Cognitive control and its contribution to language difficulties in children with Specific Language Impairment

PhD Thesis

Ladányi Enikő

Supervisor: Ágnes Lukács, PhD

Budapest, 2018
# Table of Contents

Acknowledgements ..................................................................................................................... 1

0. Introduction ............................................................................................................................... 2

1. Specific language impairment ..................................................................................................... 3
   1.1. Language problems in SLI ..................................................................................................... 4
       1.1.1. SLI in different languages ............................................................................................... 5
       1.1.2. SLI in Hungarian .............................................................................................................. 6
   1.2. Theoretical accounts of SLI .................................................................................................. 10
       1.2.1. Grammar-specific accounts of SLI ............................................................................... 10
       1.2.2. Domain-general accounts of SLI .................................................................................. 11

2. Cognitive control and language ................................................................................................ 15
   2.1. The concept of cognitive control .......................................................................................... 15
   2.2. Sentence comprehension and cognitive control ................................................................. 19
       2.2.1. Individual differences results about the relationship between cognitive control and sentence comprehension ........................................................................................................... 22
       2.2.2. Neuroimaging results about the relationship between cognitive control and sentence comprehension ......................................................................................................................... 24
       2.2.3. Neuropsychological results about the relationship between cognitive control and sentence comprehension ......................................................................................................................... 26
       2.2.4. Developmental results about the relationship between cognitive control and sentence comprehension ......................................................................................................................... 28
   2.3. Word production .................................................................................................................. 30
       2.3.1. Neuroimaging results about the relationship between cognitive control and word retrieval ........................................................................................................................................... 31
       2.3.2. Neuropsychological results about the relationship between cognitive control and word retrieval ........................................................................................................................................... 32
       2.3.3. Developmental results about the relationship between cognitive control and word retrieval ........................................................................................................................................... 33

3. Cognitive control in specific language impairment .................................................................. 34
   3.1. The Listening span task ...................................................................................................... 35
   3.2. Backward digit span task .................................................................................................... 36
   3.3. The Odd one out task ......................................................................................................... 36
   3.4. The N-back task ................................................................................................................ 37
   3.5. The Stroop task .................................................................................................................. 38
   3.6. Fluency tasks ...................................................................................................................... 38
   3.7. Category judgment under conflict ..................................................................................... 40
   3.8. Cognitive control skills in children with SLI: summary .................................................... 41

4. Aims and thesis points ............................................................................................................... 41

5. Studies ...................................................................................................................................... 48
   Study 1 ...................................................................................................................................... 48
   Study 2 ...................................................................................................................................... 57
   Study 3 ...................................................................................................................................... 70
   Study 4 ...................................................................................................................................... 98

6. General discussion ..................................................................................................................... 142

7. References ................................................................................................................................ 148
Acknowledgements

First of all, I would like to thank my supervisor, Ágnes Lukács that I could be her PhD student. She inspired me when I was uninspired, helped me when I was stuck, motivated me when I was unmotivated, corrected me when I was wrong, and she always praised me when I did something well. I couldn't wish for a better supervisor.

I want to thank my co-authors who contributed to the work presented in the dissertation. I will never forget our on-site testing marathons with Kata Fazekas and cannot be grateful enough to Ferenc Kemény for helping me when I was pestering him with diverse programming and statistical problems, and to Bence Kas for being so knowledgeable in the word of hardcore linguistic theories.

Many thanks to my colleagues, Lilla Zakariás for contributing to the very beginning of the research presented in the dissertation and for the inspiring discussions ever since and Berci Polner for helping me with statistics.

I would like to thank Bori Győri for her work in data collection. Special thanks to all the children, their parents and teachers who participated in our studies. Without their contribution this work would not have been possible.

The work was supported by the Research Grant OTKA K 83619 from the Hungarian National Science Foundation awarded to Ágnes Lukács.

I would also like to thank my family for their unconditional support and tolerance when I could barely tolerate myself.

Finally, thanks to Dávid for his help and loving support during the last years of my PhD.
0. Introduction

Children with specific language impairment (SLI) have language problems (e.g., problems in morpho-syntactic acquisition and later in sentence production and comprehension, as well as in word learning and in the retrieval of already acquired words) without apparent causes (e.g., hearing impairment, neurological problems or mental disability), but several studies suggest that impairments are present also in other, nonlinguistic domains (e.g., short-term memory, working memory, processing of rapidly changing or less salient verbal information, implicit learning). Recent works suggest that the impairment of cognitive control – the ability to resolve conflict between contradicting representations – is also associated with SLI but the findings are inconsistent. It is also not clear whether the impairment of non-linguistic abilities contribute to the language problems or they co-occur independently of each other. Cognitive control might arguably have a role in language processes in those cases when several contradicting representations are activated and the conflict between these representations has to be resolved. For instance, conflict can appear between highly activated competing word representations during word retrieval or between contradicting syntactic analyses of a sentence during sentence comprehension. Some studies found – both in typical and atypical populations – that cognitive control plays a role in the resolution of such conflicts, although research is limited in the topic.

Motivated by the contradictory findings about cognitive control in SLI and by the still open question of the relationship between cognitive control and linguistic processes, the dissertation aims to investigate cognitive control and its relationship with language functions in children with SLI and in typically developing children in four studies. More specifically, first, we wanted to find out whether cognitive control is weaker in children with SLI than in their typically developing peers. Second, we explored whether word retrieval and sentence comprehension is especially difficult for children with SLI when these processes require
cognitive control. Finally, we also tested the relationship between cognitive control and language abilities to see whether weaker cognitive control contributes to language difficulties in children with SLI and whether language performance is associated with cognitive control in general (both in children with SLI and in TD children).

Before presenting our studies, we will introduce SLI by summarizing research exploring their symptoms and causes of these symptoms. Since the dissertation focuses on cognitive control as a potential causal factor in language problems of SLI, the second chapter will review studies exploring the relationship between cognitive control and language in general to demonstrate which are the language processes in which cognitive control might play a role. Finally, research on cognitive control in children with SLI will be summarized to present both supporting and contradicting evidence for the impairment of cognitive control in SLI and for the hypothesis that cognitive control impairments contribute to language problems in SLI. After presenting the Theses of the dissertation, four empirical studies on which the thesis points are based will be reported. Results will be summarized and evaluated in a General discussion at the end of the dissertation.

1. Specific language impairment

Children with specific language impairment (SLI) have problems in various language areas that are not accounted for by obvious impairments in other cognitive domains or perceptual deficits, neurological disorders, emotional or social problems, environmental deprivation or intellectual disability. As a result of these language problems children with SLI are also at considerable risk for social and behavioral problems (Beitchman, Nair, Clegg, & Ferguson, & Patel, 1986; Cantwell & Baker, 1987; Paul & Cohen, 1984; Rice, Sell, & Hadley, 1991) as well as educational difficulties (Catts, 1993; Hall & Tomblin, 1978). The problem is relatively frequent; in the U.S. its prevalence is estimated to be 7% (Tomblin, Records, Buckwalter, Zhang, Smith, & O’Brien, 1997).
Although the literature still uses the term *specific* language impairment, today there is an agreement that mild impairments of motor organization and other cognitive abilities such as working memory can be present in these children. Following common practice in SLI research, the term specific language impairment will be used in the dissertation for referring to these children without committing to the view that only language is impaired in SLI. In the following section we introduce SLI by summarizing language problems in different languages, then different accounts of SLI will be reviewed.

1.1. Language problems in SLI

Although non-linguistic impairments are also present in SLI, the most affected area is clearly language. There is a remarkable amount of research investigating language abilities of children with SLI showing large heterogeneity in the symptoms appearing in SLI but some general characteristics also emerge. Children with SLI start to produce words and word combinations later than their typically developing peers (Trauner, Wulfeck, Tallal, & Hesselink, 1995) and the acquisition of verbs is especially difficult: children use only a few verbs and they replace less frequent verbs with more frequent ones (Watkins, Rice, & Moltz, 1993). The pronunciation of phonemes is also atypical, the speech of children with SLI is often only understood by their family for a while. They produce simple sentences, often with grammatical errors. Limitations in vocabulary, sentence production and comprehension abilities are still present in primary school or at later ages (although research about older children is limited). Word retrieval problems also appear manifested in word finding difficulties both in spontaneous speech and in word production tests. Overall, language production is always impaired in SLI but in some cases comprehension can stay relatively intact.
1.1.1. SLI in different languages

When researchers started to investigate SLI, studies were mainly conducted with children whose mother tongue was English. Later, however, several studies aimed to investigate children speaking languages with different grammatical features. Research targeting different languages is crucial, because the symptoms of SLI can vary depending on the mother tongue, therefore it would be impossible to describe SLI in general without these studies and they can help to understand the cause of the symptoms as well since different languages enable distinct experimental manipulations.

Studies exploring children with SLI speaking different languages show a high variance in the symptoms beyond the abovementioned general characteristics. In English, there is ample evidence that the most prevalent domain of errors is morpho-syntax in SLI. Children with SLI often omit verb inflexions: the third person singular -s and the –ed marking past tense, as well as auxiliaries and determiners, and they tend to use incorrect case marking on pronouns (e.g.: Rice, Wexler, & Cleave, 1995; Loeb & Leonard, 1991). Although the impairment of morpho-syntax is general, the symptoms can be very different depending on the features of the language. In contrast with English, Italian children with SLI do not omit verb suffixes, they rather substitute them with another form from the paradigm (e.g., dorme ‘he/she sleeps’ instead of dormono ‘they sleep’). Beyond these problems with verb inflection, the most characteristic errors are the omission of certain function words, determinants and clitics (unstressed elements functioning as pronouns right before or after the verb; e.g., Gina lo [CLIT] vede ‘Gina saw him’) in Italian (e.g., Bortolini, Caselli, & Leonard, 1997; Bortolini, Caselli, Deevy, & Leonard, 2002). In Swedish children with SLI tend to omit the possessive –s but they don’t have problems with the plural –s which is probably caused by the fact that the plural –s always appear together with a vowel – in contrast with the possessive –s which stands without a vowel making it more difficult to perceive. Similarly to
English and Italian, Swedish children with SLI have problems with determiners but in contrast with English and Italian, beyond omission errors, substitutions are also frequent. Verb inflection errors appear in Swedish as well: children with SLI sometimes use the infinitive form instead of the inflected form of the verb or present instead of past tense (Leonard, Salameh, & Hansson, 2001; Hansson & Leonard, 2003). In German, characteristic errors also involve verb inflection but are related to word order as well. In German, the verb is the second word in the sentence in structures without an auxiliary (Peter kauft [VERB-3sg] ein Geschenk ‘Peter buys a gift’). But if there is an auxiliary in the sentence it goes to the place of the finite verb and the verb is at the end of the sentence in a non-finite form (Peter will [AUX-3sg] ein Geschenk kaufen [VERB-inf]). A typical error in SLI is the production of the finite verb at the end of the sentence – at an incorrect position – often with an incorrect suffix. The omission and substitution of determiners is also frequent as well as the incorrect use of plural markers (e.g., Clahsen, 1999; Bartke, 1998).

Difficulties with morpho-syntax appear in comprehension as well. Sentence comprehension was found to be impaired in English as a result of using strategies like interpreting the animate noun/first noun in the sentence as the subject independently of the grammatical structure (Evans & MacWhinney, 1999; Evans, 2002).

In summary, the investigation of languages with different features shows that the impairment of the same domain – morpho-syntax – can be manifested by fundamentally different error types depending on the properties of the language (Lukács, Kas, & Pléh, 2014).

1.1.2. SLI in Hungarian

Since our participants’ mother tongue is Hungarian, we found it important to present previous research about Hungarian children with SLI in more detail. Hungarian differs from the languages mentioned above in several respects. It has a very rich agglutinative morphology, grammatical relations are marked usually with suffixes – several suffixes can
follow the word root, in contrast with English in which word order plays the most important role, therefore word order is relatively free (e.g., É. Kiss, 2002). Relatively free word order means that sentences with different word orders can describe the same event (*A macska szereti a halat* ‘The cat-NOM likes the fish-ACC’ vs. *A halat szereti a macska* ‘The fish-ACC likes the cat-NOM’ meaning that ‘The fish is what the cat likes’). The use of different word orders is, however, determined by context and they show slight pragmatic differences. The most neutral and most frequently used order is the subject-verb-object order, while with other word orders different parts of the sentence can be emphasized (see a detailed description of Hungarian sentence structure in É. Kiss, 2002). These features make Hungarian very different from English and also from other languages discussed above, motivating studies of Hungarian children with SLI to explore the pattern of language difficulties in a new type of language and for exploring phenomena which could not be investigated in other languages.

The first study of SLI in Hungarian (Vinkler & Pléh, 1995) described the grammatical abilities of a 8-year-old boy with SLI. The most characteristic linguistic symptoms appeared in the use of morpho-phonological alternations (pl. *zuhamyoz* instead of *zuhamyozik* ‘(s)he takes a shower’), case markings (*Összevesznak a csokit* ‘They are arguing the chocolate’ instead of *Összevesznak a csokin* ‘They are arguing about the chocolate’) and of locative suffixes and postpositions (*Hol laksz?* ‘Where do you live?’ – *A sárga ház* ‘The yellow house’ instead of *A sárga házban* ‘In the yellow house’). His vocabulary was also smaller relative to his TD peers. This first study reporting results about Hungarian children with SLI later inspired more systematic and detailed experimental studies.

A decade after this first case study, Ágnes Lukács’s research group started to conduct experimental studies about SLI investigating noun and verb morphology in the domain of production and comprehension. These studies supported Vinkler and Pléh (1995)’s finding
about the impairment of noun morphology (e.g., Lukács, Leonard, & Kas, 2010, Lukács, Kas, & Leonard, 2013). Lukács et al. (2010) took advantage of the possibility to use more than one suffixes after a word root in Hungarian (in contrast with for instance in English), and investigated whether children with SLI show difficulties when they have to produce two suffixes in one word. According to their results younger (4 to 7 years) but not older (8 to 11 years) children with SLI made significantly more errors if two suffixes follow the noun (házakat ‘the houses’ in which ház is the word root –ak is the plural suffix and –at is the accusative case marker). Interestingly, word frequency had a strong effect on the performance of the SLI group suggesting that they rely more (though not exclusively) on memorized items in the lexicon. A recent study (Kas & Józsa, 2017) showed impairments in the production of possessive structures as well. 4-6 years old children with SLI were significantly weaker at producing the correct form of the possessee (*A kutyáknak van egy finom almái ‘The dogs have a tasty apples’ instead of A kutyáknak van egy finom almája ‘The dogs have a tasty apple’).

Beyond the domain of nouns, studies showed an impairment also in verb morphology in Hungarian children with SLI. For instance Lukács, Leonard, Kas and Pléh (2009) showed that 7;6–11;2 year old children with SLI were impaired in verb agreement (A gyerekek simogat a malacot ‘Children pets the pig’ instead of A gyerekek simogatják a malacot ‘Children pet the pig’ or Tegnap építek egy tornyot ‘Yesterday I build a tower’ instead of Tegnap építettem egy tornyot ‘Yesterday I build a tower’). The group difference disappeared, though, when non-word repetition was included as a covariate in the analysis. During the non-word repetition task nonexistent but phonotactically appropriate words with increasing length have to be repeated and it is one of the most acknowledged method to measure short-term memory – the ability to keep verbal information active for processing. The result that non-word repetition scores eliminate group differences suggests that difficulties in verb
agreement can be at least partly attributed to weaker verbal short-term memory in children with SLI. According to another study (Leonard, Lukács, & Kas 2012) the production of both perfective and imperfective verbs is also impaired in a 4;10–7;2 years old group of children with SLI (Amikor lefényképezték, a tehén itta a tejet ‘When this picture was taken, the cow was drinking the milk’ instead of Amikor lefényképezték, a tehén megitta a tejet ‘When this picture was taken, the cow had drunk the milk’).

Sentence comprehension and production were also found to be impaired in Hungarian children with SLI. Since in Hungarian grammatical functions are primarily marked with morphemes, word order can be manipulated independently of grammatical functions. Kas, Lukács, & Szentkuti-Kiss (in prep) investigated the comprehension of transitive sentences with different word orders in a picture selection task and found that a younger (5-6 years) and an older (9-12) group of children with SLI showed a generally weaker sentence comprehension performance and particular difficulties appeared in the case of the object-subject order. The authors account for this by frequency differences between word orders. While object-subject order is much less frequent order of arguments in Hungarian than the subject-object order, it is easier to analyze a sentence with this structure (Kas et al., in prep). Kas & Lukács (2008) investigated the comprehension of sentences with a relative clause using an act out task, and found significantly weaker performance in children with SLI.

In summary, previous studies of Hungarian-speaking children with SLI show impairments in tasks assessing morpho-syntax, partially accounted for by an impairment in a non-linguistic domain (short-term memory). These studies did not focus on language areas where representations compete, and the role of cognitive control has not been investigated yet either. In the dissertation we aimed to expand previous research on sentence comprehension by investigating the role of cognitive control in the comprehension of sentences with anaphors – a sentence type in which cognitive control might play a role – (see Study 3, Thesis...
point 4). Furthermore we focused on word retrieval, which got less attention in research about Hungarian so far, and studied the contribution of cognitive control also to this domain of language (see Study 2 and 4, Thesis points 3 and 5).

1.2. Theoretical accounts of SLI

Several theories (partly discussed above) tried to account for language in SLI. As it was already mentioned, children with SLI constitute a very heterogeneous group. Both the degree and type of impairments have a large variability within the populations, making research on children with SLI challenging. The creation of subgroups was tried but without any clear success so far. Partly due to this difficulty, underlying causes of SLI are still unknown and various theories exist in the literature. Explanations can be divided into two main groups.

1.2.1. Grammar-specific accounts of SLI

The first set of theories proposes a selective impairment in grammar or in one aspect of grammar assuming that the rest of the cognitive system is intact (Gopnik & Crago 1991; van der Lely & Stollwerck 1997; Clahsen, 1999; Rice et al., 1995). These theories are primarily based on English, therefore their main aim is to account for the omission of verb morphemes which is the most prevalent symptom of SLI in English. The feature blindness hypothesis (Gopnik & Crago, 1991) explains this deficit with the impairment of grammatical features like tense and agreement. Rice et al. (1995) state that children with SLI stay for an extended period of time in the so-called optional infinitive stage in which the infinitive form of verbs are used as if they were grammatical in contexts in which they are not grammatical (e.g., walk instead of walked). Motivated by studies which did not show the use of bare infinitives in obligatorily suffixing languages, Wexler modified their theory (Wexler, 2003; 2011; Wexler, Schütze, & Rice, 1998; Wexler, Schaeffer, & Bol, 2004): the agreement/tense omission model, ATOM states that children with SLI stay for an extended period of time in a
developmental stage in which they can only control a single semantically uninterpretable feature: either tense or agreement with the subject. Similarly to the ATOM model, the agreement deficit hypothesis states that features related to agreement are impaired in children with SLI (Clahsen & Hansen, 1997; Clahsen, Rothweiler, Woest, & Marcus, 1992; Eisenbeiss, Bartke, & Clahsen, 2006). Van der Lely’s theories (the representational deficit for dependent relations (RDDR) and its updated version: the computational grammatical complexity hypothesis (CGC); van der Lely, 1994; 1998; van der Lely & Stollwerck, 1997) also suggests that the feature checking mechanism of generative grammar is impaired in children with SLI. The RDDR theory states that a subgroup of children with SLI – those with grammatical SLI – are unable to check features of distant relationships, thus having problems with creating syntactic dependencies and they are not always able to check the features of nouns and verbs within the syntactic structure. Therefore they show problems in agreement but also in other dependent relations like binding anaphors to their antecedents (see more about the theory and anaphors in Study 3.). According to the updated version of the theory, the CGC, SLI is characterized by a deficit in structural complexity, and the theory was extended to morphology and phonology beyond syntax (Marshall & Lely, 2007; Lely, Jones, & Marshall, 2011).

All of the abovementioned views gained supporting evidence from several studies but neither of them is able to account for either cross-linguistic differences or for all within-language symptoms of SLI. Furthermore, they do not have an explanation for non-linguistic impairments which are clearly present in SLI. Accounts attempting to explain both the linguistic and non-linguistic impairments constitute the second type of theories.

1.2.2. Domain-general accounts of SLI

According to the other group of theories more general impairments underlie the language problems of children with SLI. Some of these theories state that a general
processing deficit, a cognitive capacity limitation or the general slowing of processing is the main cause (Marchman & Bates, 1994; Leonard, 1998), while others assume that the impairment of a specific cognitive subsystem which is crucial for language acquisition (e.g., working memory in Gathercole & Baddeley (1990) or the processing of rapidly changing auditory stimuli in Tallal & Piercy (1973)) leads to the language symptoms. There are also researchers suggesting that a general learning/memory mechanism is impaired in SLI which is responsible for the language problems (e.g., the impairment of procedural memory in Ullman & Pierpont (2005)).

