.

Lukács, Á., Győri, M., & Rózsa, S. (2012). A TROG pszichometriai jellemzőinek magyar vizsgálata, a normák kialakítása [The psychometric analysis of Hungarian data from the TROG]. In D. V. M. Bishop (Ed.), *TROG—Test for Reception of Grammar. Handbook* (pp. 47–86). Budapest: OS Hungary.

Lukács, Á., Kas, B., & Leonard, L. B. (2013). Case marking in Hungarian children with specific language impairment. *First Language, 33*, 331–353.

Lukács, Á., Ladányi, E., Fazekas, K., & Kemény, F. (2016). Executive functions and the contribution of short-term memory span in children with specific language impairment. *Neuropsychology, 30*, 296–303.

Lukács, A., Leonard, L. B., Kas, B., & Pléh, C. (2009). The use of tense and agreement by Hungarian-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 52*, 98–117.

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language and Communication Disorders, 43*, 181–200.

Marton, K., Campanelli, L., Eichorn, N., Scheuer, J., & Yoon, J. (2014). Information processing and proactive interference in children with and without specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 57*, 106–119.

Marton, K., Kelmenson, L., & Pinkhasova, M. (2007). Inhibition control and working memory capacity in children with SLI. *Psychologia, 50*, 110–121.

Marton, K., & Schwartz, R. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46*, 1138–1153.

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*, 416–433.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11*, 671–717.

Montgomery, J. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research, 38*, 187–199.

Montgomery, J. (2000). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 293–308.

Montgomery, J. (2004). Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. *International Journal of Language and Communication Disorders, 39*, 115–134.

Montgomery, J. W., & Evans, J. L. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 52*, 269–288.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology, 19*, 78–94.

Nicol, J. (1988). *Coreference processing during sentence comprehension* (PhD thesis, Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences).

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5–19.

Norbury, C., Bishop, D. V. M., & Briscoe, J. (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics, 23*, 247–268.

Noterdaeme, M., Amorosa, H., Mildenberger, K., Sitter, S., & Minow, F. (2001). Evaluation of attention in children with autism and children with a specific language disorder. *European Journal of Child Psychology, 10*, 58–66.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective and Behavioral Neuroscience, 5*, 263–281.

Novick, J. M., Trueswell, J. C., & Thomspon-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass, 4*, 906–924.

Novogrodsky, R., & Friedmann, N. (2010). Not all dependencies are impaired in syntactic-SLI: Binding in children with a deficit in Wh-movement. In K. Franich, L. Keil, K. Iserman, & J. Chandlee (Eds.), *Proceedings of the 34th Boston University Conference on Language Development (BUCLD)*. Somerville, MA: Cascadilla Press.

Perovic, A., Modyanova, N., & Wexler, K. (2013). Comparison of grammar in neurodevelopmental disorders: The case of binding in Williams syndrome and autism with and without language impairment. *Language Acquisition, 20*, 133–154.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai [Hungarian methods for studying verbal working memory]. *Magyar Pszichológiai Szemle, 60*, 479–505.

Rákosi, G., & Tóth, E. (2016). The pronoun interpretation problem in child Hungarian. *Acta Linguistica Hungarica, 63*, 63–96.

Raven, J., Court, J. H., & Raven, J. C. (1987). *Raven's Progressive Matrices*. London: H. K. Lewis.

Rice, M. L., Wexler, K., & Redmond, S. (1999) Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 943–961.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2003). Assignment of reference to reflexives and pronouns in picture noun phrases: Evidence from eye movements. *Cognition, 89*, B1–B13.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 725–738.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language and Hearing Research, 42*, 1275–1278.

Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241*, 468–469.

van der Lely, H. K. (1996). Specifically language impaired and normally developing children: Verbal passive vs adjectival passive sentence interpretation. *Lingua, 98*, 243–272.
van der Lely, H. K. (2005). Domain-specific cognitive systems: Insight from grammatical specific language impairment. *Trends in Cognitive Sciences, 9*, 53–59.
van der Lely, H. K. (1998). SLI in children: Movement, economy and deficits in the computational syntactic system. *Language Acquisition, 72*, 161–192.
van der Lely, H. K., & Harris, M. (1990). Comprehension of reversible sentences in specifically language-impaired children. *Journal of Speech and Hearing Disorders, 55*, 101–117.
van der Lely, H. K., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition, 62*, 245–290.
Waters, G., & Caplan, D. (1996). Processing resource capacity and the comprehension of garden path sentences. *Memory and Cognition, 24*, 24–55.
Windsor, J., & Hwang, M. (1999). Testing the generalized slowing hypothesis in specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1205–1218.
Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language, 108*, 40–55.
Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience & Biobehavioral Reviews, 33*, 1168–1177.