The first significant theory assuming a non-grammatical impairment in SLI is the **Rapid Auditory Processing Deficit Hypothesis**, which states that the core problem in SLI is the processing deficit of rapidly changing verbal or non-verbal acoustic stimuli (Tallal & Piercy, 1973; Tallal, 1976; Tallal, Stark, & Mellits, 1985). Leonard’s **Surface Hypothesis** (Leonard, 1989) also assumes that a perceptual deficit is responsible for language problems in SLI: he proposes that the development of hearing is delayed in SLI which makes the perception of less salient morphemes (as the English –ed verb suffix) difficult and as a result of this delay children with SLI do not get the sufficient amount of information to build morphological paradigms. Leonard later rephrased his theory as the **Morphological Richness Hypothesis** based on newer results from agglutinating languages (Leonard, 1998). The theory assumes that children with SLI have a limited processing capacity and a more general limitation in language learning ability. As languages differ in the ways of expressing grammatical relations, a cue which is important in one language might be less important in another. For example unlike in Hungarian, in English word order is a strong cue for marking grammatical roles in a sentence and suffixes are less important. Typically children are sensitive to cues which are less important in their language but children with SLI might ignore less important cues due to their processing limitations. For instance children with SLI
acquiring English will devote their limited resources to acquisition of word order rules and suffixes get less attention which will lead to morphological errors (e.g., omission of –ed marking past tense), while Hungarian children will show less difficulties when only one suffix has to be produced together with the word root and difficulties are expected only when suffixes are long, rare, complex, or their functions are non-transparent (see supporting results in Lukács et al., 2010). This theory gained a lot of support since it is able to account for a wide range of linguistic symptoms and cross-linguistic variance as well, but it cannot explain all of the non-linguistic symptoms appearing in SLI.

Another processing deficit theory, the Critical Mass Hypothesis of Bates and colleagues (Marchman & Bates, 1994) (and based on their view, Conti-Ramsden & Jones, 1997; Windfuhr, Faragher, & Conti-Ramsden, 2002; Conti-Ramsden, 2003), emphasizes the general slowing of processing in children with SLI. It states that knowledge of grammatical rules depends on a ‘critical mass’, a critical vocabulary size and since children with SLI are slower at acquiring words due to their general processing deficit, it takes much longer for their vocabulary to reach the critical size for grammatical generalizations to emerge. Although the theory is also able to account for some linguistic and non-linguistic results, it cannot explain differences in the vulnerability of various grammatical functions and cross-linguistic differences, neither all of the non-linguistic problems often present in SLI.

The most prevalent non-linguistic symptom of SLI is the impairment of short-term memory. The theory assuming a short-term memory impairment in children with SLI was formulated by Gathercole and Baddeley (1990), however the role of weaker verbal memory processes in language impairments was suggested previously (language impairment: Kirchner & Klatzky, 1985; dyslexia: Jorm, 1979). Gathercole and Baddeley (1990) follow the working memory (WM) model of Baddeley and Hitch (1974) who define WM as a system comprising of three subsystems: the phonological loop, the visuo-spatial sketchpad and the central
executive. Gathercole and Baddeley (1990) assume that phonological loop is impaired in children with SLI which is responsible for the temporary storage of phonological material and it can be divided to a phonological store and an articulatory rehearsal process responsible for keeping the phonological representations in the store. Verbal short-term memory impairment is a stable characteristic of SLI (e.g., Archibald & Gathercole, 2006; Ellis Weismer, Evans, & Hesketh, 1999; Gathercole & Baddeley, 1990; Hesketh & Conti-Ramsden, 2013) therefore a reduction in verbal short-term memory capacity became one of the diagnostic criteria of SLI. Some studies also investigated complex working memory in children with SLI which will be discussed in the third section of the Introduction.

A recent theory, the Procedural Deficit Hypothesis (PDH; Ullman & Pierpont, 2005), is aimed at explaining all the linguistic and non-linguistic impairments in one framework with suggesting that a general learning/memory mechanism, namely procedural memory is impaired in SLI. The hypothesis is based on Ullman’s Declarative Procedural model (DP; Ullman, 2001, 2004, 2016), which states that there are two functionally and anatomically distinct memory systems: the procedural memory is responsible for sequence and rule learning while declarative memory underlies lexical learning. The PDH states that abnormalities in the brain network underlying procedural memory lead to impairments of linguistic and non-linguistic functions that depend on it which can cause symptoms like those which appear in SLI. The PDH might be able to explain part of the heterogeneity of SLI: depending on the affected structures and the degree of impairment various linguistic and nonlinguistic problems can appear.

Domain-general theories discussed above are all supported by experimental results but neither of them is able to account alone for all the linguistic and non-linguistic symptoms appearing in SLI. Based on these findings it is very likely that several cognitive functions are impaired in SLI and the combination and degree of these impairments determines the
language development of a child together with the features of his/her mother tongue. Therefore it is important to investigate children with different mother tongues and beyond language abilities various cognitive abilities must be assessed. In accordance with these ideas, in the dissertation *cognitive control* was investigated in Hungarian children with SLI. Cognitive control is our ability which helps to resolve conflict between contradicting representations and it is a good candidate for a potentially impaired cognitive ability in SLI since there are some evidence suggesting weak cognitive control in children with SLI (see the third section) and it was also found to be important for language processes by several studies (reviewed in the second section), but only a few studies targeted the role of cognitive control in SLI so far.

### 2. Cognitive control and language

#### 2.1. The concept of cognitive control

Cognitive control is our ability which is responsible for the resolution of *conflict* or interference between contradicting representations (Novick, Trueswell, & Thompson-Schill, 2005, 2010; Miller & Cohen, 2001). Conflict appears when incompatible information is available about how to characterize a stimulus or how to respond to it. An everyday example of conflict arising is the situation when we would automatically shake hands when we introduce ourselves but the other person starts to give us a kiss as a salutation and we have to override our automatic hand shaking movement and give kisses instead. Several tasks were created to measure cognitive control ability experimentally as well. A traditional paradigm for testing conflict resolution is the Stroop task (Stroop, 1935), in which participants have to respond based on the ink color of color terms appearing on the screen. While in congruent trials the meaning of the word matches the ink color of the word (*red* written in red), in incongruent trials they don’t match (*green* written in blue). In the latter case the meaning of the word (*green*) is automatically activated and conflicts with the task of responding based on
the word’s ink color (blue). In these cases processing difficulty (mirrored by increased response times, errors) arises because participants have to override automatically generated but irrelevant representations (e.g., MacLeod, 1991). Lesion studies using the Stroop task found that participants with an injury of the frontal lobe often show severe difficulties (extremely long reaction times or many errors) in the incongruent condition of the Stroop task while their performance is typical in the congruent and control conditions (e.g., Hamilton & Martin, 2005) which suggests that frontal areas are responsible for cognitive control. Brain imaging studies using the Stroop or other cognitive control tasks support this assumption and made possible the more detailed investigation of the anatomical background of cognitive control. Within the frontal lobe, the left inferior frontal gyrus (LIFG) was found to be especially important. For instance Milham, Banich, & Barad (2003) found this area to be activated during the incongruent condition of the Stroop task compared to the control condition.

In a recent study Hsu, Jaeggi and Novick (2017) explored the role of LIFG in cognitive control in detail and they concluded that the LIFG probably plays a role in the resolution of conflict between various types of stimuli acting as a conflict-resolution “hub”, but other regions of the brain are recruited as well depending on the type of representations and on the task (see more details about the study below). Therefore, it is very likely that the LIFG plays a crucial role in cognitive control but its exact function and exact localization within the LIFG is unknown at the moment and needs further research.

Beyond the LIFG other brain areas were also proposed to contribute to cognitive control. According to Fedorenko (2014) – summarizing studies investigating the role of domain-general cognitive control in language comprehension in a review article – the following areas were found to be important: the dorsolateral prefrontal cortex (along the inferior frontal sulcus/middle frontal gyrus), parts of the insular cortex, regions along the
precentral gyrus (going inferiorly to the posterior aspects of the inferior frontal gyrus, IFG), pre-supplementary and supplementary motor areas, parts of the anterior cingulate, and regions in and around the intraparietal sulcus (e.g., Posner & Petersen, 1990; Cabeza & Nyberg, 2000; Corbetta & Shulman, 2002; Cole & Schneider, 2007; Duncan, 2010).

It is also not clear how domain-general the cognitive control processes recruited during various language processes are. Earlier studies proposed a more general process responsible for conflict resolution both between verbal and nonverbal stimuli (e.g. Novick et al., 2005). Now they suggest a more specialized process working only in the verbal domain but as we will see below some results suggest that the same process is responsible for conflict resolution between sentence analyses and nonverbal stimuli (for instance for arrows in the flanker task; see Ye and Zhou, 2009). In accordance with the shift in the view about the domain generality of cognitive control processes, at the beginning of our research we applied also non-verbal cognitive control tasks (see Study 1) but while both findings in the literature and our unpublished results supported the involvement of more specific processes, we decided to focus on the investigation of the role of verbal conflict resolution in various language processes. Therefore our tasks and also most of the tasks applied in the studies reviewed below measure conflict resolution between verbal stimuli (letters, numbers, words). Taken together, these findings do not allow any conclusions about the domain-generality of cognitive control processes responsible for conflict resolution in language. Our working hypothesis is that it is a general process at least for verbal stimuli, meaning that the same process is responsible for the resolution of conflict between syntactic structures, word representations, numbers or letters. We do not exclude, however the possibility that the same

---

1 The flanker task, however, might also require verbal conflict resolution. During the task participants are asked to press arrows according to the direction of the target arrow presented on the screen flankered by arrows pointing either to the same direction or to the reversed direction. The task is considered to be non-verbal but if participants verbalize the directions (right, left), it becomes a verbal task requiring
process is responsible for the resolution of conflict both between verbal and non-verbal representations.

Until now the dissertation used the term cognitive control to refer to the ability which resolves conflict between contradicting representations but it is important to note that studies will be mentioned which use tasks requiring cognitive control but their authors refer to these tasks as tests measuring some other construct (complex working memory, executive functions, or inhibition in most of the cases). To clarify the relationship of these constructs with cognitive control these concepts are presented shortly below.

Tasks measuring complex working memory aim to test the active processing component of working memory (called central executive in the multi-component model of Baddeley & Hitch, 1974), focus of attention in the embedded process model (Cowan, 1995, 1999), executive attention in the executive attention view (Engle, 2002) together with storage capacity. These tasks require the manipulation and organization of working memory content which might include conflict resolution (see Marton, Schwartz, Farkas and Katznelson (2006), Marton, Kelmenson, & Pinkhasova (2007) and Marton, Eichorn, & Zakariás (2016) about the role of conflict resolution in working memory tasks) therefore they are highly relevant for the topic of the dissertation.

Executive functions (EFs) also overlap with both the concept of cognitive control and complex working memory. EFs are defined as a family of top-down mental processes recruited when we have to concentrate or pay attention, when responding automatically or relying on instinct or intuition would be ill-advised, insufficient, or impossible (Diamond, 2013). Most of the researchers differentiate subcomponents within EFs. A well-known study investigating the separability of EFs is Miyake, Friedman, Emerson, Witzki, & Howerter (2000)’s work. The authors used a latent variable analysis to investigate the separability of three executive functions: shifting, inhibition and updating. Shifting is defined as the ability
to change back and forth between multiple tasks, operations, or mental sets. *Updating* helps to replace old, no longer relevant information with newer, more relevant information in working memory. The third target EF, *inhibition* is responsible for the inhibition of dominant, automatic, or prepotent responses when necessary. According to their results, these are three distinct sub-functions within EFs, and they are separable but related constructs. Among the three EFs, inhibition is defined very similarly to cognitive control and often the same task is used to assess cognitive control and inhibition in different works. Novick’s research group (whose cognitive control definition we are following) also noted in a recent study (Hussey, Harbison, Teubner-Rhodes, Mishler, Veloskey & Novick, 2017) that the concept of cognitive control is similar to Miyake et al.’s (2000) construct of inhibition but they don’t use the term because they aim to be neutral about whether cognitive control involves inhibition of task-irrelevant representations or promotion of task-relevant ones or a combination of both processes. Therefore, although executive functions, inhibition and working memory refer to similar processes, we will use the term cognitive control in line with the literature motivating our studies.

In what follows studies investigating the role of cognitive control in language processes will be reviewed with focusing on the two domains targeted in our studies presented in the dissertation: sentence comprehension and word production\(^2\). Results from healthy adults, brain-damaged patients and healthy children will be reviewed.

### 2.2. Sentence comprehension and cognitive control

Cognitive control might play an important role in sentence processing in those cases when multiple incompatible analyses generate conflict due to temporal or standing ambiguity.

\(^2\) Note that cognitive control might play a role in other language processes too but due to time limitations we concentrated on these two domains during the research presented in the dissertation.
Ambiguity is prevalent in languages, for example every time when a string of words is compatible with more than one syntactic structure (*The children gave the flowers... to the girl*) – until the appearance of the phrase *to the girl*, structurally, the phrase *the flowers* can be analyzed both as the object and the recipient of the verb *gave*, but it does not always lead to a conflict. In this example *the flowers* will be interpreted as the object of the verb *gave* and not as the recipient in the sentence fragment before the disambiguating phrase (*to the girl*) is processed thanks to world knowledge (i.e. flowers rarely *get* anything, they are usually *given* to someone).

Although in most cases the sentence can be analyzed without any difficulties, sometimes problems arise. For instance if it turns out that *flowers* is the recipient of the verb *gave* (*The boy gave the flowers some water*), then our word knowledge will support an analysis in which *flowers* is the object but the end of the sentence support another analysis in which *flowers* is the recipient of the verb. These types of sentences are interesting for linguistic research for a long time. Early studies assumed that we commit to one analysis based on specific principles of syntactic structure building during sentence processing and if this analysis fails, the sentence has to be reanalyzed (Frazier & Rayner, 1982). More recent theories of sentence processing state that all the possible analyses are activated simultaneously, and they are also weighted by different factors, like the frequency of a structure, linguistic as well as non-linguistic context of the sentence and also by the world knowledge of the person comprehending the sentence (see Boland & Blodgett, 2001; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994). Although several theories of syntactic ambiguity resolution expose the problem of conflicting representations, it is still not clear whether such a conflict is resolved within the language system or by more general processes. The cognitive control account assumes that cognitive control is required during the processing of these sentences for conflict resolution.
Conflict originates from structural ambiguity in the abovementioned examples. Although these types of sentences were investigated most extensively in the literature, there are other possible sources of conflict during sentence comprehension. An ambiguous word (homonym) also activates contradicting representations for instance in the sentence \textit{The bank is under reconstruction, therefore it is forbidden to run or walk on it}. Since the ‘money institution’ interpretation of the word is more frequent and it is compatible with the context of the word, that will be the dominant meaning until the end of the sentence disambiguates it. According to the cognitive control view the two contradicting meanings generate conflict and cognitive control helps in its resolution. Another example is when neither the words nor the sentence is ambiguous but the sentence is semantically implausible. For instance in case of the sentence \textit{The thief kept the policeman in the police station} two interpretations will be activated: one based on our word knowledge and one based on the sentence structure. For the correct interpretation of the sentence conflict between these two interpretations have to be resolved.

Since anaphors refer back to earlier meanings in the sentence or discourse, they can also generate conflict when multiple antecedents are available. Here the source of conflict is that the anaphor is supposed to activate all compatible antecedents in the sentence regardless of the grammatical analysis. For instance in the sentence \textit{My friend said that the cat hates her}, \textit{her} activates both the noun phrase \textit{my friend} and \textit{the cat} and these two representations create conflict which has to be resolved to correctly interpret \textit{her} referring back to \textit{my friend}. Syntactic processes guiding anaphor resolution have been extensively studied (Chomsky, 1981, 1986; Chien & Wexler, 1990; Grodzinsky & Reinhart, 1993) and the role of working memory in the comprehension of sentences with anaphors has also been highlighted (Mongomery & Evans, 2009) but to our knowledge, the contribution of cognitive control has not been investigated yet. Therefore we aimed to study the role of cognitive control in the
The relationship between sentence comprehension and cognitive control was investigated with different methods in different populations. First results with healthy adults will be reviewed (individual differences in behavioral studies and brain imaging studies), then results from people with brain injuries will be presented and at the end of this section developmental research in the topic will be summarized.

2.2.1. Individual differences results about the relationship between cognitive control and sentence comprehension

One method to investigate the relationship between cognitive control and sentence comprehension is to study individual differences in both domains and test whether the measures are associated or not. There are only a few works focusing on cognitive control specifically but several studies aimed to investigate the relationship between working memory and sentence comprehension. As it was mentioned in section 2.1., working memory tasks, especially those taxing complex working memory might require cognitive control, therefore these results might be important regarding the topic of the dissertation. Daneman & Merikle (1996) conducted a meta analysis of 77 individual differences studies investigating the role of working memory in sentence comprehension. Working memory tasks measured either simple span (e.g.: digit span or word span in which sequences of digits or words have to be repeated) or complex working memory. The complex working memory task was most often a reading or listening span test. During the reading span participants are asked to read sentences and at the end of the block they had to recall the last words of each sentence. Some studies also included comprehension questions, a true/false judgment or a sensibility judgment after the sentences. The length of the blocks increased during the task. Complex WM span is the largest set for which the subject could successfully recall all of the final
words. In listening span task, which is similar to the reading span task, the sentences are presented auditorily and not visually. Results of operation span tasks were also included to the analysis. During this task, participants have to decide about equations instead of sentences whether they are correct or not and memorize a word or letter presented after the equations. The overall design of the task is the same as in the reading span task. Language abilities were measured with global or standardized tests of comprehension and vocabulary knowledge or with more specific tests of sentence processing, like tests assessing people's ability to identify the referent of a pronoun or to process ambiguous sentences. The authors found that complex working memory measures are better predictors of language performance than simple span tasks and they are better predictors of performance on specific tests of sentence comprehension than on global language tests. This association appeared in the case of the operation span task as well which is crucial because listening and reading span tasks require sentence comprehension, therefore associations with sentence comprehension tasks might appear only due to the common sentence comprehension component. This confound can be eliminated if the operation span task is used instead of the listening/reading span tasks. Since both the working memory tasks and the sentence comprehension tasks require the resolution of conflict between contradicting representations, the result of this meta-analysis supports the view that cognitive control might play a role in sentence comprehension. To investigate this hypothesis further, more specific tasks investigating cognitive control and sentence processing under conflict would be necessary but there are only a few studies investigating individual differences with such targeted tasks.

The unpublished doctoral dissertation of Mendelsohn (2002) cited by Novick et al. (2005) found that the performance on a cognitive control task, the verbal sorting task – in which words had to be categorized based on constantly changing criteria, such as meaning, grammatical category or syllable count – predicted the performance on tasks requiring the
resolution of lexical and syntactic ambiguity. Similar relationships are reported by Novick, Trueswell, January and Thompson-Schill (2004) who found that performance on a sentence comprehension task requiring ambiguity resolution is related to performance on an item recognition task measuring cognitive control. During the item recognition task (also called letter verification task) participants were visually presented with a set of four letters followed by one letter which was either among the previous four letters or was not. The task was to decide whether the actual letter was part of the immediately preceding set or not. In high conflict trials the actual letter did not appear in the preceding set but it appeared in the one before that. Therefore participants had to override familiarity resulting from recent appearance to correctly reject the letter with the help of cognitive control. In summary, these results suggest that cognitive control might play a role in ambiguity resolution during sentence processing.

2.2.2. Neuroimaging results about the relationship between cognitive control and sentence comprehension

Neuroimaging studies also point to the importance of cognitive control in sentence comprehension. These works usually measure the effect of conflict on brain activations in a cognitive control task and in a sentence processing task involving conflict. If the same area is activated during the cognitive control and the sentence processing tasks (i.e. overlapping activations appear), that suggests that cognitive control is involved in conflict resolution during sentence comprehension.

Following the general logic of imaging studies exploring the relationship between cognitive control and sentence comprehension, January, Trueswell and Thompson-Schill (2009) investigated overlapping activations in the LIFG for conflict resolution in an auditory sentence comprehension task with ambiguous sentences and in the Stroop task. Results showed that both the comprehension of ambiguous sentences and the incongruent condition
of the Stroop task increased the activation of the LIFG relative to control conditions, supporting the view that cognitive control is recruited for conflict resolution during sentence comprehension.

In a more recent study, Hsu et al. (2017) investigated brain activations during a sentence comprehension task and three cognitive control tasks: the item recognition task, the n-back task and the Stoop task. During the n-back paradigm words were presented after each other on the screen. Participants were asked to decide whether the actual word was the same as the one that appeared three trials before or not. For the manipulation of cognitive control lure trials were included, in which the stimulus appearing three trials before was not a match but the one appearing two or four trials before matched the actual word. For the correct rejection of these trials, participants needed to override the familiarity bias which requires cognitive control. The main aim of the study was to explore both overlapping and divergent brain activations during these tasks, in order to study both domain-general and domain-specific aspects of cognitive control. The results show consistent activation of the LIFG within individuals during conflict processing in various linguistic and nonlinguistic tasks. Furthermore, task-specific areas were also recognized, therefore the authors propose a domain general conflict-resolution “hub” which cooperates with other task-specific networks. However, the exact nature and scope of general and more specific areas are not clear yet.

Ye and Zhou (2009) explored another possible source of conflict in sentence comprehension by using semantically implausible sentences. In these sentences syntactic analysis contradicted world knowledge (e.g., *The thief kept the policeman in the police station*) and for a correct interpretation participants had to override the automatic, world knowledge-based analysis. Brain activation during the comprehension of these sentences was compared to brain activation during comprehension of sentences with a similar structure but with a semantically plausible meaning. In addition to the sentence comprehension task, a
Stroop task and a flanker task were assessed. During the flanker task participants were asked to press the right or the left arrow depending on the direction of the arrow displayed in the center of the screen. For congruent trials, the target arrow was surrounded by arrows pointing to the same direction as the target arrow, while for incongruent trials, the arrows on the side were pointing to the opposite direction. In the incongruent condition cognitive control was assumed to help to focus on the target arrow and overcome the conflicting information generated by the distractor arrows. A significant overlap was found in the activations during the conflict trials compared to the control trials in all three tasks, which means that the same brain areas are recruited for resolving conflict during sentence comprehension, a non-syntactic language related task (Stroop task) and a non-language related task (flanker task).

To summarize, results from brain imaging studies clearly support the hypothesis that cognitive control is involved in conflict resolution in sentence comprehension.

2.2.3. Neuropsychological results about the relationship between cognitive control and sentence comprehension

Another line of research in which the relationship between cognitive control and certain linguistic processes was observed is the work related to patients with brain injury. Individual differences studies and neuroimaging studies discussed above are able to show whether the same conflict resolution processes play a role during sentence comprehension and cognitive control tasks but they cannot be used to determine whether cognitive control is necessary for conflict resolution in language. Neuropsychological studies, however, are suitable for the investigation of the question. If a patient with a focal injury of the LIFG – the area responsible for cognitive control – shows impairments in conflict resolution in language but displays good performance when conflict resolution is not necessary, that is a strong
support for the hypothesis that cognitive control is necessary for conflict resolution in language.

Patients with aphasia with a focal lesion of the LIFG often show surprisingly good general language skills but in some situations they have extreme difficulties. Luria (1973) called this impairment “dynamic aphasia” because the impairment was sometimes present but other times it disappeared. One explanation for this variability in language abilities is that the problem is not a consequence of a primary impairment of language abilities, but it arises from a cognitive control impairment, and as a result, difficulties only appear when there is a conflict between linguistic representations, and cognitive control would be required for their resolution.

Novick and his research group tested this hypothesis and found in multiple studies that patients with a focal injury in their LIFG have problems with the comprehension of sentences involving syntactic ambiguity (Patient N.J. in Novick, January, Trueswell, & Thompson-Schill, 2004; Patient I.G. in Novick, Kan, Trueswell, Thompson-Schill, 2009). They also assessed I.G.’s performance on a letter verification task to measure I.G.’s cognitive control. The results showed an impaired performance in the conflict condition while I.G. reached normal scores in the control condition.

Vuong and Martin (2011) report results from two patients with an impairment in their LIFG. They employed a sentence comprehension task with sentences with a lexical ambiguity which could be resolved based on sentence context (She mixed the punch). Patients were successful in resolving lexical ambiguities but showed unusually long processing times when the word’s generally less frequent (subordinate) meaning was supported by the context (He drank the port). The patient’s cognitive control was also measured with a Stroop task, a picture-word interference task (pictures of objects had to be named while a word also appeared on the screen which was semantically related to the
picture in the conflict condition and non-related in the control condition) and a recent negatives task (in each trial three words were presented and then one word appeared on the screen; the patient had to decide whether the word was a part of the current set; in the conflict conditions, the probe word was related (semantically or phonologically) to one word from the current set or from the previous set). Both patients showed a significantly longer increase in reaction times for the conflict conditions of the tasks than controls.

Taken together, studies with LIFG impaired patients show that conflict resolution during sentence processing is impaired together with cognitive control. This suggests that cognitive control is necessary for conflict resolution in language.

2.2.4. Developmental results about the relationship between cognitive control and sentence comprehension

The relationship between cognitive control and language abilities was studied in children as well. The prefrontal cortex involving the LIFG is one of the latest maturing brain areas; its development can last until late adolescence (e.g., Huttenlocher & Dabholkar, 1997) which implies the slow development of cognitive control.

If cognitive control is involved in certain aspects of language, then we expect difficulties in these language areas involving conflict in young children.

To investigate the role of cognitive control in sentence comprehension in children, several studies used an act-out paradigm in which conflict is generated by applying temporally ambiguous sentences as instructions (e.g., Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2001; Trueswell, Sekarina, Hill, & Logrip, 1999; Choi & Trueswell 2010; Woodard, Pozzan & Trueswell, 2016). Most of these studies use sentences in which temporal ambiguity is generated by a phrase which can be interpreted both as a destination of the action described by the verb and as the modifier of the noun phrase. An example is the sentence *Put the frog on the napkin into the box* in which the phrase *on the napkin* could be
interpreted as the destination of the event or as the modifier of the frog until the phrase into the box appears. Adult data show that adults initially interpret the phrase on the napkin as a destination but when into the box appears they override this initial preference and interpret the phrase as a modifier. In contrast, young children often make errors while executing the task, relying on the destination interpretation (e.g., they put the frog on the napkin instead of putting it into the box) (Trueswell et al., 1999; Choi & Trueswell, 2010). According to the cognitive control view (Choi & Trueswell, 2010; Novick et al., 2005), these phenomena can be accounted for by weak cognitive control, because of which children are less successful in overriding the initially dominant destination interpretation. This explanation is supported by Woodard et al. (2016)’s results as well, who found that individual differences in the comprehension of temporally ambiguous sentences in 4-6 year old children are associated with differences in their certain cognitive control scores.

The developmental aspect of the relationship between cognitive control and sentence comprehension was investigated with sentences with a non-canonical structure as well. My master thesis (Ladányi, 2012) investigated the comprehension of sentences with relative clauses with different structures. One factor which was manipulated was the structure of the main clause. In one condition the main clause had an SVO structure while in the other condition it had an OVS structure. In Hungarian, as in many other languages, the SVO structure is more frequent than the OVS. According to some models of sentence comprehension if the lexical units of the sentence are compatible with the more frequent (in this case SVO) analysis then this dominant analysis will be activated (Ferreira, 2003). But if the sentence has a non-canonical (not SVO) structure then this frequency-based default analysis has to be overwritten with the correct non-canonical structure, resulting in conflict, and thus potentially requiring cognitive control. As a result of this, young children with weak cognitive control are expected to have difficulties in the comprehension of sentences with a
non-canonical structure. Indeed, we found a significantly better performance in the case of the canonical SVO sentences than in OVS sentences in pre-school and primary school children. We also measured cognitive control with an n-back task, a Stroop task, a non-verbal Stroop task and the backward digit span task in these children and we found a significant correlation between the performance on the sentence comprehension task and the backward digit span, as well as the n-back tasks suggesting that cognitive control plays a role in the comprehension of sentences with a non-canonical sentence structure.

To summarize, children seem to have difficulties in the comprehension of sentences that require cognitive control supporting the hypothesis that weaker cognitive control leads to weaker performance on the comprehension of sentences requiring cognitive control.

### 2.3. Word production

Cognitive control also seems to play an important role in word production under conflict (Kan & Thompson-Schill, 2004; Schnur, Schwartz, Brecher, & Hodson, 2006; Schnur, Schwartz, Kimberg, Hirshorn, Coslett, & Thompson-Schill, 2009). When we want to produce a word, several competing word representations are activated beyond the target representation. When one wants to produce the word *table*, words with similar meaning (*desk*) and semantically connecting words (*tablecloth, plate, chair, lamp, eat, write...*) are also activated. In many cases, the activation levels of the competing words are too weak to affect the retrieval of the target word. In some cases, however, the activation level of the competing word can be similar or higher than the activation level of the target word. In these cases we can experience a tip of the tongue phenomenon (we cannot produce a word because its activation is not high enough relative to the competing word’s activation, but we have several pieces of information available about it) or slips of the tongue (we produce another word instead of the intended one because its activation level is higher for some reason than...
the activation level of the target word). Studies presented below argue that cognitive control plays a role in resolving conflict in these situations.

In this domain, studies are not focusing on relationships between individual differences in word production and cognitive control but rather take results from brain imaging and lesion studies which are summarized below.

2.3.1. Neuroimaging results about the relationship between cognitive control and word retrieval

Schnur et al. (2009) used a blocked-cyclic naming task to study word production under conflict in healthy adults with a brain imaging study. In this task, pictures of simple objects appear on the screen and participants are asked to name pictures in the context of other pictures either from the same category (homogeneous block, e.g., *pear, apple, melon . . .*) or from different categories (mixed block, e.g., *pear, chair, blouse . . .*). Conflict is assumed to be higher in the homogeneous blocks because previously named semantically similar pictures are still active and they get extra activation due to their semantic relationship to the target word therefore the difference between these competing representations and the target word does not reach a critical difference. In accordance with the cognitive control view, cognitive control is required to resolve this conflict. Schnur and her colleagues (2009) found that the LIFG was more active in the homogeneous condition than in the mixed condition suggesting that the LIFG is recruited for conflict resolution in word retrieval. Since the two conditions differed only in the degree of conflict and the LIFG is traditionally considered to be responsible for cognitive control, that result strongly supports the view that cognitive control helps in resolving conflict appearing during word production.

In another study, Kan & Thompson-Schill (2004) manipulated name agreement to change the level of conflict during word retrieval. Name agreement is determined by the number of names available for describing a picture. High name agreement pictures can be
named with one word only, or even when there are more available names, one of them is used much more frequently than the others (a picture of an apple is usually named as an *apple*). In the case of low agreement pictures, more than a single name is available with similar probabilities of use (a picture of a stove can be named as *stove*, *oven* or *range*). In the latter case conflict is expected to be higher since the alternatives are competing with each other for selection. Cognitive control might be recruited for biasing the selection in these cases. The results of Kan & Thompson-Schill (2004) support this idea: the LIFG was active in the low naming agreement condition but not in the high naming agreement condition.

**2.3.2. Neuropsychological results about the relationship between cognitive control and word retrieval**

Schnur and her colleagues conducted the same blocked-cyclic naming task presented above with patients with an injury in their LIFG (Schnur et al., 2006, Schnur et al., 2009). Patients showed greater difficulties than typical adults in the semantically homogeneous blocks mirrored by more errors and long reaction times while they showed typical performance in the case of the semantically mixed blocks.

A case study was reported as well investigating word production under conflict conducted by Novick et al. (2009) with a LIFG impaired patient, I.G., whose sentence comprehension and item recognition results were already presented above. In this study the authors used name agreement to manipulate the level of conflict. They found that I.G. could not name pictures properly in the low name agreement condition while I.G.’s performance was typical in the high name agreement condition.

These results show that patients with an impairment in their LIFG – which is responsible for conflict resolution – have difficulties with word retrieval but only under conflict, suggesting that cognitive control is recruited for word production under conflict.
2.3.3. Developmental results about the relationship between cognitive control and word retrieval

For the investigation of conflict resolution in the lexical domain, the blocked-cyclic naming paradigm was used in children as well. Boelens and La Heij (2017) assessed the task in a younger (5-7 years old) and older (10-12 years old) group of children and found significantly longer naming times for semantically homogeneous blocks than for mixed ones, but the size of the effect did not differ between age groups. Snyder and Munakata (2013) used the same paradigm to investigate the effect of conflict on word production in 6 year olds and found that children showed significantly longer naming latencies in the semantically homogeneous than in the semantically mixed condition. Charest (2017) also manipulated the context of pictures in a picture naming task and found that 3-years-old children showed longer naming times in the semantically homogeneous context than in the semantically mixed context.

In sum, studies with children show that – similarly to adults – they take longer to name pictures under conflict. We would expect a greater conflict effect in children than in adults due to weaker cognitive control, we did not find, however, any studies comparing the effect of conflict in children and adults directly. Nevertheless, the result that the conflict effect is similar in 5-7 and 10-12 year old children suggests that cognitive control is already developed enough to resolve conflict in word retrieval in young children.

Motivated by these findings, the role of cognitive control in word production under conflict is investigated in two of the studies presented in the dissertation. Study 2 manipulates the level of conflict in a picture naming task and investigates the performance of children with SLI and TD children in different conditions (see Study 2 and Thesis point 3), while Study 4 tests word production with two tasks potentially requiring cognitive control and studies the relationship between the performance on these tasks and on cognitive control tasks (see Study 4 and Thesis point 5).
3. Cognitive control in specific language impairment

In the previous section studies using various methods with different populations were presented suggesting that cognitive control might play a crucial role in certain language processes. Studies presented in the first section showed that non-linguistic functions can be impaired in SLI beyond linguistic functions, the impairment of working memory being one of the main characteristics in the population. Since working memory tasks require cognitive control (as discussed in the previous section; see also Marton et al. 2006 and 2007 discussed below), one can hypothesize that cognitive control is also impaired in SLI and these impairments might account for some of the language problems appearing in children with SLI. There are, however, only a few studies investigating the SLI population’s cognitive control and even fewer which aim to discover the relationships between cognitive control and language functions. Below results from tasks requiring cognitive control will be reviewed. Note that these tasks are not referred to as cognitive control tasks by the authors of the studies but based on the cognitive control literature, all of them require cognitive control. Importantly, all of these tasks require other abilities beyond cognitive control (e.g., almost all the tasks rely on short term memory and performance on the fluency tasks is strongly affected by vocabulary size), therefore these abilities should be controlled for and the results of those studies which do not take this into account should be interpreted with caution. Since the same task is often used to measure different constructs in the different studies due to terminological or theoretical discrepancies between the studies (e.g., the n-back task is referred to as measuring working memory, cognitive control, updating and inhibition in different studies), the results below are grouped by tasks instead of the functions they assess (see also section 2.1. about terminological questions).
3.1. The Listening span task

The listening span task is extensively used in SLI studies to assess cognitive control or related abilities. It is typically applied to measure executive loaded/complex working memory which is our ability to store and manipulate information at the same time. It also requires, however, cognitive control for the selection of the last word of the sentence which has to be memorized and the inhibition of last words from previous sets and task-irrelevant words from the previous sentences.

Several studies found significantly lower span in children with SLI than in TD children in this task (Ellis Weismer et al. (1999) in 5 to 9 years old children, Archibald & Gathercole (2006) in 7 to 11 years old children, Mainela-Arnold & Evans (2005) in 8 to 12 year old children, Vugs, Hendriks, Cuperus, & Verhoeven (2014) in 4-5 years old children).

Motivated by earlier studies suggesting the involvement of WM limitations in language problems of children with SLI Marton et al. (2006) aimed to find out the source of difficulties in working memory tasks in children with SLI. They manipulated the storage (sentence length) and processing (grammatical complexity) demands of a listening span task independently and studied their effects on the performance in 7 to 11-year-old Hungarian children with SLI. Children with SLI showed a significantly weaker performance on all measures of all tasks but the pattern of results was similar in the two groups: grammatical complexity led to greater performance drop both in children with SLI and in TD children. Although grammatical complexity had a similar effect on accuracy scores, the characteristic type of errors differed between the two groups: TD children usually omitted words or substituted them with words with similar meaning while children with SLI produced words which appeared earlier in the task in a non sentence final position instead of the proper sentence-final word. After finding the same result in another group of children with SLI (Marton et al., 2007) the authors suggest that problems with inhibiting irrelevant words are
responsible for weak performance on complex working memory tasks. This idea is in accordance with our hypothesis, that cognitive control is impaired in children with SLI.

Montgomery and Evans (2009) also found significantly weaker performance on a listening span task in 6 to 12 years old children with SLI than in age matched TD peers and, importantly, the SLI group’s performance was also associated with their performance on the comprehension of complex sentences. This suggests that difficulties with complex sentence comprehension might be partly caused by SLI children’s difficulties with cognitive control. Associations might appeared, though, due to the shared sentence processing component of the two tasks (see section 2.2.1), therefore this result should be interpreted with caution.

3.2. Backward digit span task

Another task targeting executive loaded/complex working memory in the verbal domain but without involving sentence comprehension is the backward digit span task. Cognitive control might be used for inhibiting irrelevant numbers from previous sets and numbers from the same set which should be produced later. Children with SLI tend to show weaker performance than their age-matched controls according to several studies (e.g., Lum, Conti-Ramsden, Page, & Ullman, 2012 in 8-11 year-old children; Vugs et al., 2014 in 4-5 year old children).

3.3. The Odd one out task

Researchers aimed to investigate executive loaded/complex working memory in SLI in the visuospatial domain as well. One task designed to assess this function is the odd one out task (Henry, 2001) which – similarly to the listening span task – requires the simultaneous storing and manipulation of visuospatial information and cognitive control as well for resolving conflict between the to be memorized and interfering locations in the task. Results are more controversial in the visuospatial than in the verbal domain: Vugs et al. (2014) found significantly weaker performance in 4-5 year old children with SLI than in their
TD peers while Archibald and Gathercole (2006) did not find any differences in 9-11-year-olds and neither did Engel de Abreu, Cruz-Santos, and Puglisi (2014) in 7-9 year old children. Based on these few studies, the controversial results can be explained by the age differences of the children assessed. Children with SLI might have a mild impairment in visuospatial executive loaded working memory which disappears in older children due to compensation strategies, or the ability may simply mature slower than in TD children.

3.4. The N-back task

The n-back task is often used to measure cognitive control and in the frame of the multicomponent working memory model, it primarily measures updating but requires monitoring and inhibition as well. Cognitive control is required to resolve conflict between the item based on which the response have to be made and previously presented items.

Evans and Pollack (2011) conducted an n-back study with measuring event related potentials (ERPs) in 12-14 year old children with SLI and TD children. Stimuli were auditory (words) in one condition and visual (human faces) in the other and participants completed both a 1-back and a 2-back condition. In the auditory domain children with SLI showed comparable performance to TD children in the 1-back condition but were less accurate in the 2-back condition and also showed an atypical pattern in their ERP signal. In the visual domain, no group difference appeared in the behavioral performance of the two groups but their ERP signal was atypical, suggesting that the same performance was achieved with higher processing costs in the SLI group.

Im-Bolter, Johnson, & Pascual-Leone (2006) used a visual n-back task in which configurations of three dots were presented on the screen and children had to decide whether the configuration is the same as the one shown n items earlier. They used 0-back (children had to press the button when a certain configuration appeared in the sequence), 1-back (children had to press the button when the same configuration came up one after the other)
and 2-back (children were asked to press the button when the stimulus was the same as the one viewed two stimuli ago) conditions. Performance of the two groups did not differ in the 0-back and 2-back conditions: both groups showed good performance in the 0-back condition and were at about chance in the 2-back condition. In the 1-back condition the SLI group’s performance was significantly below the TD group’s performance.

These few studies suggest that children with SLI have difficulties with the n-back task especially in the more challenging conditions suggesting the impairment of cognitive control.

### 3.5. The Stroop task

The Stroop task (Stroop, 1935) - in which conflict between the color and meaning of the word has to be resolved in the original version - is a classical measure of cognitive control (see description of the task above). Reichenbach, Bastian, Rohrbach, Gross and Sarrar (2016) investigated 5-6 year old children who cannot read yet, and therefore created a modified version requiring no reading skills. During this task, children first had to rapidly name the color of black and white objects of fruits and vegetables. In the second part, they were asked to do the same but now the objects were colored incongruently to their natural color (e.g.: a blue strawberry). In this case conflict arises between the color of the object presented on the screen and the prototypical color of the object. Results showed that both the SLI and the TD groups were slower in the second part of the task but children with SLI were not significantly slower than TD children, suggesting that they were as good as TD children at resolving interference originating from the incongruent color of the picture. Based on this one task, children with SLI do not seem to have problems with overcoming interference in a Stroop task.

### 3.6. Fluency tasks

Fluency tasks also involve executive functions/cognitive control. During this task words have to be produced according to different criteria (words starting with a given letter in
the letter fluency task, words belonging to a given category in the category fluency). Cognitive control is assumed to play a role in resolving conflict between the item which one wants to produce and irrelevant items, already produced or to be produced items (See a more detailed description of the fluency paradigms in Study 4).

Kail & Leonard (1986) assessed a category fluency task in 6-14 year old children with SLI and in two groups of TD children matched in age and in language skills to the SLI group with finding no difference between the groups in the number of words produced. A category fluency task was performed by 6-12 year old children with SLI and age-matched TD children in Weyandt and Willis (1994)’s study too who did not find any group differences either.

Other studies, however found significant differences both in the category and the letter fluency task between children with SLI and TD children. According to Weckerly, Wulfeck, and Reilly (2001) 8-12 years old children with SLI were able to produce significantly less items than their age-matched TD peers both in the case of the letter and category fluency tasks. 8-14 year-old children with SLI produced significantly less items than their TD peers in Henry, Messer, and Nash (2012)’s study as well in a fluency task with both category and letter fluency conditions. In a following study (Henry, Messer & Nash, 2015) – reporting more detailed analyses on the fluency scores of the same group of children – the authors found that the SLI group produced less correct words, more errors and less switches both in the letter and the category fluency task than the TD group. Relationships with executive functions were also investigated and an association appeared between the number of errors in the letter fluency task and inhibition performance. Inhibition was measured with a task in which the child had to repeat the word produced by the experimenter (doll or car) in the first part of the task. In the second part, they had to produce doll when the experimenter said car and produce car when the experimenter said doll. Accuracy in the second part was used as the measure of inhibition. Rodriguez, Santana, and Exposito (2016) also found a
significantly weaker performance in the number of words produced in 5-11 years old children with SLI than in TD children both in a letter and a category fluency task.

In sum, the results of studies are controversial in the case of the category fluency task but children with SLI generally show a weaker performance when words starting with the same phoneme have to be produced.

For measuring fluency in the non-verbal domain Henry et al. (2012) used the design fluency task in which children are asked to draw as many different designs as possible in one minute by connecting dots with 4 straight lines. In the first condition there were only filled dots while in the second condition filled and empty dots were displayed and children had to connect only empty dots. 8-14 years old children with SLI were able to produce significantly less designs than their TD peers in the two conditions. This result suggests that fluency is impaired not just in the verbal domain but in the nonverbal one as well.

### 3.7. Category judgment under conflict

While the tasks mentioned so far used cognitive control/working memory/executive function tasks to measure cognitive control in children with SLI, Marton, Campanelli, Eichorn, Scheuer, & Yoon (2014) manipulated conflict within a linguistic task. They used a category judgment task with conflict manipulations in which a category name appeared on the screen (e.g., *Family*) followed by either a target word belonging to the category (e.g., *Mother*) or a distractor item (e.g., *Ball*) and the child had to decide whether the item belongs to the category or not with a button press. In conflict trials the distractor item was a target word in the preceding category, therefore, for the correct rejection of the item, the conflict between the ‘yes’ answer based on the previous category and the ‘no’ answer based on the current category have to be resolved. Accuracy was generally significantly lower in conflict trials than in distractor trials when a new word was presented but the difference between these conditions was significantly bigger in children with SLI, suggesting the impairment of
3.8. Cognitive control skills in children with SLI: summary

In summary, children with SLI show significant impairments on various tasks measuring cognitive control. There were only two tasks on which children with SLI showed a comparable performance to TD children: on the odd one out task, difficulties seem to disappear by school age and no impairment appeared in the Stroop task (although this is based on only one study available in the literature). Overall, there are very few studies investigating performance on cognitive control tasks and only one of them studied the relationship of cognitive control scores with language measures. In conclusion, it is still an open and debated question whether cognitive control is impaired in SLI, and whether it is associated with impairments in language.

4. Aims and thesis points

As the above review suggests, cognitive control might play a role in several language processes although research is limited in the area and there are several areas of language in which the role of cognitive control has not been studied yet. There are even less studies exploring cognitive control and its role in language problems of children with SLI. The main aim of the dissertation is to contribute to this line of research with investigating whether cognitive control is impaired in SLI and whether cognitive control and certain language abilities are related to each other in children with SLI and in TD children. In the studies presented below, three hypotheses are going to be tested. First, we hypothesized that cognitive control is weaker in children with SLI than in their typically developing peers (see our results related to this hypothesis in Thesis 1 and 2). Our second hypothesis was that children with SLI show more pronounced difficulties in language processes that require conflict resolution like word retrieval under conflict or comprehension of sentences with an
anaphor (see results related to this hypothesis in Thesis 3, 4 and 5). According to our third hypothesis, performance on cognitive control tasks and language tasks in which conflict appears will be associated with each other showing the involvement of cognitive control in certain language processes (see results related to this hypothesis in Thesis 4 and 5).

Although the focus of this dissertation is SLI, the results of these studies can enrich our knowledge about the role of cognitive control in language processing in general. Beyond the theoretical importance, these questions have an important clinical relevance as well. If cognitive control is impaired in children with SLI and these impairments contribute to problems in language, targeted training of cognitive control should be used in language therapy.

Thesis 1. Children with SLI show impairments on some, but not all cognitive control tasks.

Study 1 measured performance of children with SLI and TD children on eight cognitive control tasks: a listening span task, an odd-one out task, a verbal and nonverbal n-back task, a verbal and nonverbal Stroop task and a verbal and nonverbal fluency task. A significantly weaker performance appeared on the listening span and verbal fluency tasks suggesting weaker cognitive control in children with SLI than in TD children. Study 3 and Study 4 assessed performance on a backward digit span task, an n-back task and a Stroop task and found a weaker performance in the SLI group than in TD children on the backward digit span and the n-back tasks.

Publications related to these this point:


Ladányi E., & Lukács Á. (under revision in JSLHR). Cognitive control impairment and its contribution to word production difficulties in specific language impairment.

**Thesis 2.** Weaker performance on some of the cognitive control tasks is accounted for by weaker verbal short-term storage capacity in children with SLI while the group difference is present even after accounting for weaker short-term storage capacity in other cognitive control tasks.

In *Study 1 and Study 4* the role of verbal short-term memory was also investigated in cognitive control impairments of children with SLI. In *Study 1* performance differences on the listening span and verbal fluency tasks disappeared when simple verbal span was included as a covariate. These results suggest that cognitive control is intact in children with SLI, while short-term memory is impaired leading to weaker performance on cognitive control tasks. In contrast, in *Study 4* children with SLI showed a weaker performance on the backward digit span and n-back task even after accounting for their weaker short-term memory capacity.

Publications related to this thesis point:


Thesis 3. Children with SLI are generally slower in naming pictures but they resolve conflict as successfully as their TD peers during word retrieval.

In Study 2, a picture naming task was created in which we manipulated the level of conflict 1) by presenting pictures in a semantically homogeneous vs. semantically mixed context 2) by presenting pictures with low name agreement – i.e. pictures which may be named with more than one name (e.g.: a picture of a sofa) – vs. pictures with high name agreement – for which a single dominant name exists (e.g., a picture of an apple). The naming of pictures presented in a homogeneous context and of those with low name agreement is assumed to require conflict resolution for inhibiting already named semantically similar words in the first case and other possible names of the picture in the second case. Conflict resolution is expected to manifest in longer reaction times relative to the low conflict condition. Our results showed that 1) for children with SLI it generally took longer to name the pictures and 2) both the SLI and the TD group needed more time to produce the names of pictures with high conflict than that of low conflict 3) but the difference was not bigger in the case of the SLI group than in the TD group. These results suggest that children with SLI were able to resolve conflict during word production as successfully as TD children.

Publication related to this thesis point:


Thesis 4. Children with SLI show weaker performance on the comprehension of
sentences presumably requiring conflict resolution (comprehension of sentences with an anaphoric expression) than their TD peers and sentence comprehension performance is associated with cognitive control scores in children with SLI.

In Study 3, 7-11 year old children with SLI and age-matched TD children completed a sentence-picture verification task with anaphoric expressions and performed an n-back, a backward digit span and a Stroop task. The reference of anaphoric expressions is not always obvious therefore more than one referents can be activated and the correct interpretation of the sentence might require cognitive control for the resolution of conflict between these alternative referents. If cognitive control is weaker in children with SLI and it is necessary for conflict resolution during the comprehension of anaphoric sentences then children with SLI will perform weaker on both the cognitive control and sentence comprehension tasks and these performances will be associated. Children with SLI showed significantly weaker performance on the sentence comprehension task, on the n-back task and on the backward digit span task. Sentence comprehension scores were associated both with the n-back and the backward digit span scores but according to a linear regression analysis the n-back performance alone was the best model of sentence comprehension performance. Neither the backward digit span, nor the non-word repetition span (both measuring verbal short-term storage capacity) contributed to the model meaning that there is a relationship between cognitive control and the comprehension of sentences with anaphors, and this association does not appear due to the shared verbal storage component of the two tasks.

Publication related to this thesis point:


Thesis 5. Children with SLI show significantly weaker performance on word production
tasks presumably requiring conflict resolution (the letter fluency test and the size-color-shape rapid automatized naming test) than their TD peers and performance on these word production tasks is associated with cognitive control performance in children with SLI and TD children.

In Study 4, 7 to 11 years old children with SLI and age-matched TD children performed two word production tasks which might require cognitive control (the fluency task and the rapid automatized naming (RAN) task) and their cognitive control was also assessed with a backward digit span task, an n-back task and a Stroop task. Cognitive control might be recruited for the resolution of conflict between the irrelevant words/already produced words/to be produced words and the target word during the fluency task and between already produced sizes/colors/shapes and the target size/color/shape as well as between lexical units referring to the size, the color and the shape and the current target item (e.g., when the name of the shape is activated when the color is the target word). If cognitive control plays a role in conflict resolution during these word production tasks then performance on these tasks would be associated with cognitive control performance. Furthermore, if cognitive control is impaired in children with SLI, than they will show a weaker performance both on the word retrieval tasks and the cognitive control tasks than their TD peers. Children with SLI showed a weaker performance on the n-back and backward digit span tasks relative to the TD group and weaker performances on the backward digit span task were associated with lower letter fluency scores and longer naming times in the RAN task while weaker n-back performance was associated with longer naming times in the RAN task. Performance on both word retrieval tasks was best predicted by the backward digit span and non-word repetition scores which is the measure of short-term memory. Our results support the hypothesis that domain-general cognitive control plays a role in word production under conflict but short-term memory also has an important role.
Publications related to this thesis point:


Ladányi E., & Lukács Á. (under revision in JSLHR). Cognitive control impairment and its contribution to word production difficulties in specific language impairment.
5. Studies

Study 1.

Executive Functions and the Contribution of Short-Term Memory Span in Children With Specific Language Impairment

Ágnes Lukács
Budapest University of Technology and Economics

Enikő Ladányi
Budapest University of Technology and Economics and Hungarian Academy of Sciences

Kata Fazekas and Ferenc Kemény
Budapest University of Technology and Economics

Objective: An increasing number of results show that specific language impairment (SLI) is often associated with impairments in executive functions (EF), but the nature, extent, and generality of these deficits is yet unclear. The aim of the paper is to present results from verbal and nonverbal tasks examining EF in children with SLI and their age-matched typically developing (TD) peers. Method: 31 children with SLI were tested on verbal and nonverbal versions of simple and complex span, fluency, N-back, and Stroop tasks. Their performance was compared with 31 TD children matched on age and nonverbal IQ. The design allows us to examine whether executive functions are similarly affected in SLI in verbal and nonverbal tasks. Results: The SLI group showed difficulties in verbal versions of complex span (listening span task) and fluency but not in inhibition (Stroop tasks) relative to TD age-matched children. Including simple verbal span (digit span) as a covariate eliminated group differences on both verbal tasks. Conclusions: Children with SLI were found to be impaired on several verbal measures of EF, but these differences were largely due to more fundamental deficits in verbal short-term span.

Keywords: specific language impairment, executive functions, short-term memory

Specific language impairment (SLI) is defined as a developmental disorder where language abilities are impaired in the absence of any hearing deficits, neurological disorders, intellectual disability, or other obvious nonlinguistic impairments that would explain the language problems. An increasing number of results, though, show that SLI is often associated with impairments in executive functions (EF), whereas the nature, extent, and generality of these deficits is yet unclear (see also the recent debate about terminology and diagnosis and comments: Bishop, 2014; Reilly et al., 2014). Language, as any complex goal-directed human behavior, relies heavily on nonlinguistic higher order executive functions responsible for the efficient coordination, selection, strategic use, and sustainment of relevant information in time. The role of working memory in language acquisition, production, and processing is now well-established (e.g., Engel de Abreu, Gathercole, & Martin, 2011; Lewis, Vasishth, & Van Dyke, 2006; Montgomery, Magimairaj, & Finney, 2010), and there is a growing body of research demonstrating that executive functions play an important role in production and comprehension at both the word and the sentence level where representations compete (e.g., Novick, Truewell, & Thompson-Schill, 2005, 2010). These findings point to the importance of examining executive functions in SLI together with their relationship with language abilities and potential contribution to language symptoms.

Executive Functions in SLI

Executive functions are a set of cognitive processes associated with the coordination, control, and regulation of other cognitive functions and adaptive and efficient goal-directed behavior, especially in complex and new tasks and situations that require sustained conscious attention, or in dual- and multiple-task situations involving divided attention (e.g., Anderson, 2002; Burgess, 2000; D’Esposito et al., 1995; Friedman et al., 2006; Huizinga, Dolan, & van der Molen, 2006; Miller & Cohen, 2001; Miyake et al., 2000). EF include generation of new responses (fluency); planning; switching between different tasks, mental sets, or actions; inhibition of irrelevant stimuli and inappropriate responses; generation of new responses and concurrent storage; updating; and manipulation of working memory representations of context-relevant information (Miyake et al., 2000). The concept and specific compo-
nents of executive functions are still subjects of considerable debate in the literature (e.g., Anderson, 2002; Baddeley, 1996; Engle & Kane, 2004; Miyake et al., 2000; Norman & Shallice, 1986; Smith & Jonides, 1999).

Evidence for EF impairments in SLI is controversial but present in several areas. Parental and self-ratings assessing everyday activities involving executive functions (on the Behavior Rating Inventory of Executive Function, Self-Report and Preschool Versions; Gioia, Espy, & Isquith, 2003; Guy, Isquith, & Gioia, 2004) in SLI suggest that deficits in executive function are severe enough to affect academic and everyday lives of both adolescents (Hughes, Turkstra, & Wulfeck, 2009) and preschool children with SLI (Wittke, Spaulding, & Scheckman, 2013). The studies summarized below give us a more detailed, though far from unequivocal, picture of how specific EF subfunctions are affected according to experimental results.

Attention

A number of studies suggest that children with SLI have problems in both selective and sustained attention. Problems have been documented in sustained attention in preschool children (4 to 6 years) in the visual (Finneran, Francis, & Leonard, 2009) and in the auditory domain (Spaulding, Plante, & Vance, 2008), together with deficits in the temporal engagement of attention (in 5- to 8-year-old children; Dispaldro et al., 2013).

Planning

The few studies addressing the question of planning suggest problems in this domain as well. Marton (2008) observed more perseverative errors, failures to develop a rule, rule violations, and more impulsive responses in children with SLI between 5 and 7 years than in typically developing (TD) children on both the Wisconsin Card Sorting Test and the Tower of London test. Weyandt and Willis (1994) also found impaired planning performance on the Tower of Hanoi in a group of children with developmental language disorder (6 to 12 years). As successful planning relies on several other executive functions, these results probably reflect deficits not only in planning, but also in controlled attention, switching, inhibition, and goal maintenance.

Inhibition

Evidence for inhibition difficulties in SLI is controversial, but most results show problems with inhibitory control. Bishop and Norbury (2005) found deficits on tasks testing both verbal and nonverbal inhibition of prepotent responses in school-age children with SLI. Im-Bolter, Johnson, and Pascual-Leone (2006) documented impairment in an antisaccade task and also on incompatible trials of the children’s trail making task (with intact switching) in 7- to 12-year-old children with SLI. In a continuous performance task, Finneran et al. (2009) found a reduced capacity to monitor the target stimulus and inhibit distractor stimuli in preschool children with SLI. On the other hand, Noterdaeme, Amorosa, Mildenberger, Sitter, and Minow (2001) found no group differences on the go/no-go task in school-age children and adolescents with SLI.

Preschool children with SLI, according to Spaulding’s (2008, 2010) results, were more susceptible to distraction by speech, environmental sounds, and visual stimuli. They more frequently made incorrect button presses on stop trials in a stop signal task, and fewer correct button presses on go trials. Distractibility and inhibition scores also correlated with standardized language tests, showing that attention and inhibition play an important role in language processes.

Executive-Loaded Working Memory

Tasks associated with executive-loaded working memory functions require storage and manipulation of information at the same time. Results from complex span tasks show impaired verbal abilities between 7 and 12 years (Marton & Schwartz, 2003; Montgomery et al., 2010; Weismer, Evans, & Hesketh, 1999), while results for nonverbal abilities are controversial. In Marton’s (2008) study, preschool children with SLI and TD children did not differ in visuospatial short-term memory, but SLI children’s performance lagged behind controls on the executive-loaded visuospatial tasks. Archibald and Gathercole (2007), on the other hand, did not find any evidence for spatial short-term and working memory deficits in school-age children with SLI.

Switching (Shifting)

As explained above, problems with planning may also indicate a switching deficit, although Im-Bolter and colleagues (2006) found no shifting deficits in the SLI group. Dibbets, Bakker, and Jolles (2006) examined task-switching in a functional MRI study in the nonverbal domain. At the behavioral level, they observed no group differences, but children with SLI showed larger compensatory activations in areas associated with cognitive control.

Fluency

There are surprisingly few studies of (even verbal) fluency in SLI, and the results are controversial too. Some studies found no group differences in the number of items listed in a semantic fluency task in children between 6 and 13 (Kail & Leonard, 1986; Weyandt & Willis, 1994). In the nonverbal domain, Bishop and Norbury (2005) also failed to find a deficit in an ideational fluency task. On the other hand, Weckerly, Wulfeck, and Reilly (2001) showed both semantic and phonemic fluency impairment in SLI at school age, while performance patterns, and clustering and switching strategies were similar in the two groups. As these latter measures are more closely associated with frontal executive functions, Weckerly and colleagues (2001) account for the deficit in terms of a linguistic information processing deficit.

In light of this controversial pattern of findings, a recent comprehensive study by Henry, Messer, and Nash (2012) systematically tested verbal and nonverbal aspects of executive functions in 41 children with SLI between 8 and 14 years and TD peers. Using
10 tasks, they compared group performances while controlling for age, nonverbal, and verbal IQ. They were tested on measures of executive-loaded working memory (the listening recall subtest of Working Memory Test Battery for Children by Pickering & Gathercole, 2001 in the verbal domain, and the odd-one-out task by Henry, 2001 in the nonverbal domain), fluency (verbal and nonverbal fluency subtests of Delis-Kaplan Executive Function System [D-KEFS]; Delis, Kaplan, & Kramer, 2001), planning (verbal and nonverbal tasks on the sorting test of D-KEFS), inhibition (a new verbal inhibition, motor inhibition test with verbal and nonverbal subtests developed by the authors), and switching in the verbal (trail-making test from D-KEFS) and nonverbal (switching test was intra/extradimensional shift from Cambridge Neuropsychological Test Automated Battery; Cambridge Cognition, 2006) domain. Children with SLI lagged behind nonlanguage-impaired controls on six tasks, and performed comparable with TD children on verbal and nonverbal switching, verbal planning, and verbal inhibition.

As the above literature review shows, it is difficult to draw a clear picture of executive problems in SLI. The aim of this paper is to extend previous findings and present results from verbal and nonverbal tasks examining executive functions in Hungarian-speaking children with SLI and their age-matched TD peers. Because the concept and components of executive functions are still subjects of considerable debate in the literature (e.g., Anderson, 2002; Baddeley, 1996; Engle & Kane, 2004; Miyake et al., 2000; Smith & Jonides, 1999) and it is difficult to find tasks that are associated with only one EF subcomponent, instead of taking a theoretical stance, we rely on tasks that are associated with at least partially different aspects of EF. We designed this study to extend systematic studies of EF in SLI following Henry et al.’s (2012) logic in developing verbal and nonverbal tasks for several functions, trying to focus on tasks that build relatively little on other executive functions. Children were tested on verbal and nonverbal versions of simple and complex span tasks, fluency, N-back, and Stroop tasks. Simple span tasks were included to allow controlling for their potential contribution to EF deficits. This design allows us to examine whether (a) deficits in specific executive functions are only manifest when they are mediated by verbal stimuli, or they point to domain-general dysfunctions that are present in nonlinguistic tasks as well; (b) there is a general EF deficit in SLI, or there are selectively vulnerable specific subfunctions like inhibition or updating.

Method and Measures

Participants

Thirty-one children with SLI participated in the study (eight girls, 23 boys).1 Demographic and screening data for the two groups are shown in Table 1. They were recruited from two special school classes and two special preschool groups for children with language impairment. Children were referred to these groups and classes by speech and language therapists working in clinical practice. In each institution, recruitment took between 2 and 3 months. No eligible children declined participation. All children met inclusive and exclusive criteria for SLI that are standardly used in selecting SLI children in research (see, e.g., Dollowhan, 2007; Leonard, 1998/2014, Tager-Flusberg and Cooper, 1999). Each child scored above 85 on the Raven Colored Progressive Matrices (Raven, Court, & Raven, 1987), a measure of nonverbal intelligence. No child had a hearing impairment or a history of neurological impairment. No children in the SLI group had any known comorbidities. Each child scored at least 1.25 SDs below age norms on at least two of four language tests administered. The four tests included two receptive tests: the Hungarian standardizations of the Peabody Picture Vocabulary Test (Csányi, 1974) and the Test for Reception of Grammar (Bishop, 1983, 2012; Lukács, Győri, & Rózsa, 2012) and two expressive tests: the Hungarian Sentence Repetition Test (Magyar Mondatútámadási Teszt; Kas & Lukács, in preparation), and a nonword repetition test (Racsmány, Lukács, Németh, & Pélh, 2005). The 31 children in the control group were TD children matched on chronological age (each child in the TD group was within 3 months of age of a child in the SLI group) and nonverbal IQ (children from a larger group of age-matched TD children were only included in the control group if their IQ scores were within 5 points of their match in the SLI group). TD children were recruited from three schools and two preschools with no special selection processes for children. 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.

Simple and Complex Span Tasks

As a baseline, we included simple span tasks in our design together with complex span tasks that require concurrent storage and processing involving executive-loaded working memory. All span tasks were similar in that they contained sequences of different length, and each length was associated with four items. Sequences were presented in increasing length, and the child had to repeat at least two out of four items to proceed to the next length level. 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. Testing started with sequences of two items in all tasks; the longest possible sequence contained nine items in the simple and six items in the complex span tasks.

Simple Span: Digit Span and Corsi Blocks

In the digit span task, participants are auditorily presented with a sequence of numbers (using a computerized task), and they are asked to repeat the numbers in the same order as they heard them. The Corsi blocks task measures spatial span with the help of nine cubes in random arrangement on a tray. The experimenter touches a certain number of cubes in a given sequence, and the participant is asked to touch the same cubes in the same order.

Complex Span Tasks: Listening Recall and Odd-One-Out

In the listening recall task, participants listen to sets of short sentences. After hearing a sentence, they have to tell whether it is true or false, and at the end of each set, they are asked to recall the

---

1 Boys are systematically more vulnerable to SLI, estimates vary between 3–4:1 for boys:girls (e.g., Robinson, 1991; Cheuk et al., 2005).
Table 1
Demographic Data and Scores for Screening Tests in the SLI and TD Groups

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
<th>F</th>
<th>Sig</th>
<th>n²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>7.81</td>
<td>7.84</td>
<td>0.004</td>
<td>0.949</td>
<td>0.000</td>
</tr>
<tr>
<td>Raven IQ (standard score)</td>
<td>107.06</td>
<td>103.10</td>
<td>1.058</td>
<td>0.224</td>
<td>0.025</td>
</tr>
<tr>
<td>Nonword repetition (span)</td>
<td>5.90</td>
<td>2.97</td>
<td>&lt;0.001</td>
<td>0.507</td>
<td></td>
</tr>
<tr>
<td>Peabody Picture Vocabulary Test (raw scores)</td>
<td>110.45</td>
<td>86.39</td>
<td>&lt;0.001</td>
<td>0.216</td>
<td></td>
</tr>
<tr>
<td>Test for Reception of Grammar blocks (raw scores)</td>
<td>16.77</td>
<td>11.90</td>
<td>&lt;0.001</td>
<td>0.445</td>
<td></td>
</tr>
<tr>
<td>Sentence repetition (raw scores)</td>
<td>35.35</td>
<td>17.90</td>
<td>&lt;0.001</td>
<td>0.554</td>
<td></td>
</tr>
</tbody>
</table>

Note. STL = specific language impairment; TD = typically developing; Sig = significant. Means are highlighted in boldface.

N-Back Tasks

N-Back tasks are most strongly associated with updating, but correct performance often also relies on monitoring and inhibition. In order to minimize the role of inhibition and focus on updating in this task, we did not include lures in the design.

In the N-back tasks, participants are presented with a sequence of stimuli, and their task is to indicate (by pressing “Enter”) when the current stimulus matches the one presented n steps earlier. Stimuli were presented electronically using the E-Prime 2.0 software (Schneider et al., 2012). We used one- and two-back conditions for each stimulus type, in two blocks with about a 1-min 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).

We developed a verbal and a nonverbal version of the task. In the verbal condition, stimuli were letters (participants typically use a strategy where they rehearse the letters with their names instead of just relying on their visual shape), and in the nonverbal condition, stimuli were pictures (pictures of fractals that are difficult to verbalize). The design of the task and the instructions were the same for both types.

We calculated the number of hits (when the participant correctly presses the Enter on an “N-back trial,” i.e., when the current item is identical to the target item, with a maximum of 10 hits per block) and the number of false alarms (the participant presses Enter on a not “N-back trial,” i.e., the actual stimulus is not identical to the one presented n before) for one-back and two-back blocks separately.

Stroop Tasks

Stroop tasks are designed to tap into inhibition. We created two versions, a verbal and a nonverbal. In both versions, pictures of animals appear on the screen with a simultaneously presented auditory stimulus. The auditory stimulus was either the name of an animal (verbal condition) or the recorded sound of the animal. Stimuli were presented electronically using the E-Prime 2.0 software (Schneider et al., 2012). There were four pictures: a picture of a cow, a horse, a rooster, and a cat. In accordance, there were four animal names in the verbal, and four recorded animal sounds in the nonverbal condition. The auditory name or sound matches the picture (e.g., a picture of a cow appears and the word cow or a cow sound is heard) in the congruent condition, but does not match it in the incongruent condition (e.g., a picture of a cow and the word horse or a horse sound is heard). In the verbal control condition, there is no sound presented with the pictures, whereas participants only hear sounds in the nonverbal control condition. Participants have to press a button (marked with stickers of the animals) corresponding to the picture they see in the verbal task and the voice they hear in the nonverbal task. There are three blocks (control, incongruent, congruent) of 60 trials in the tasks. The order of the trials is random within blocks, and the three blocks also follow each other in a random order. We had two measures for both the verbal and the nonverbal version of the task. For accuracy measures, the number of correct answers for the incongruent items was subtracted from the number of correct answers for the congruent items. In the case of reaction times (RTs), we subtracted the median RT for the congruent items from the median RT for the incongruent items.

Fluency Tasks

In the verbal fluency task, children were asked to generate as many (a) actions or things that people do, (b) things they can buy at a supermarket, and (c) words starting with k as they can in 1 min for each condition. Nonverbal fluency was tested by the design fluency subtest of D-KEFS (Delis et al., 2001). This task uses a booklet containing boxes with dot patterns, and the child is asked to draw as many different designs (each in a separate box) as he or she can in 1 min, connecting the dots with four lines. In Condition A, there are only filled dots, in Condition B, boxes contain both empty and filled dots, and the task is to connect empty dots only. Condition C also contains empty
and filled dots, and the task is to connect them in an alternating sequence by the four lines. Cronbach’s alpha is 0.915 for the nonverbal task (Delis et al., 2001). For both tasks, we calculated the overall number of correct answers and the number of errors for each participant.

Results

Statistical Analysis

We had a verbal and a nonverbal version for all tasks. For each task, we conducted a $2 \times 2$ mixed-model analysis of variance (ANOVA) with type (Verbal vs. Nonverbal) as the within-subjects variable, and group (SLI vs. TD) as the between-subjects variable. Simple spans were tested to control for their potential contribution to complex tasks by including them as a covariate in case of significant group differences. Table 2 summarizes results for all tasks in the two groups, and results of one-way between group comparisons. To control for Type I errors, we divided the alpha level by the number of executive functions tested in the current study. We focused on five functions: simple span, complex span, Stroop, N-back, and fluency, hence the alpha value was set to 0.01.

Simple Span: Digit Span and Corsi Blocks

The ANOVA revealed that children had higher spans on the nonverbal than on the verbal task, as shown by a marginally significant type main effect, $F(1, 60) = 6.550$, $p = .013$, $\eta_p^2 = 0.098$. Children with SLI showed a significantly lower performance, as revealed by a main effect of group, $F(1, 60) = 10.384$, $p = .002$, $\eta_p^2 = 0.148$. The interaction of Type $\times$ Group was also significant, $F(1, 60) = 26.199$, $p < .001$, $\eta_p^2 = 0.304$.

To further analyze the Type $\times$ Group interaction, we conducted two separate univariate ANOVAs with digit span/Corsi span as a dependent variable, and group (TD vs. SLI) as between-subjects variable. The ANOVAs revealed that the control group outperformed the clinical group on the digit span task, $F(1, 60) = 34.055$, $p < .001$, $\eta_p^2 = 0.362$, but not on the nonverbal Corsi blocks task ($p = .892$). Because there were no group differences on Corsi span, it was not included as a covariate in group comparisons of nonverbal tasks below.

Complex Span: Listening Span and the Odd-One-Out

The $2 \times 2$ mixed-model ANOVA revealed a significant main effect of type, $F(1, 60) = 51.921$, $p < .001$, $\eta_p^2 = 0.464$, with generally higher performance on the nonverbal odd-one-out task. The TD group outperformed the SLI group, $F(1, 60) = 8.961$, $p = .004$, $\eta_p^2 = 0.130$. The Type $\times$ Group interaction was also approaching significance, $F(1, 60) = 8.961$, $p = .081$, $\eta_p^2 = 0.050$.

To further analyze the interaction, we conducted a separate ANOVA for each task with group as the between-subjects variable. The ANOVAs revealed that performance of the TD group was significantly higher on the verbal listening span task, $F(1, 60) = 17.024$, $p < .001$, $\eta_p^2 = 0.221$, but not on the nonverbal odd-one-out task ($p = .144$).

To investigate whether the group difference in the verbal task is due to limitations in simple span, we reran the above analysis on listening span with digit span as covariate. This way, the main effect of group was not significant ($p = .109$).

Table 2

<table>
<thead>
<tr>
<th>Measures</th>
<th>TD Mean</th>
<th>TD SD</th>
<th>TD F</th>
<th>TD Sig</th>
<th>TD $\eta_p^2$</th>
<th>SLI Mean</th>
<th>SLI SD</th>
<th>SLI F</th>
<th>SLI Sig</th>
<th>SLI $\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit span</td>
<td>4.58</td>
<td>0.92</td>
<td>3.39</td>
<td>0.67</td>
<td>34.055</td>
<td>&lt;.001</td>
<td>0.362</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corsi span</td>
<td>4.29</td>
<td>1.01</td>
<td>4.26</td>
<td>0.86</td>
<td>0.018</td>
<td>0.395</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening recall</td>
<td>2.23</td>
<td>0.92</td>
<td>1.29</td>
<td>0.86</td>
<td>17.024</td>
<td>&lt;.001</td>
<td>0.362</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odd-one-out</td>
<td>2.97</td>
<td>1.30</td>
<td>2.52</td>
<td>1.09</td>
<td>2.188</td>
<td>0.144</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy verbal Stroop</td>
<td>-3.00</td>
<td>6.76</td>
<td>-5.61</td>
<td>13.08</td>
<td>0.976</td>
<td>0.327</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy nonverbal Stroop</td>
<td>-8.55</td>
<td>15.37</td>
<td>-11.29</td>
<td>16.18</td>
<td>0.468</td>
<td>0.497</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT verbal Stroop</td>
<td>324</td>
<td>290</td>
<td>306</td>
<td>359</td>
<td>0.048</td>
<td>0.827</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT nonverbal Stroop</td>
<td>397</td>
<td>371</td>
<td>317</td>
<td>337</td>
<td>0.795</td>
<td>0.376</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal one-back hit</td>
<td>8.84</td>
<td>1.83</td>
<td>8.35</td>
<td>2.39</td>
<td>0.803</td>
<td>0.374</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonverbal one-back hit</td>
<td>9.23</td>
<td>1.12</td>
<td>8.90</td>
<td>1.70</td>
<td>0.780</td>
<td>0.381</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal one-back false alarm</td>
<td>0.48</td>
<td>1.15</td>
<td>2.16</td>
<td>4.75</td>
<td>3.655</td>
<td>0.061</td>
<td>0.057</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonverbal one-back false alarm</td>
<td>3.81</td>
<td>7.91</td>
<td>3.81</td>
<td>8.78</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal two-back hit</td>
<td>5.48</td>
<td>2.63</td>
<td>3.97</td>
<td>2.70</td>
<td>5.010</td>
<td>0.029</td>
<td>0.077</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonverbal two-back hit</td>
<td>4.65</td>
<td>2.67</td>
<td>4.58</td>
<td>2.87</td>
<td>0.008</td>
<td>0.927</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal two-back false alarm</td>
<td>1.39</td>
<td>1.99</td>
<td>2.87</td>
<td>5.99</td>
<td>1.714</td>
<td>0.195</td>
<td>0.028</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonverbal two-back false alarm</td>
<td>2.90</td>
<td>3.00</td>
<td>7.32</td>
<td>11.67</td>
<td>4.172</td>
<td>0.045</td>
<td>0.065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal fluency correct</td>
<td>31.55</td>
<td>12.26</td>
<td>22.45</td>
<td>11.16</td>
<td>9.329</td>
<td>0.003</td>
<td>0.135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonverbal fluency correct</td>
<td>8.84</td>
<td>5.01</td>
<td>10.10</td>
<td>6.45</td>
<td>0.735</td>
<td>0.395</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal fluency errors</td>
<td>1.48</td>
<td>2.79</td>
<td>2.19</td>
<td>2.63</td>
<td>1.063</td>
<td>0.307</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonverbal fluency errors</td>
<td>7.48</td>
<td>5.53</td>
<td>6.06</td>
<td>5.46</td>
<td>1.033</td>
<td>0.314</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. TD = typically developing; SLI = specific language impairment; Sig = significance; RT = reaction time. Calculation method for each measure is described in the text.
Stroop Tasks
The ANOVA revealed no significant effects on either accuracy (number of correct answers for the incongruent items subtracted from number of correct answers for the congruent items; p = .033 for the type main effect, p = .233 for the group main effect and p = .980 for the Type × Group interaction) or RT measures (median RT for congruent items subtracted from the median RT for incongruent items in milliseconds; all p values > .436).

N-Back Tasks
In the case of one-back condition, no effects were significant for hit rate: all p values >.111. Both groups of children produced a significantly higher number of false alarms in the nonverbal task, as revealed by a main effect of type, \( F(1, 60) = 8.059, p = .006, \eta^2_g = 0.118 \). Neither the main effect of group (p = .542), nor the Type × Group interaction was significant (p = .342).

In the two-back condition, no effects were significant for accuracy measures (p = .760 for type, p = .181 for group, and p = .053 for the Type × Group interaction).

The number of false alarms was significantly higher in the nonverbal two-back task, \( F(1, 60) = 13.084, p < .001, \eta^2_g = 0.179 \). Results revealed no significant differences by group, \( F(1, 60) = 3.784, p = .056, \eta^2_g = 0.059 \), and no Type × Group interaction either, \( F(1, 60) = 3.166, p = .080, \eta^2_g = 0.050 \).

Fluency
Fluency results were analyzed on composite measures calculated by summing up the three measures for both the verbal and the nonverbal task. We calculated the sum of correct answers, the sum of incorrect answers, and the sum of repetitions.

In the case of correct answers, the ANOVA revealed a significant main effect of type, \( F(1, 60) = 153.330, p < .001, \eta^2_g = 0.719 \), with significantly higher number of correct responses on the verbal task. There was also a significant interaction between type and group, \( F(1, 60) = 13.371, p < .001, \eta^2_g = 0.182 \), while the main effect of group was not significant, \( F(1, 60) = 4.380, p = .041, \eta^2_g = 0.068 \).

Further post hoc ANOVAs showed that there was a significant difference between the groups on the verbal, \( F(1, 60) = 9.329, p = .003, \eta^2_g = 0.135 \), but not on the nonverbal task (p = .395). The former was eliminated by simple span as covariate (p = .505 for the group main effect).

In the case of category errors, there was a significant main effect of type, \( F(1, 60) = 46.770, p < .001, \eta^2_g = 0.438 \), with a higher number of errors in the nonverbal task. Neither the group main effect (p = .671), nor the Type × Group interaction (p = .145) was significant. No effects were significant in the case of repetition errors: all p values >.273.

Discussion
Extensive examination of executive functions in a group of children with SLI revealed impairments in some, but not all executive functions, mostly in the verbal domain. The SLI group showed difficulties in verbal complex span (Listening span task), and fluency, but not in inhibition (Stroop tasks) and updating (N-back tasks) relative to TD age-matched children. While group differences were observed in initial analyses on these two verbal tasks, including measures of simple verbal span (digit span) as a covariate eliminated them, suggesting that fundamental difficulties in short-term memory (STM) contribute to difficulties in verbal complex working memory and other executive functions (e.g., fluency). Although difficulties were most evident in the verbal domain, they were also observed in simple group comparisons in one measure in the nonverbal domain: children with SLI made a higher number of false alarms in the nonverbal N-back task. This result suggests that in spite of the fact that we did not find difficulties with inhibition in a direct test of inhibition (the Stroop task), with increasing task difficulty and higher working memory demand, these might become evident too.

Our results extend previous studies on executive function in SLI in important ways. Examining EF in everyday situations (Hughes et al., 2009; Wittke et al., 2013) suggests deficits in different areas of executive functions, and experimental studies of attention (Dis-paldo et al., 2013; Finnner et al., 2009; Spaulding et al., 2008) point to a deficit in SLI despite testing different groups and using different tasks. As our review in the introduction illustrates, results on EF in other areas are more controversial in this relatively new area of research, and perhaps as more studies are conducted with larger and well-defined groups in different age-ranges, we will gain a better picture. The current study was a step in that direction. Our findings show that deficient performance on verbal EF tasks can in fact be a secondary phenomenon rooted in verbal short-term memory impairment, and also suggest that problems with executive functions might only become apparent when the task is complex and involves higher working memory load as well as engaging other executive functions.

Results presented in the paper are not easy to integrate with the earlier set of findings in the literature. Because these findings were controversial and used different tasks and age groups in most areas, instead of reiterating them in the light of our own findings, we try to speculate on the reasons behind such a mixed picture. First, it is well-known that SLI is a category that includes children of different age, different patterns of symptoms, different severity of impairment, and potentially of different etiology too. Language impairment is often associated with other developmental disorders, most frequently with attention-deficit/hyperactivity disorder, dyslexia, and autism spectrum disorder (for a review, see Leonard, 1998/2014), all of which are known to involve impairments in several executive functions and working memory. Although the definition of SLI excludes some of these associated problems, the boundaries are not clear-cut, and this results in a lot of heterogeneity across samples and studies. Second, studies of executive functions in SLI use different tasks that differ in their complexity and difficulty: sometimes a lack of a group difference could be due to applying a task that is very easy even for children with SLI or too difficult even for TD children (e.g., a three-back task).

Third, most tasks involve more than just one EF subfunction. For example, although the N-back task is often used to test updating, to successfully respond, participants also need to inhibit responses to distractor stimuli. Even the subfunctions themselves are often dependent on each other, and may differ greatly in their complexity: planning and shifting, for example, seem to essentially build on inhibition and updating. The fuzziness of the theoretical constructs (both of SLI and executive functions) makes it difficult to establish the nature of the deficits. As all three factors are
general problems in the area of research on SLI on the one hand, and on EF on the other, they are among the limitations of our study as well.

Also, as pointed out by one of the reviewers of the manuscript, childhood socioeconomic status (SES) influences EF (see, e.g., Hackman, Gallop, Evans, & Farah, 2015). Children with SLI were not matched to TD peers on SES in our study. Although this is not an issue that is usually addressed in studies of children with SLI, this could be a concern, but we have no reason to believe that there were significant differences between children with SLI and TD children in this regard (e.g., schools and preschools were in similar neighborhoods in both groups). Besides taking SES into account, future research should focus on larger and more homogeneous groups of children with language impairment, as well as try and tease apart EF subcomponents more effectively.

Conclusion

Children with SLI were found to be impaired on several verbal measures of EF, but these differences were largely due to more fundamental deficits in verbal short-term span. In the nonverbal domain, inhibition deficits were only present when the task involved a high working memory load. Future studies should explore the exact nature of deficits in nonverbal EFs in SLI. Also, when verbal and nonverbal functions seem to be affected in SLI, it is important to examine whether they contribute to language deficits, or are only associated with them. Our pattern of findings together with earlier results suggests that diagnosis and therapy of SLI should also consider potential limitations in executive functions.

References


Received August 27, 2014
Revision received May 27, 2015
Accepted June 22, 2015
**Study 2.**

Lexical conflict resolution in children with specific language impairment

Enikő Ladányia,b,*, Ágnes Lukácsb

a Paris Descartes University, Laboratoire Psychologie de la Perception, 45 Rue des Saints-Peres, 75006 Paris, France
b BME Department of Cognitive Science, Budapest University of Technology and Economics, Egry József u. 1, H-1111 Budapest, Hungary

ARTICLE INFO

Article history:
Received 26 January 2015
Received in revised form 6 April 2016
Accepted 6 April 2016
Available online 12 April 2016

Keywords:
Specific language impairment
Word retrieval
Conflict resolution
Cognitive control

ABSTRACT

The aim of our study is to examine the effect of conflict on naming latencies in children with specific language impairment (SLI) and typically developing (TD) children and to explore whether deficits in conflict resolution contribute to lexical problems in SLI. In light of previous results showing difficulties with inhibitory functions in SLI, we expected higher semantic conflict effect in the SLI than in the TD group. To investigate this question 13 children with SLI and 13 age- and gender-matched TD children performed a picture naming task in which the level of conflict was manipulated and naming latencies were measured. Children took longer to name pictures in high conflict conditions than in low conflict conditions. This effect was equally present in the SLI and TD groups. Our results suggest that word production is more effortful for children when conflict resolution is required but children with SLI manage competing lexical representations as efficiently as TD children. This result contradicts studies, which found difficulties with inhibitory functions and is in line with findings of intact inhibitory abilities in children with SLI. Further studies should rule out the possibility that in SLI lower level of conflict resulting from weaker lexical representations masks impairments in inhibition, and investigate the effect of linguistic conflict in other areas.

ã 2016 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Lexical impairments in specific language impairment

Children with specific language impairment (SLI) show linguistic deficits that are not accounted for by obvious impairments in other cognitive domains. Usually morphosyntactic and syntactic problems are emphasized (e.g.: Bishop, 1997; Leonard, 1998/2014) but lexical impairments are reported as well. Several studies show that first words appear later in children with SLI than in typically developing children (TD) and their vocabulary size lags behind age-based expectations at older ages too (Bishop, 1997; Watkins, Kelly, Harbers, & Hollis, 1995; Trauner, Wulfeck, Tallal, & Hesselink, 1995). SLI can also appear later without early vocabulary deficits, and early vocabulary problems do not always lead to language impairment later (Henrichs et al., 2011; Poll & Miller, 2013; Rescorla, 2011; see Ellis Weismer, 2007 and Leonard, 1998/2014 for a review).

* Corresponding author at: Paris Descartes University, Laboratoire Psychologie de la Perception, 45, Rue des Saints-Peres, 75006 Paris, France.
E-mail addresses: eladanyi@cogsci.bme.hu (E. Ladányi), alukacs@cogsci.bme.hu (Á. Lukács).
Children with SLI are less efficient than TD peers in experimental word learning tasks as well (e.g. Rice, Buhr, & Nemeth, 1990; Rice, Buhr, & Oetting, 1992; Rice, Oetting, Marquis, Bode, & Pae, 1994). Word retrieval seems problematic in picture naming tasks: they make more errors than TD children even when they know the word (Kail & Leonard, 1986; Mcgregor & Leonard, 1995), and they have longer naming latencies (Anderson, 1965; Ceci, 1983; Kail & Leonard, 1986; Katz, Curtiss, & Tallal, 1992; Lahey & Edwards, 1996; Leonard, Nippold, Kail, & Hale, 1983; Miller et al., 2001; Wiig, Semel, & Nystrom, 1982; Windsor & Hwang, 1999a). Difficulties with lexical processing also appear: children with SLI show more errors (Crosbie, Howard, & Dodd, 2004) and longer reaction times in lexical decision (Edwards & Lahey, 1996; Windsor & Hwang, 1999b) and in word monitoring tasks (Montgomery & Leonard, 1998; Montgomery, Scudder, & Moore, 1990; Stark & Montgomery, 1995). Taken together, these findings show that the acquisition, production and processing of words are slower and more error-prone in SLI than in typical development (see also Leonard & Deevy, 2004).

1.2. Accounts of lexical impairments in SLI

As the above review suggests, lexical deficits are present in various forms in the language of children with SLI, and there is no agreement in the literature on the potential causes of these impairments. A possible explanation is that lexical problems are caused by differences in the features of lexical representations. Several studies argue that in language impairment, less information is available about lexical items and the items are not as well-organized as they are in the mental lexicons of typically developing children (Kail & Leonard, 1986; Lahey & Edwards, 1999; Mcgregor, Newman, Reilly, & Capone, 2002; Sheng & Mcgregor, 2010). Mcmurray, Samelson, Lee and Tomblin (2010) attributes language impairments to faster lexical decay after word retrieval in these populations.

Others relate lexical problems to non-linguistic impairments: several earlier studies suggest that a generally slower processing contributes to lexical impairments as well (Kail, 1994; Leonard et al., 2007; Miller et al., 2001; Windsor & Hwang, 1999c). Lahey and Edwards (1996) argue that children name pictures more slowly because their non-linguistic response processes are impaired. Non-linguistic problems are suggested to contribute to the processing of ambiguous words by Norbury (2005), who proposes that less effective suppression mechanisms can contribute to weaker performance of children with SLI and she also emphasizes deficits of memory and attention skills as likely factors at play. Mainela-Arnold, Evans and Coady (2008) and Mainela-Arnold, Evans and Coady (2010) argue, based on findings from a gating task, that top-down attentional processes are impaired in children with SLI. In a gating task participants are presented with increasingly longer chunks of words starting from their beginning, and they are asked to guess the word after each trial. Children with SLI showed similar performance as TD children with one difference: they produced competing alternatives even after they found the appropriate word. According to the authors, this pattern is the result of weaker representations that are more vulnerable to lexical competition or, alternatively, it is caused by the deficit of top-down competition-resolving processes, or by a combination of the two (Mainela-Arnold et al., 2008).

Mainela-Arnold et al., (2010) also showed that performance of children with SLI lags behind TD children both in a word definition task and in a delayed word repetition task, with positive correlations between the two performance measures. They propose that because of reduced attentional capacity — reflected by impaired performance on the delayed repetition task — children with SLI have weak phonological representations which have a negative effect on semantic representations as well. Impairments of higher order top-down processes in word retrieval can also be linked to a new line of research in the psycholinguistic literature that emphasizes the role of cognitive control functions in linguistic processes, which we review below.

1.3. Competition and the role of cognitive control in word retrieval

According to recent studies, cognitive control, i.e. the ability to orchestrate our actions and thoughts with our internal goals (Miller & Cohen, 2001) is necessary for language use in many areas including syntactic ambiguity resolution (January, 1992; Lahey & Edwards, 1996; Leonard, Nippold, Kail, & Hale, 1983; Miller et al., 2001; Wiig, Semel, & Nystrom, 1982; Windsor & Hwang, 1999a). Difficulties with lexical processing also appear: children with SLI show more errors (Crosbie, Howard, & Dodd, 2004) and longer reaction times in lexical decision (Edwards & Lahey, 1996; Windsor & Hwang, 1999b) and in word monitoring tasks (Montgomery & Leonard, 1998; Montgomery, Scudder, & Moore, 1990; Stark & Montgomery, 1995). Taken together, these findings show that the acquisition, production and processing of words are slower and more error-prone in SLI than in typical development (see also Leonard & Deevy, 2004).
for the mixed block and a smaller drop for homogeneous blocks and this pattern remains similar in the third and the fourth cycles (e.g., Belke et al., 2005; Biegler, Crowther, & Martin, 2008; Navarrete, Del Prato, & Mahon, 2012; Schnur et al., 2006).

Several explanations have been proposed to account for the patterns found in the blocked cyclic naming task. Selection-by-competition accounts (e.g., Belke et al., 2005) state that in mixed blocks, reaction times become lower from the second cycle because of the activation of the same word representations during the previous cycle(s) (repetition priming) and this effect is smaller in homogeneous blocks because of the co-activation of several semantically similar representations. In the homogeneous condition, the target word requires extra time to reach the critical difference threshold, which leads to longer naming latencies. Schnur et al. (2006) suggest that cognitive control might also have a role in the process. They argue that a high level of competition leads to conflict between the target word and competing words, constituting a signal that engages cognitive control mechanisms. Longer reaction times from the second cycle result, at least partly, from the time needed for this control mechanism to take effect. The formulation of hypotheses and the experimental design in the current study was motivated by the findings and theoretical framework of Schnur et al. (2006); following their terminology, we will use the word ‘conflict’ to refer to situations when multiple representations are expected to be activated to a similar degree. We will refer to the increase in naming latencies in conditions with high level of competition relative to low level of competition as the ‘conflict effect’ throughout the paper.

Oppenheim, Dell, and Schwartz (2010), on the other hand, suggest a different account of the semantic blocking effect. They take word production to be an error-based implicit learning process resulting in incremental changes in the connection weights between semantic features and word representations. Due to this process, semantic-lexical connections are strengthened for the selected word and weakened for non-selected but related words during word production. For instance, the selection of table for naming the picture of a table strengthens the links between the semantic features of a table (e.g., made of wood, has four legs, used for eating or working) and the word table, but also leads to a parallel weakening of links between those same semantic features and other semantically related words which were not selected (e.g., chair, tablecloth). Therefore it yields a decrease in reaction times from the second cycle due to the strengthened connections but in semantically mixed blocks reaction times should show a smaller decrease because of the continuous weakening of connection strengths of every word except the actual target word. For instances of high interference, the theory proposes a booster mechanism which amplifies the activation of each word until a winner can be selected. The model would predict the presence of a semantic blocking effect already in the first cycle, the lack of which is explained by conscious strategies applied by the participant, according to Oppenheim et al. (2010). Crowther and Martin (2014) argues (following the proposition of Belke (2013) and Belke and Stielow (2013)) that the model should be supplemented by a top-down control mechanism biasing the activation of items within the response set. This idea is supported by their results showing correlations between the size of the semantic blocking and Stroop effects (Crowther & Martin, 2014). Thus we can conclude that both of these theories are compatible with the idea of cognitive control processes involved during the naming of homogeneous blocks of the semantic blocking paradigm.

The most popular picture naming paradigm for the manipulation of semantic competition is the abovementioned semantic blocking paradigm, but important results were found with a task in which participants are required to name pictures with low vs. high name agreement (Kan & Thompson-Schill, 2004; Novick, Kan, Trueswell, & Thompson-Schill, 2009). Name agreement is determined by the number of names available for describing a picture. High name agreement pictures can be named with one word only, or even when there are multiple available names, one of them is used a lot more frequently than the others (a picture of an apple is usually named as an apple). In the case of low agreement pictures there are more than one available names, with similar probabilities of use (a picture of a stove can be named as stove, oven or range). In the latter case conflict is expected to be higher while the alternatives are still competing with the target word for selection. Kan and Thompson-Schill (2004) as well as Novick et al. (2009) assumes that a cognitive control mechanism is responsible for biasing the selection in these cases.

In sum, studies with healthy adults showed that pictures with high conflict are named significantly slower than pictures with low conflict. The involvement of cognitive control in these conflict effects are supported by correlations between word retrieval and cognitive control measures and an association of higher conflict with an increased level of activation in the left inferior frontal gyrus, an area usually active during other tasks involving cognitive control (Kan and Thompson-Schill, 2004; Schnur et al., 2006). Furthermore, patients with aphasia with a left inferior frontal gyrus impairment take longer or even fail to produce pictures with high conflict (Biegler et al., 2008; Novick et al., 2009; Schnur et al., 2006, 2009). These results suggest that cognitive control has a critical role in successful word selection when more representations are activated in healthy adults, and impairments of cognitive control can contribute to word retrieval difficulties in patients with aphasia.

To our knowledge, no studies explored the effect of conflict manipulations during picture naming either in typically developing children, or in children with language impairment so far. Beyond its theoretical importance, the question has potential clinical relevance as well: shedding light on the specific sources of lexical problems in children with SLI enables the development of targeted trainings. If general cognitive control abilities are impaired in children with SLI, contributing to deficits in the language domain, these abilities should also be the focus of therapy beyond language abilities.

1.4. The current study

Motivated by the above findings on the role of cognitive control in word retrieval in healthy adults and in patients with aphasia, our aim in the current study employing a picture naming task was to test the hypothesis that cognitive control
problems in children with SLI contribute to their word retrieval difficulties. Since we have not found any developmental results on this question in the literature, we also aimed to explore the relationship between cognitive control and lexical conflict resolution in TD children. Our hypothesis regarding the role of cognitive control impairments in SLI was motivated by three sets of previous findings: 1) word retrieval problems observed in SLI (as reviewed above) 2) cognitive control impairments contributing to word retrieval difficulties in adults (also reviewed above) and 3) problems with cognitive control observed in children with SLI (e.g. Finneran, Francis & Leonard, 2009; Henry, Messer, & Nash, 2012; Lum & Bavin, 2007; Seiger-Gardner & Brooks, 2008; Spaulding, 2010).

Based on earlier results from adults and patients with aphasia, we expected a decrease in reaction times through cycles both in homogeneous and mixed blocks because of the repetition priming effect, with a smaller decrease in the homogeneous blocks due to the conflict effects. Considering the name agreement manipulation, we expected longer reaction times in the case of pictures with low name agreement than with high name agreement. Overall, generally higher reaction times were expected in the case of high conflict conditions (homogeneous blocks, low name agreement) than low conflict conditions (mixed blocks, high name agreement). Since cognitive control develops well into adolescence (Davidson, Amso, Anderson, & Diamond, 2006), these effects of conflict could be even stronger in children than in adults. Also, we expected that if children with SLI have problems with cognitive control, conflict effects are going to be stronger and manifest in longer reaction times for high conflict conditions in SLI than in TD.

The last prediction, however, can be modulated by differences of facilitatory and inhibitory processes in children with SLI relative to TD children. As it was mentioned in the introduction, the organization of the lexicon and the connections between lexical representations can be different in children with SLI and TD children. These differences can be conceptualized in different ways by the selection-by-competition and the incremental learning accounts.

On the basis of the selection-by-competition account, a relevant difference is expected in the connections between elements of the mental lexicon. First, as it was discussed in the first section of the paper, these connections between semantic nodes and word representations can be weaker in children with SLI leading to generally slower reaction times during all conditions. Second, connections between semantically related elements can be weaker resulting in slower spreading activation which would lead to smaller conflict effects, i.e. smaller difference between homogeneous and mixed conditions as well as smaller increase in the effect through the cycles of homogeneous blocks. Therefore even if cognitive control abilities are impaired, the smaller level of conflict could be handled by the weaker cognitive control abilities leading to conflict effects comparable to those observed in TD children. Third, the difference between connections both between semantic nodes and word representations on the one hand and between semantically related elements on the other can have consequences on the name agreement manipulations as well in the SLI group. When a picture with multiple potential names is named for the first time with the target word, competing names are activated due to their relationships with the semantic nodes as well as to their relationships with the target word. But once the target word is selected, the competing alternatives do not activate to a high degree again due to the weaker connections, which would yield smaller or missing name agreement effects through the later cycles. And fourth, effects caused by name agreement differences can be generally smaller in children with SLI because they do not have strong relationships between semantic nodes and alternative names (they use the word couch for naming the couch, competing alternatives, like squab or settee might be less activated than in TD children).

The incremental learning account also predicts differences between the TD and SLI groups in lexical organization and in the connections between lexical representations that are expected to manifest in the blocked cyclic naming paradigm. Weight adjustments may not be as efficient in children with SLI as in TD children, predicting similar patterns as the selection-by-competition account by proposing weaker relationships between elements. It is also possible that the booster mechanism is less efficient in children with SLI, which would lead to longer reaction times in the case of semantically homogeneous blocks in that group relative to the TD group. The incremental learning account, however, does not have explicit theoretical predictions for the name agreement manipulations.

### Table 1

Means (and standard deviations) for demographic data and scores for screening tests in the SLI and TD groups. Results from one-way ANOVAs are shown for group differences.

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
<th>F</th>
<th>Sig</th>
<th>(\eta_p^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.9</td>
<td>1.4</td>
<td>8.10</td>
<td>1.4</td>
<td>0.16</td>
</tr>
<tr>
<td>Raven IQ (standardized scores)</td>
<td>108.54</td>
<td>9.71</td>
<td>100.54</td>
<td>5.83</td>
<td>6.473</td>
</tr>
<tr>
<td>Nonword repetition (raw scores)</td>
<td>6.23</td>
<td>1.01</td>
<td>3</td>
<td>1.47</td>
<td>42.506</td>
</tr>
<tr>
<td>PPVT (raw scores)</td>
<td>124.5</td>
<td>13.75</td>
<td>97</td>
<td>20.86</td>
<td>15.792</td>
</tr>
<tr>
<td>TROG blocks (raw scores)</td>
<td>18.38</td>
<td>1.45</td>
<td>13.15</td>
<td>2.67</td>
<td>38.533</td>
</tr>
<tr>
<td>Sentence repetition (raw scores)</td>
<td>37.38</td>
<td>3.23</td>
<td>19.76</td>
<td>8.21</td>
<td>51.862</td>
</tr>
</tbody>
</table>
2. Methods

2.1. Participants

Twenty-six Hungarian-speaking children participated in our study. The SLI group consisted of thirteen children (4 girls, 9 boys) who were selected from a special school for children with language impairments. Their mean age was 8;10 with a standard deviation of 1;3. Only children without hearing or neurological impairments and with normal intelligence (performance above 85 scores on Raven Colored Progressive Matrices; Raven, Court, & Raven, 1987) were screened for inclusion in the SLI group. All children were included based on criteria that are commonly used and represent accepted practice in selecting children with SLI in research (see e.g. Leonard, 1998/2014, Tager-Flusberg & Cooper, 1999): linguistic abilities were assessed with four tests and children who performed at least 1.5 SD below age-based expectations on at least two out of the four tests were included in the SLI group. These four tests consisted of two receptive and two expressive tests. The receptive tests were the Hungarian versions of the Peabody Picture Vocabulary Test (PPVT: Csányi, 1974; Dunn & Dunn, 1981) and the Test for Reception of Grammar (TROG: Bishop, 1983; Lukács, Györi, & Rózsa, 2012). The expressive tests were the Hungarian Sentence Repetition Test (Kas & Lukács, in preparation), and a nonword repetition test (Racsmány, Lukács, Németh, & Pléh, 2005). All children meeting these criteria were included in the study. No eligible children declined participation.

Typically developing children were matched individually to children in the SLI group on chronological age and gender. 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.

2.2. Stimuli

We used 36 line drawings of common objects well-known to children at this age from the picture set used by the norming study of Bates et al. (2003). Bates and her colleagues selected 520 pictures from various databases for their study. The pictures were comparable in picture quality, visual complexity, and potential cross-cultural validity of the depicted item. The pictures were normed in a picture naming study with adults in seven languages (also in Hungarian) measuring naming latency, name agreement (defined as the proportion of using one dominant name from all valid names) and various features of the dominant response (frequency, length, complexity – monomorphemic vs. plural/compound . . . ) (Bates et al., 2003). The 36 pictures were selected for the study to manipulate competition during naming both with varying the semantic context (based on Schnur et al., 2006) of the pictures and the name agreement of a picture (based on Kan & Thompson-Schill, 2004).

Pictures were taken from six semantic categories with six exemplars in each; half of these pictures had low name agreement (i.e. had more than one similarly plausible names) and the other half had high name agreement (i.e. had one dominant name). Data for name agreement for the pictures were available in Hungarian from the study of Bates et al. (2003) who published their data online (http://crl.ucsd.edu/experiments/ipnp/1database.html). (The design that determined picture selection is shown in Fig. 1.). Frequencies of the target names were similar in the high and low name agreement conditions (F(1,34) = 0.001, p = 0.97, ηp² = 0.000; log frequencies ranged from 2.26 to 5.35 (mean 4.39, SD 4.7)); based on a Hungarian frequency dictionary (http://szotar.mokk.bme.hu/szoszablya/searchq.php; Halácsy et al., 2004; Kornai, Halácsy, Nagy, Trón, & Varga, 2006). The length of words varied between one and four syllables in each condition. Because of having to control several factors at the same time, we ended up with a word set containing words with higher mean length in the low agreement condition (mean: 2.5 syllables, SD: 0.99) than in the high agreement condition (mean: 1.9 syllables, SD: 0.8). We decided to loosen this specific criteria instead of other ones because according to Bates et al. (2003), in naming, word length effects on naming latency are confounded with other factors (with word frequency or name agreement) and length has very small or no effect independently. Kawamoto, Liu, Mura, and Sanchez (2008) found longer naming latencies for words starting

---

**Fig. 1.** Pictures were selected from six semantic categories with six pictures from each category, therefore 36 pictures were used in the task. These pictures were presented in two blocks: once together with members of the same category and once with members of other categories. Therefore altogether 12 blocks of pictures were presented. Half of the 36 pictures had a name with low name agreement and the other half had a name with high name agreement.
with a plosive than for other initial phonemes; in the current study number words starting with a plosive was equal in the low and high name agreement conditions.

Each picture was presented once together with pictures from the same semantic category (semantically homogeneous blocks) and once together with pictures from various semantic categories (semantically mixed blocks). One block consisted of six pictures, which were repeated four times (in four cycles) each time with a different order of pictures yielding 24 items in each block (Examples of target answers for the pictures in a homogeneous and a mixed block are given in the Appendix, together with the names of categories from which homogeneous blocks were generated). Four cycles were included to raise the level of conflict during homogeneous cycles. Altogether six homogeneous and six mixed blocks were presented for the children. Due to the name agreement manipulations, half of the pictures belonged to the low agreement condition with several equally plausible names (e.g. a picture of a couch can be named in Hungarian with the following words: szőfa ‘sofa’, kanapé ‘couch’, diván ‘divan’, ágy ‘bed’) and the other half was a high agreement picture with only one plausible name (e.g. a picture of an apple in Hungarian is almost always named with the word alma ‘apple’).

2.3. Procedure

We used the E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2012) for presenting the stimuli and for collecting data. Reaction times were measured with a microphone that triggered a voice key. Answers were coded as ‘correct’, ‘incorrect’ or ‘technical error’ by the experimenter on paper. Before the experiment, children were instructed to name each picture and asked not to say anything else except the names of the pictures. Pictures were presented on a computer screen and remained on the screen until the child gave a response. Within a block, pictures followed each other with a one second pause between them; between blocks, children could take a break as long as they needed. The order of the 12 blocks was randomized across participants. The task lasted for approximately 20 min.

3. Results

Results were analyzed using SPSS (SPSS, 2009), version 18.0. We conducted a $2 \times 2 \times 2 \times 2$ repeated measures analysis of variance with Group (SLI vs. TD) as between-subject variable and Homogeneity (Homogeneous vs. Mixed), Agreement (Low vs. High) and Cycle (1 vs. 4) as within-subject variables. Note that for investigating the effect of Cycle we included reaction times only for the first and the fourth cycles in the analysis (Means and standard deviations for reaction times for the four cycles by Group and conditions are shown in Table 2.). Only reaction times for correct answers (names which are plausible for the picture) were included; trials where the voice key was triggered inappropriately or was not triggered because the answer was too quiet were also excluded, as well as reaction times under 300 ms and above 3000 ms. After the exclusion of RTs based on these criteria, 87% of all trials were included in the analysis.

The ANOVA showed a significant main effect of Agreement ($F(1,24)=13.845$, $p<0.001$, $\eta^2_{p}=0.366$); low agreement pictures took longer to name than high agreement pictures. All other main effects were nonsignificant (Homogeneity: $F(1,24)=1.162$, $p=0.242$, $\eta^2_{p}=0.046$; Cycle: $F(1,24)=0.28$, $p=0.675$, $\eta^2_{p}=0.001$; Group: $F(1,24)=0.105$, $p=0.794$, $\eta^2_{p}=0.004$). A significant interaction appeared between Homogeneity, Agreement and Group ($F(1,24)=8.841$, $p=0.007$, $\eta^2_{p}=0.269$), Homogeneity and Cycle ($F(1,24)=27.079$, $p<0.001$, $\eta^2_{p}=0.530$), and Agreement, Cycle and Group ($F(1,24)=5.092$, $p=0.033$, $\eta^2_{p}=0.175$). All other interactions were nonsignificant (Homogeneity x Group: $F(1,24)=0.364$, $p=0.552$, $\eta^2_{p}=0.015$; Agreement x Group: $F(1,24)=0.157$, $p=0.695$, $\eta^2_{p}=0.007$; Cycle x Group: $F(1,24)=1.112$, $p=0.302$, $\eta^2_{p}=0.044$; Homogeneity x Agreement: $F(1,24)=0.906$, $p=0.351$, $\eta^2_{p}=0.036$; Homogeneity x Cycle: $F(1,24)=1.038$, $p=0.318$, $\eta^2_{p}=0.041$; Agreement x Cycle: $F(1,24)=3.349$, $p=0.080$, $\eta^2_{p}=0.122$).

For breaking down the Homogeneity x Agreement x Group interaction, we analyzed the effect of Homogeneity and Agreement in $2 \times 2$ repeated measures ANOVAs separately for the two groups. In the SLI group the main effect of Agreement was significant ($F(1,12)=5.076$, $p=0.044$, $\eta^2_{p}=0.297$): low agreement pictures were named significantly slower than high

### Table 2

Mean reaction times (and standard deviations) in milliseconds in the picture naming task by Group and by conditions.

<table>
<thead>
<tr>
<th></th>
<th>High name agreement</th>
<th>Low name agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SLI</td>
<td>TD</td>
</tr>
<tr>
<td>Homogeneous cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>1095</td>
<td>1087</td>
</tr>
<tr>
<td>2.</td>
<td>1089</td>
<td>1183</td>
</tr>
<tr>
<td>3.</td>
<td>1274</td>
<td>1165</td>
</tr>
<tr>
<td>4.</td>
<td>1249</td>
<td>1170</td>
</tr>
<tr>
<td>Mixed cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>1167</td>
<td>1126</td>
</tr>
<tr>
<td>2.</td>
<td>1114</td>
<td>1062</td>
</tr>
<tr>
<td>3.</td>
<td>1130</td>
<td>1163</td>
</tr>
<tr>
<td>4.</td>
<td>1083</td>
<td>1094</td>
</tr>
</tbody>
</table>
agreement pictures. The main effect of Homogeneity was nonsignificant \((F(1,12)=0.160, \text{n.s.)}\). A significant interaction appeared between Homogeneity and Agreement \((F(1,12)=5.832, p=0.033, \eta_p^2=0.327)\). After investigating the effect of Agreement in homogeneous and mixed blocks separately, we found that Agreement did not have a significant effect on homogeneous blocks \((F(1, 12)=1.804, \text{n.s.)}\) but low agreement pictures were named significantly slower in mixed blocks: \(F(1, 12)=6.305, p=0.027, \eta_p^2=0.344\). In the TD group the main effect of Agreement was significant \((F(1, 12)=9.219, p=0.010, \eta_p^2=0.434)\): low agreement pictures were named significantly slower than high agreement pictures. The main effect of Homogeneity was not significant \((F(1, 12)=1.069, \text{n.s.)}\), neither was the Homogeneity x Agreement interaction \((F(1, 12)=2.937, \text{n.s.)}\) (Results for the SLI and TD groups are shown in Fig. 2).

To investigate the significant Homogeneity x Cycle interaction further, we examined the effect of Homogeneity in the first and fourth cycles separately (collapsing mean reaction times over Agreement and Group). We found that in the first cycle naming pictures in the mixed condition took significantly longer than in the homogeneous condition \((F(1, 25)=8.546, p=0.007, \eta_p^2=0.255)\). We observed the opposite pattern in the fourth cycle: reaction times were significantly higher for pictures from the homogeneous condition \((F(1, 25)=14.579, p=0.001, \eta_p^2=0.368)\). (Mean reaction times by Homogeneity and Cycle are shown in Fig. 3).

The Agreement x Cycle x Group interaction was further investigated by examining the effect of Agreement and Cycle in the two groups separately. In the SLI group the main effect of Agreement was significant \((F(1, 12)=5.085, p=0.044, \eta_p^2=0.298)\): low agreement pictures were named significantly slower than high agreement pictures. The main effect of Cycle was not significant \((F(1, 12)=0.346, \text{n.s.)}\) while the Agreement x Cycle interaction was significant \((F(1, 12)=20.625, p=0.001, \eta_p^2=0.632)\). After investigating the effect of Agreement separately in the first and fourth cycles, we found a significant effect only in the case of the first cycle where low agreement pictures were named significantly slower than high agreement pictures (first cycle: \(F(1, 12)=11.066, p=0.006 \eta_p^2=0.480\), fourth cycle: \(F(1, 12)=0.153, \text{n.s.)}\). In the TD group the main effect of Agreement was significant \((F(1, 12)=9.226, p=0.010, \eta_p^2=0.435)\): low agreement pictures were named significantly slower than high agreement pictures. But the main effect of Cycle \((F(1, 23)=0.866, \text{n.s.)}\) and the Agreement x Cycle interaction \((F(1,12)=0.57, \text{n.s.)}\) was not significant (Mean reaction times in the SLI and TD groups by Cycle and Agreement are shown in Fig. 4).

We also tested whether conflict resolution abilities (the size of the conflict effect) were associated with individual differences in PIQ and language measures (performance on the screening tests). No significant correlations were observed either with PIQ, or with language measures (all ps > 0.1).

### 4. Discussion

Our aim in the current study was to investigate the effect of lexical conflict on word production in children with SLI compared to age-matched TD children. For this aim we manipulated the level of conflict in three ways in a picture naming task. Pictures were presented either in a 1) semantically homogeneous or in a semantically mixed context, they appeared in both contexts 2) four times across the task resulting in increasingly higher level of conflict in the homogeneous blocks and 3) pictures had either low name agreement with multiple plausible names or high name agreement with one dominant name. Results show a very similar pattern for the two groups. We found that pictures with lower name agreement, i.e. with multiple equally plausible names took longer to name for both children with SLI and for TD children. Furthermore, when pictures appeared for the first time in the block both groups named them faster in a semantically homogeneous context than in a mixed context but this pattern was reversed after three repetitions: in the fourth cycle, pictures in the semantically mixed context were named faster. Name agreement affected the two groups differently in the fourth cycle: TD children named pictures with multiple available names slower than pictures with one dominant name but no such difference appeared in children with SLI. The same difference appeared between the two groups when only the homogeneous blocks were
considered: pictures with multiple possible names were named slower than pictures with one dominant name by TD children but reaction times of children with SLI were similar for these two types of pictures in homogeneous semantic contexts.

Arguably, different effects of name agreement in the two groups can be partly accounted for by differences in the organization of the mental lexicons of children with SLI and TD children that might lead to different priming and conflict effects. We summarized some of these potential differences at the end of the Introduction both in line with the competition-by-selection and the incremental learning accounts. One of these potential differences based on the competition-by-selection theory was that competition can be smaller in children with SLI in the fourth cycle of the low agreement condition because competing alternatives of the target name are reactivated to a smaller degree than in TD children after once the word was selected successfully. The lack of reactivation can be the result of weaker relationships between semantic nodes and word representations and between semantically related word representations.

The lack of an agreement effect in the homogeneous semantic context in children with SLI is difficult to interpret. The finding that a conflict effect is not present in the case of low agreement pictures appearing in the fourth cycle in children with SLI as well as other differences between the mental lexicon of TD and SLI children are likely contributions. Lexical conflict has two different sources in the low agreement homogeneous condition: multiple available names on the one hand and homogeneous semantic context on the other hand. As discussed above, children with SLI were relatively quick to name low agreement pictures in the fourth cycle, probably resulting from the lack of strong reactivation of alternative names. Lower reaction times overall in the low agreement condition in the SLI group are thus mainly accounted for by this group difference in the fourth cycle. The effect of the homogeneous context might also be reduced in SLI. As it was discussed among the potential differences between SLI and TD lexicons predicted by the selection-by-competition account the Introduction, children with SLI might have fewer and weaker connections between semantically related word representations. While conflict in a homogeneous semantic context originates from strong relationships between representations of one semantic category that raise the activation level of all category members when one member is retrieved through spreading activation, weaker associations yield smaller competition and a reduced conflict effect.

The potential differences predicted by the selection-by-competition account between the lexicons of children with SLI and TD children are partly supported by our data. Generally weaker relationships between semantic nodes and word representations in the SLI group would predict generally higher reaction times which we did not find. The hypothesis about
less dominant alternative names of the object which would result in generally smaller name agreement was not supported either. In contrast, weaker relationships between semantically related word representations potentially leading to weaker reactivation of alternative names after the target was once selected (points two and three in the Introduction) can account for the RT differences between the two groups. The predictions of the incremental learning account would only apply for semantic blocking effects but our results show differences more related to name agreement manipulations for which the theory does not have clear predictions.

Overall, our results show that children, both with and without SLI exhibit some, but not all lexical conflict effects found in previous adult studies. The number of available names for a picture affected reaction times in children in a similar way as it affected adults in the study of Kan and Thompson-Schill (2004): just like adults, children were slower in naming pictures with multiple available names than pictures with one dominant name. Overall reaction times were higher for children (with a mean of 830 ms in adults and 1170 ms in children) but the magnitude of the effect of name agreement was comparable in the two studies (100 ms in adults and 80 ms in children).

The manipulation of semantic context and number of presentations resulted in a different pattern in our study than in a previous adult study with the same design (Schnur et al., 2006). Adults named pictures with a similar speed when they appeared for the first time in the block independently of the semantic context and reaction times grew faster with repetitions. This priming effect on naming latencies was smaller when pictures appeared in a homogeneous semantic context, yielding generally slower reaction times in the semantically homogeneous than in the semantically mixed context. In contrast, in our study semantic context of the pictures had different effects at the beginning and at the end of the blocks and the number of presentations had different effects on semantically homogeneous and mixed blocks. Homogeneous context facilitated the naming in the beginning of the blocks relative to mixed context and while reaction times increased with cycles in homogeneous blocks, they decreased in mixed blocks. Because of reverse effects of the semantic context at the beginning and at the end of the blocks, semantic context did not have an overall effect on naming latencies. Based on the framework of Schnur et al. (2006) discussed in the Introduction of our paper, we can state that activation of the semantically similar names facilitated naming speed in children relative to the semantically mixed context in the beginning of the blocks, when names appeared only once and thus competition was low. At later stages of the block, after producing the semantically similar names three times, competition became higher, demanding cognitive control mechanisms for successful retrieval, which led to higher reaction times than in the first cycle. The different pattern of results in children and adults can be attributed to the facilitatory effect of homogeneous semantic context in the first cycle in children but not in adults. This can probably be accounted by the faster reaction times of the adult population which allow less space for further facilitation by the priming effect of semantically similar names.

Results can be interpreted in the frame of the incremental learning account (Oppenheimer et al., 2010) supplemented by cognitive control mechanisms (Crowther & Martin, 2014) as well. We suggested at the end of our introduction that weight adjustments might be less efficient in children with SLI than in TD children. This consideration can be applied for the difference between children and adults as well (although the theory does not have clear predictions for the developmental aspects of incremental learning during word production). Due to slower weight adjustments in children than in adults, weight decrease between semantic nodes and competing word representations might take longer therefore the spreading activation between semantically related nodes can have a facilitatory effect. This effect is not expected to appear in adults because it is suppressed by the faster weight decreases resulting in comparable reaction times to the first cycle of the mixed condition.

Although we did not find an overall effect of semantic context in children, the effect of semantic context appeared both in children and adults at the end of blocks. Reaction times of adults were generally faster in Schnur et al. (2006) study (mean reaction times of adults is ~800 ms and the mean reaction time of children is ~1170 ms). Our study showed a more pronounced effect of semantic context in children, shown by bigger difference between reaction times for homogeneous vs. mixed semantic context in the case of the fourth appearance (~60 ms for adults in the Schnur et al. (2006) study and 122 ms for children in our study).

In sum, we found generally higher naming latencies in children than did previous studies in adults with similar manipulations. This is not surprising based on developmental research about word retrieval showing that retrieval speed reaches its plateau after age 10 (Wiegel-Crump & Dennis, 1986). Lexical conflict had similar effects in children and in adults, although the effect of semantic context was modulated in children by the priming effect of homogeneous semantic context when a picture appeared for the first time in the block. An age-related difference also appeared in the size of the effects of name agreement and semantic context manipulations. In adults, several factors were associated with a lexical conflict effect reflected by an increase in reaction times in picture naming: multiple available names (versus just one available name, Kan & Thompson-Schill, 2004) and homogeneous semantic context (i.e. names from the same category, versus mixed semantic context, i.e. names from different categories, Schnur et al., 2006). Children in our study showed a higher conflict effect than adults for the semantic context manipulation and a similar or even smaller effect for the name agreement manipulation. The higher conflict effect for the context manipulation was expected based on earlier results showing protracted development of cognitive control abilities until adolescence (Davidson et al., 2006) and it can be accounted for by less effective cognitive control abilities of children. The smaller effect of name agreement manipulation in children was an unexpected finding. A potential explanation for the small name agreement effect lies in the differences between the mental lexicons of children and adults. We suggested at the end of our Introduction that children with SLI might have weaker connections between semantic nodes and subdominant words leading to weaker competition in the case of low agreement pictures. Our data did not support this prediction – as the effect of name agreement manipulation was comparable in
children with SLI and in TD children – but this difference might be present between the lexicons of children and adults. Multiple available names might generate a smaller degree of conflict in children because the alternatives might have weaker connections with the semantic nodes: if a child uses couch for naming the couch, competing alternatives, like squash or settee might get less activation in children than in adults. This smaller level of conflict can be resolved even with less developed cognitive control abilities.

Although we cannot draw strong conclusions based on the comparison of conflict effects in different studies, these results suggest that conflict resolution processes are similar in children between 7;1 and 10;7 years and adults, although their efficiency might be different. Adult imaging studies with the same manipulations in picture naming tasks showed that brain areas associated with cognitive control are recruited for conflict resolution during word retrieval (Kan & Thompson-Schill, 2004; Schnur et al., 2006). Similarities between adult and child results in lexical conflict resolution suggest that general cognitive control processes are recruited in lexical conflict resolution of children as well. Systematic comparison of effects of name agreement and homogeneity on children and adults would be necessary to explore this question further, together with testing associations between general cognitive control abilities and lexical conflict resolution in children directly.

Our main aim in the current study was to explore the possibility that conflict resolution is especially difficult for children with SLI and their word retrieval problems can be partly accounted for by the impairments of conflict resolution processes. We expected that conflict manipulations will have a bigger effect in children with SLI than in TD children. Contrary to our expectations, conflict effects were similar in the SLI and TD groups, suggesting that lexical conflict resolution, at least when it involves semantic conflict, is not impaired in children with SLI, and thus lexical problems in SLI probably have other sources. This finding is in accordance with studies finding intact cognitive control abilities in children with SLI (e.g. Henry et al., 2012; Lukács, Ladányi, Fazekas, & Kemény, 2015; Noterdaeme, Amorosa, Mildenberger, Sitter, & Minow, 2001). Nevertheless, it should also be taken into consideration that the mental lexicons of children with SLI and TD children might be different leading to different levels of semantic conflict. If in SLI connections are weaker between semantic nodes and word representations as well as between semantically related word representations then the degree of conflict will be also smaller. Successful resolution of a smaller level of conflict might be achieved even if conflict resolution is impaired yielding similar performance patterns in SLI and TD.

Further studies are needed to investigate lexical processes and conflict resolution in SLI. The comparison of children with SLI with a vocabulary matched control group would be a fruitful line of future research: it would control for group differences in the present study potentially originating from differences between lexicon sizes of children with SLI and TD children. Previous work (Mainela-Arnold et al., 2008, 2010) found problems related to the inhibition of phonological representations in SLI, therefore the effect of phonological conflict on word retrieval should also be further investigated systematically. Another promising line of future research would be to study the effect of lexical conflict with a design or with a set of experiments that allows the investigation of facilitatory and inhibitory effects in a more targeted way together with directly examining the relationship between lexical conflict resolution and cognitive control.

Acknowledgments

This research was supported by research grant OTKA K 83619 ‘Nonlinguistic abilities in Specific Language Impairment’ from the Hungarian National Science Foundation to Ágnes Lukács.

Appendix A.

Required answers for pictures of a mixed block: thumb, lemon, wardrobe, glass, doll, fan, wardrobe, thumb, doll, fan, glass, lemon, thumb, doll, fan, wardrobe, lemon, glass, lemon, fan, doll, wardrobe, thumb, glass.

Required answers pictures of a homogeneous block: lemon, strawberry, peach, cherry, apple, pear, apple, peach, pear, cherry, lemon, strawberry, peach, pear, cherry, lemon, strawberry, apple, cherry, peach, strawberry, apple, lemon, pear.

The categories of the homogeneous blocks: fruits, parts of the human body, electrical devices, furniture, toys, kitchen utensils.

References


Kas, B., & Lukács, Á. *Magyar Mondatátanömdési Test [Hungarian sentence repetition test].* Manuscript in preparation.


Study 3.

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

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), auditory 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.
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 abil-
ities in SLI motivated by the assumption that the impairment of cognitive control
might contribute to language problems. As we have discussed previously, con-
cepts referring to abilities controlling and coordinating our thoughts and actions
vary within the literature, and the lack of clear conceptualization of these phe-
omena 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 informa-
tion in children with SLI (e.g., Finneran, Francis, & Leonard 2009; Henry, Messer,
& Nash, 2012; Im-Bolter, Johnson, & Pascual-Leone, 2006; Marton, 2008; Mart-
ton, 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 con-
trol is involved in anaphor resolution, and thus an impairment of cognitive control
in SLI contributes to sentence processing difficulties in the case of sentences con-
taining 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 au-
thors 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. There-
fore, 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 pro-
nouns (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 cogni-
tive 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. 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...
Typically developing children were matched individually to children in the SLI group on chronological age and sex, and they were matched groupwise on non- verbal 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

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>SLI</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Age</td>
<td>8.93</td>
<td>1.18</td>
<td>7.08</td>
<td>11.33</td>
<td>8.95</td>
<td>1.18</td>
<td>7.08</td>
<td>11.25</td>
</tr>
<tr>
<td>Stand. scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven IQ</td>
<td>106.27</td>
<td>9.94</td>
<td>85</td>
<td>125</td>
<td>101.9</td>
<td>10.09</td>
<td>85</td>
<td>130</td>
</tr>
<tr>
<td>Raw scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonword repetition</td>
<td>6.33</td>
<td>0.96</td>
<td>4</td>
<td>8</td>
<td>3.43</td>
<td>1.17</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>PPVT</td>
<td>124.033</td>
<td>12.21</td>
<td>102</td>
<td>146</td>
<td>96.73</td>
<td>18.91</td>
<td>66</td>
<td>132</td>
</tr>
<tr>
<td>TROG blocks</td>
<td>17.77</td>
<td>1.7</td>
<td>13</td>
<td>20</td>
<td>13.33</td>
<td>2.26</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Sentence repetition</td>
<td>36.83</td>
<td>4.64</td>
<td>16</td>
<td>40</td>
<td>20.57</td>
<td>8.24</td>
<td>0</td>
<td>36</td>
</tr>
</tbody>
</table>

Note: TD, Typically developing; SLI, specific language impairments; PPVT, Peabody Picture Vocabulary Test; TROG, Test of Reception of Grammar.
“Christopher Robin says that 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 × 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

<table>
<thead>
<tr>
<th>Match</th>
<th>Antecedent Mismatch</th>
<th>Agent Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pronoun, no quantifier</strong></td>
<td>Rabbit sprinkles</td>
<td>Rabbit sprinkles himself.</td>
</tr>
<tr>
<td>Róbert Gida azt mondja, hogy Nyuszi lelocsolja Őt.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Christopher Robin says that Rabbit sprinkles him.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pronoun, quantifier</strong></td>
<td>Every rabbit sprinkles</td>
<td>Every rabbit sprinkles himself.</td>
</tr>
<tr>
<td>Róbert Gida azt mondja, hogy minden nyuszi lelocsolja Őt.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Christopher Robin says that every rabbit sprinkles him.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reflexive, no quantifier</strong></td>
<td>Rabbit sprinkles himself.</td>
<td>Rabbit sprinkles himself.</td>
</tr>
<tr>
<td>Róbert Gida azt mondja, hogy Nyuszi lelocsolja magát.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Christopher Robin says that Rabbit sprinkles himself.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reflexive, quantifier</strong></td>
<td>Every rabbit sprinkles himself.</td>
<td>Every rabbit sprinkles himself.</td>
</tr>
<tr>
<td>Róbert Gida azt mondja, hogy minden nyuszi lelocsolja magát.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Christopher Robin says that every rabbit sprinkles himself.”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 $2 \times 2$ 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
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
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 \( \times \) Quantifier \( \times \) Group interaction, we conducted a 2 \( \times \) 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, $t(29) = 0.000, \text{ns}$, 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: $t(29) = 0.551, \text{ns}$; agent: $t(29) = 1.934, \text{ns}$.

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

**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, $F(1, 59) = 2.554, \text{ns}$; 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 ($r = .18, \text{ns}$), even after controlling for short-term memory span ($r = .172, \text{ns}$).

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 ($r = .261, \text{ns}$) 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: $r = .280, \text{ns}$; 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: $r = .36, \text{ns}$; partial: $r = .36, \text{ns}$).

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: $r = -.022, \text{ns}$; partial: $-.051, \text{ns}$).

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 backward/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


Study 4.


Ladányi E., & Lukács Á. (under revision in JSLHR manuscript). Cognitive control impairment and its contribution to word production difficulties in specific language impairment.
Cognitive control impairment and its contribution to word retrieval difficulties in specific language impairment

Enikő Ladányi ¹, ²
Ágnes Lukács ²

¹ Laboratoire Psychologie de la Perception, Université Paris Descartes
² Department of Cognitive Science, Budapest University of Technology and Economics

Enikő Ladányi
Laboratoire Psychologie de la Perception, Université Paris Descartes;
Department of Cognitive Science, Budapest University of Technology and Economics
eniko.ladanyi@parisdescartes.fr

Ágnes Lukács
Department of Cognitive Science, Budapest University of Technology and Economics
alukaes@cogsci.bme.hu
Abstract

**Purpose:** The study aims to investigate factors behind word retrieval difficulties in children with specific language impairment (SLI), focusing on the role of cognitive control.

**Method:** 31 children with SLI (mean age: 8;11, SD: 1;1) and 31 age- and IQ-matched typically developing (TD) children completed two word production tasks potentially requiring cognitive control – a fluency task (letter, category, action conditions) and a rapid automatized naming (RAN) task (letters, numbers, objects, and a size-color-shape condition) – together with three cognitive control tasks (backward digit span, n-back, Stroop) to test whether children with SLI show weaknesses in word production and in cognitive control and to find out whether impairments in the two domains are associated.

**Results:** We found weaker performance in the SLI than in the TD group on fluency and in the size-color-shape RAN as well as on the backward digit span and n-back tasks. Performance on the letter fluency task was associated with backward digit span while size-color-shape RAN performance was associated with backward digit span and n-back scores. Performance on both tasks was best explained by non-word repetition and backward digit span.

**Conclusions:** These results suggest that both word production and cognitive control are impaired in SLI and some word production problems can be accounted for by weaker cognitive control, although weaker short-term memory also has a crucial contribution to word retrieval difficulties during these tasks. If further research confirms the association with cognitive control, training of this ability should be included in the therapy of at least some children with SLI.

Introduction

Children with specific language impairment (SLI) show various language problems which cannot be accounted for by impairments in other cognitive domains or perceptual deficits, neurological disorders, emotional or social problems, environmental deprivation or intellectual disability. Usually morpho-syntactic and syntactic problems are emphasized (e.g.: Bishop, 1997; Leonard, 1998/2014) but lexical impairments are reported as well. The potential first sign of SLI is that the child starts to produce words later than typically developing peers (although not all the children with a delay in word production will have SLI). The vocabulary size of children with SLI lags behind age-based expectations at older ages too (Bishop, 1997; Watkins, Kelly, Harbers, & Hollis, 1995; Trauner, Wulfeck, Tallal, & Hesselink, 1995). The retrieval of already known words is also often impaired: children

In contrast with early studies investigating the source of problems in SLI emphasizing a language-specific impairment, more recent theories assume that various non-linguistic cognitive functions are also impaired in SLI and language problems can be (partly) accounted for by these non-linguistic deficits. Several non-linguistic abilities were found to be impaired. Some theories proposed a deficit in processing stimuli with certain features, like the processing of rapidly changing auditory stimuli (Tallal & Piercy, 1973; Tallal, 1976; Tallal, Stark, & Mellits, 1985) or less salient morphemes (Leonard, 1989). A general processing capacity limitation (Leonard, 1998/2014) or general slowing of processing (Marchman & Bates, 1994; Conti-Ramsden & Jones, 1997; Windfuhr, Faragher, & Conti-Ramsden, 2002; Conti-Ramsden, 2003) was also assumed, together with the impairment of working memory (Gathercole & Baddeley, 1990) or procedural learning (Ullman & Pierpont, 2005). In the last couple of years cognitive control/executive functions also became the focus of research in SLI, with controversial results so far (e.g., Im-Bolter, Johnson, & Pascual-Leone, 2006; Marton, Schwartz, Farkas, & Katsnelson, 2006; Marton, Kelmenson, & Pinkhasova, 2007; Marton, Campanelli, Eichorn, Scheuer, & Yoon, 2014, Henry, Messer, & Nash, 2012; Lukács, Ladányi, Fazekas, & Kemény, 2016). Motivated by these controversial findings, the current study will investigate cognitive control in children with SLI further, and we will explore whether cognitive control impairments contribute to word retrieval problems in children with SLI. In what follows, first, the concept of cognitive control will be introduced and research about the role of cognitive control in word retrieval will be summarized. After that, studies
investigating cognitive control and word retrieval under conflict in children with SLI will be reviewed. At the end of the introduction the aims of the current study will be presented.

**The role of cognitive control in word retrieval**

Cognitive control is responsible for the resolution of conflict or interference between contradicting representations implemented mostly by the left inferior frontal gyrus (LIFG; Novick, Trueswell, & Thompson-Schill, 2005, 2010, Miller & Cohen, 2001). Such functions have also been referred to as executive functions (EF) which are defined as a set of top-down mental processes recruited when we have to concentrate or pay attention and when responding automatically or relying on instinct or intuition would be ill-advised, insufficient, or impossible (Diamond, 2013). The influential EF model of Miyake, Friedman, Emerson, Witzki and Howarter (2000) assumes that there are three separable components within executive functions: shifting, updating and inhibition. Some researchers use these terms (i.e. cognitive control and executive functions) interchangeably, while others differentiate between them. We follow Novick and his colleagues’ view (e.g., Hussey, Harbison, Teubner-Rhodes, Mishler, Velnoskey & Novick, 2017) in this respect who consider cognitive control as a construct similar to inhibition in Miyake’s model defined as a process responsible for the inhibition of dominant, automatic, or prepotent responses when necessary. Hussey et al. (2017) argue that they do not use inhibition to refer to this process because they aim to be neutral about whether cognitive control involves inhibition of task-irrelevant representations or promotion of task-relevant ones or a combination of both processes.

Cognitive control was found to be important for word retrieval – among other language processes. In the course of retrieving a word, several semantically or phonologically similar word representations are activated beyond the target word. Conflict appears when the activation of these competing representations is as high or higher than the activation of the
target word and cognitive control is assumed to play a role in the resolution of such conflict (Kan & Thompson-Schill, 2004; Schnur, Schwartz, Brecher, & Hodson, 2006; Schnur, Schwartz, Kimberg, Hirshorn, Coslett, & Thompson-Schill, 2009). In our everyday life conflict during word retrieval is manifested in the form of word finding difficulties, the tip of the tongue phenomenon (when the target word’s activation is similar to those of the competing words) or slips of the tongue (when one of the competing words’ activation is higher than that of the target word). Although these word finding difficulties typically appear in sentences, experimental paradigms often investigate the question with word production tasks instead of studying sentence production since it is easier to manipulate the variables and collect RT-s with using only words. Different paradigms have different methods for creating situations similar to those which lead to word finding difficulties during our everyday language use.

The most frequently used method for conflict manipulation is to present pictures in semantically homogeneous (e.g. different fruits have to be named after each other) vs. semantically mixed (pictures showing objects from different semantic categories have to be named) blocks in a picture naming task (e.g., Schnur et al., 2006, Schnur et al., 2009). Conflict is assumed to be higher in the case of the semantically homogeneous blocks because competing word representations get a high activation due to their semantic relationship to the target word and their recent retrieval in the same block. Another method is to manipulate the number of possible names of pictures in a picture naming task (e.g., Novick, Kan, Trueswell, & Thompson-Schill, 2009; Kan & Thompson-Schill, 2004). In this paradigm conflict is assumed to be higher when several similarly frequently used names are available for a picture (e.g., a picture of a sofa can be also named as a couch or a loveseat) than when the picture has one dominant name. Neuropsychological studies with patients with brain injury in their LIFG (Schnur et al., 2006, Schnur et al., 2009, Novick et al, 2009) found a weak performance
(errors, long naming times) on high conflict conditions with a typical performance on low conflict conditions in such tasks. Furthermore, brain-imaging studies with adults (Schnur et al., 2006, 2009; Kan & Thompson-Schill, 2004) using such paradigms found that the same brain area – the LIFG – is recruited for conflict resolution in word production as in tasks taxing cognitive control. Both group of results strongly support the role of cognitive control in word retrieval under conflict.

Although the abovementioned picture naming paradigms are the most frequently used tasks in the investigation of conflict resolution during word retrieval, cognitive control is also required in other word production paradigms. Below we present two tasks – the fluency task and the rapid automatized naming (RAN) task – with a review of supporting results for the involvement of cognitive control.

During fluency tasks participants are asked to produce as many words as they can without any repetitions according to different criteria (e.g., words starting with a certain letter in letter fluency tasks or words belonging to a certain category in category fluency tasks). Fluency tasks are widely used in neuropsychological batteries for assessing verbal functions and executive functions, although the exact underlying processes required by the task are not clear (Shao, Janse, Visser, & Meyer, 2014). Cognitive control is potentially necessary for the successful execution of the task for overcoming conflict generated by irrelevant words, previously produced words and to be produced words.

There are less studies investigating the role of cognitive control in the fluency task than in picture naming tasks with conflict manipulations but some supporting evidence can be found in the neuropsychology literature. Several papers showed impairments in verbal fluency in patients with frontal lobe impairments (Baldo & Shimamura, 1998, Stuss, Alexander, Hamer, Palumbo, Dempster, Binns et al., 1998; Troyer, Moscovitch, Winocur,
Alexander, & Stuss, 1998; Schwartz, Baldo, 2001; Rogers, Sahakian, Hodges, Polkey, Kennard, & Robbins, 1998). Interestingly, Rogers et al. (1998) found that the impairment was more pronounced in the case of the letter fluency task than in the category fluency task in patients relative to control participants.

The above results are in accordance with the hypothesis that the impairment of cognitive control leads to weaker performance on the fluency task since the LIFG which is responsible for cognitive control is in the left frontal lobe. However, due to methodological limitations, the exact location of the brain injury is unknown in these early studies, therefore brain areas responsible for other abilities required for the fluency task might be (also) impaired and these impairments could also lead to weaker performance on the fluency task. A recent study (Robinson, Shallice, Bozzali, & Cipolotti, 2012) using more precise anatomical categories compared patients with a LIFG vs. non-LIFG frontal injury (left or right) on a category and a letter fluency task and found a weaker performance in letter fluency in LIFG patients while the number of correctly produced words did not differ between LIFG and non-LIFG groups in the case of the category fluency task. The authors suggest that the performance difference appeared because competition is higher in letter fluency than in category fluency task. In letter fluency word representations activated due to semantic similarity are often irrelevant therefore they have to be inhibited while they are probably correct answers in the case of the category fluency task (e.g., when a participant produces strawberry during a fluency task, apple, peach and raspberry will be also activated; these fruit names are irrelevant if the task is a letter fluency task in which words starting with s have to be produced, while they are relevant if the task is a category fluency task in which fruits have to be produced). This study provides a strong evidence for the involvement of cognitive control in the letter fluency task.
Rapid automatized naming (RAN) tasks may also involve cognitive control. During the task several stimuli are presented on the screen and participants are asked to name them from the left to the right, row by row, from the first item of the first row. Stimuli are usually letters, digits, objects or colors. The RAN task is stated to measure different abilities by different studies (e.g., phonological skills, lexical access, naming ability, expressive language, working memory, processing speed; Decker, Roberts, & Englund, 2013; Aguilar-Mediavilla, Buil-Legaz, Perez-Castello, Rigo-Carratala, & Adrover-Roig, 2014) and it might require cognitive control as well. As Bexkens, Wildenberg and Tijms (2015) argue, during a RAN task previously named stimuli compete for selection with the current target stimulus and these competing inappropriate word representations have to be inhibited for the successful naming of the target item. Cognitive control might be required for the resolution of conflict between competing representations and the target word.

The task is most often used in dyslexia research since RAN performance shows a strong relationship with reading abilities (Compton, 2003; Powell, Staintorp, Stuart, Garwood, & Quinlan, 2007; Lervåg & Hulme, 2009; Moll, Fussenegger, Willburger, & Landerl, 2009; Pan, McBride-Chang, Shu, Liu, Zhang, & Li, 2011; Nag & Snowling, 2012; Georgiou, Parilla, Cui, Papadopoulus, 2013; Wijayathilake & Parrila, 2014), therefore results about the role of cognitive control in RAN tasks also come from dyslexia research. Bexkens et al. (2015) found that cognitive control measured by the Simon task (Craft & Simon, 1970) was related to the RAN performance with a stronger association to the colors and objects RAN than to the numbers and letters RAN in children with dyslexia. The authors argue that the association with the colors and the objects task was stronger than with the numbers and letters task because conflict is higher in the case of the first two. Unlike numbers and letters, colors and objects refer to categories with variable and overlapping boundaries and often more than one plausible name is available for them therefore it is more likely that several
word representations will be activated for one target item which generates conflict. The relationship between the Stroop task – a traditional paradigm measuring cognitive control – and the RAN task was also investigated and an association was found with the color RAN but not with the letters and digits RAN in typical readers (Stringer, Toplak, & Stanovich, 2004) which is in accordance with Bexkens et al. (2015)’s results. Amtmann, Abbott and Berninger (2007) investigated atypical readers and found an association between the letters RAN and the Stroop task. Taken together, relationships between performance on cognitive control task and on the RAN task suggest that cognitive control might play a role in word retrieval during the RAN task.

The above studies using different word retrieval tasks in various populations suggest that cognitive control is involved in word retrieval under conflict. There is no agreement, however, in the literature whether cognitive control supporting word retrieval is special to linguistic stimuli or is more general. Some authors (e.g., Schnur et al., 2006) argue for a system specific to language while others assume the existence of a domain-general mechanism (e.g., Hsu, Jaeggi, & Novick, 2017; although in this view too, general conflict resolution mechanisms are supposed to work together with domain specific systems). One way to test whether domain-general cognitive control is involved in word retrieval would be the investigation of relationships between behavioral measures/brain activations on/during non-linguistic cognitive control tasks and word retrieval tasks. We have not found, however, any studies testing cognitive control with nonlinguistic stimuli together with word retrieval performance. Since we did not use a nonverbal cognitive control task, we do not aim to determine the domain generality of cognitive control recruited during word retrieval in the current study either.

Since – as discussed in the first section – word retrieval is impaired in SLI and some
results show that cognitive control might be impaired in SLI, cognitive control impairments may contribute to word retrieval problems in this population. In the next section, research focusing on cognitive control in children with SLI and on the relationship of cognitive control impairments with word production problems will be reviewed.

**The role of cognitive control in word retrieval problems in SLI**

A growing body of evidence shows that children with SLI have difficulties in tasks that require conflict resolution both in the verbal and in the non-verbal domain (backward digit span: Lum, Conti-Ramsden, Page, & Ullman, 2012; Vugs, Hendriks, Cuperus, & Verhoeven, 2014; listening span: Ellis Weismer, Evans, & Hesketh, 1999; Archibald & Gathercole 2006; Mainela-Arnold & Evans, 2005; Vugs et al., 2014; Marton et al., 2006; Marton et al., 2007; Montgomery & Evans, 2009; odd-one-out: Vugs et al., 2014; n-back: Evans & Pollack, 2011; Im-Bolter et al., 2006; category judgment under conflict: Marton et al., 2014). According to the results presented in the previous section, cognitive control plays a role in word retrieval under conflict, therefore word retrieval problems of children with SLI might be partly accounted for by cognitive control impairments. Research targeting this question is, however, very limited.

As far as we know, our previous picture naming study is the only one testing effects of lexical conflict manipulation in children with SLI (Ladányi & Lukács, 2016) which found that they were generally slower in naming pictures but as efficient in resolving conflict during word retrieval as TD children. In the current study we aimed to investigate word retrieval under conflict further in children with SLI and TD children using a fluency task and a RAN task. The few results available on research using these tasks in children with SLI are summarized below.

Several studies found an impairment on various fluency paradigms in children with
SLI and some of them assume that the impairment of cognitive control or executive functions have a role in these problems. Henry and colleagues (2012) found that 8 to 14 years old children with SLI were able to produce less words in a category and a letter fluency task than their TD peers. In a following study (Henry, Messer & Nash, 2015) the authors reported a more detailed analysis on the fluency performance of the participants of Henry et al. (2012): children with SLI produced less words, more errors and less switches than TD children both in letter and category fluency. The authors also investigated relationships with executive functions and found an association between the number of errors in letter fluency and inhibition. Rodriguez, Santana and Exposito (2017) also found that 5 to 11 years old children with SLI produced fewer words both in a letter and a category fluency task than TD children. Similarly, Weckerly, Wulfeck and Reilly (2001) showed that children with SLI were able to produce less words both in a category and letter fluency task than their TD peers. Both Rodriguez et al. (2017) and Weckerly et al. (2001) assumed that lower fluency in the SLI group relative to the TD group is a marker of less efficient executive functions in children with SLI.

Studies reporting similar performance on fluency tasks in children with SLI and in TD children are also present in the literature. Kail & Leonard (1986) conducted a category fluency task in 6-14-year-old children with SLI and did not find a group difference in the number of words retrieved, neither relative to an age- nor to a language-matched control group. Weyandt and Willis (1994) made the same observation investigating category fluency in 6-12 year old children. Fluency was also investigated in one of our previous studies (Lukács et al., 2016) in a group of 6 to 9 years old children with SLI and age- and IQ-
matched TD children. We found a significantly weaker score when we combined category, letter and action fluency but further analysis revealed that the group difference can be accounted for by lower short-term memory (STM) in the SLI group.

As discussed above, rapid automatized naming tasks also involve cognitive control. The few RAN results from children with SLI come from studies investigating RAN in children with dyslexia also including children with SLI. In what follows, we only summarize results from these studies for children with SLI but without dyslexia.

De Groot, Van den Bos, Van der Meulen and Minnaert (2015) found slower naming times both in letters and digits RAN in 8-13 years old children with SLI than in age-matched TD controls, with a bigger effect in the case of the letters RAN. Similarly, in Claessen, Leitão, Kane & Williams (2013) 6-8 years old children with SLI showed significantly slower naming times than TD children in an objects RAN task. Katz and colleagues (1992) also found that children with SLI were significantly slower than age-matched TD children on an objects RAN in a longitudinal study at all of the three data points (4, 6 and 8 years).

In contrast, Bishop, McDonald, Bird, & Hayiou-Thomas (2009) did not find a significant difference between 9-10-years old children with SLI and TD children in the objects and digits RAN. Neither did Vandewalle, Boets, Ghesquière and Zink (2012) in either of the RAN tasks (letters, colors, digits, objects) in their longitudinal study in which children were assessed in kindergarten, grade 1 and grade 3.

While the letters, digits, objects and colors RAN versions are the most commonly used versions of the task, there is also a variant that requires the naming of shapes together with their size and color (e.g., big yellow square). We assume that conflict is higher in this

---

3 Participants in Lukács et al. (2016) partly overlap with participants of the current study (with a higher age range in this study than in the previous one).
task than in the previously mentioned versions since three lexical units have to be retrieved for each item and each of them will have competitors: colors, sizes and shapes of previously named elements will compete with the color, size and shape of the target item. Furthermore, lexical units referring to the size, color and shape compete with each other too (e.g., if the shape or the color is activated first instead of the size which is required to be produced first, then these word representations also generate conflict). Beyond cognitive control, this task requires sequential organization as well, which is also vulnerable in SLI (Ullman & Pierpont, 2005; Lukács & Kemény, 2014) making the size-color-shape more difficult for children with SLI RAN than the other versions of the task. We are aware of only one study (Aguilar-Mediavilla et al., 2014) using the size-color-shape RAN paradigm in 6-8-years old bilingual children with SLI, finding significantly lower accuracy scores in the SLI than in the TD group.

The current study

Motivated by the contradictory findings of fluency and RAN performance in SLI, our first aim in this study was to find out whether children with SLI show weaker performance on the various conditions of the fluency (letter, category and action fluency) and the RAN (letters, numbers, objects and size-color-shape RAN) tasks.

Our second aim was to contribute to the research targeting cognitive control in children with SLI by assessing children’s cognitive control. Since there is no single generally accepted measure of cognitive control in the literature – different studies use different tasks to assess this ability – we decided to use three relatively different tasks (the Stroop task, the n-back task and the backward digit span task) which potentially require cognitive control to get a more general picture of our participant’s cognitive control ability. During the Stroop task, which is probably the most frequently used cognitive control task, conflict appears
between the color activated by the word presented on the screen and the color of the letters and it has to be resolved for the appropriate answer. The n-back task is often used as a working memory updating task but it also requires cognitive control because participants have to overwrite the representations of recently presented letters except one based on which they have to respond (e.g., the letter 2 steps before the actual one in the 2-back condition) to give a correct answer (e.g., Jaeggi, Buschkuehl, Perrig & Meier, 2010). The backward digit span is most often considered as a complex/executive-loaded working memory task but it might also require cognitive control because conflict is assumed to appear between the items which should be produced later and the target item (May, Hasher, & Kane, 1999).

Based on 1) studies finding weaker performance on fluency and RAN tasks in children with SLI than in TD children, 2) works showing weaker performance on tasks involving cognitive control in children with SLI than in TD children and 3) on studies of patients with brain injury showing the involvement of cognitive control in fluency tasks and dyslexia research supporting the involvement of cognitive control in the RAN task our third aim was to investigate the role of cognitive control in word retrieval in children with SLI and TD children. More specifically, we aimed to find out whether weaker cognitive control could account (at least partly) for impaired performance on the RAN and fluency tasks in children with SLI.

Both fluency and RAN tasks require several linguistic and non-linguistic processes beyond cognitive control. STM is necessary for both paradigms and also for cognitive control, therefore it is important to rule out the possibility that associations between the language and cognitive control tasks appear only due to their common STM requirement. Therefore we included non-word repetition scores in the investigation of associations between performances on these tasks. Furthermore, vocabulary size arguably has a role in fluency
performance and it might affect RAN performance as well. To explore the effect of cognitive control separately from vocabulary size as well as the relative contribution of the different cognitive control measures to performances on the word retrieval tasks, we conducted a linear regression analysis for those word retrieval tasks which showed an association with cognitive control tasks. Beyond the cognitive control tasks and vocabulary age and STM scores were also included in this analysis.

We hypothesized that children with SLI will show weaker performance both on language and cognitive control tasks than their TD peers. Furthermore, we assumed that cognitive control contributes to conflict resolution during word retrieval in the fluency and the RAN tasks. During the fluency task cognitive control is assumed to be involved in the resolution of conflict between irrelevant word representations/already produced/to be produced words and target word. In the RAN task conflict between the names of previously named items and the target word hypothesized to be resolved by cognitive control. Based on these assumptions, we hypothesize that scores on the fluency and RAN tasks will be associated with performance indices on tasks measuring the efficiency of cognitive control even if we take into account the effect of vocabulary size, STM and age.

**Methods**

*Participants*

31 white Caucasian Hungarian children with SLI and 31 age-, gender- and IQ-matched typically developing (TD) children (8 girls in each group) participated in the study (see age and IQ of the groups in Table 1). Children with SLI were recruited from two different schools; both are special institutions for children with speech problems. One school is in Kőszeg - a small town in Hungary - while the other school is in Budapest, the capital of Hungary. All of the children with SLI were receiving speech therapy in the school. Typically
developing children matched to children with SLI attending the school in the small town were recruited from a school in another small town – Szentes – in Hungary and TD pairs of children attending the school in Budapest were recruited from five different schools in Budapest.

As a first step of the recruitment process speech therapists selected a larger group of children in both schools who had no history of neurological impairment, psychiatric or social problems and their hearing was normal. Non-verbal intelligence (Raven Colored Progressive Matrices; Raven, Court, & Raven, 1987) and language skills were assessed in these children by the speech therapists. Linguistic abilities were assessed with four tests targeting both receptive and expressive skills. The receptive tests were the Hungarian adaptations of the Peabody Picture Vocabulary Test (PPVT: Dunn & Dunn, 1981; Csányi, 1974) and the Test for Reception of Grammar (TROG; Bishop, 1983; Lukács, Győri, & Rózsa, 2012). The expressive tests were the Hungarian Sentence Repetition Test (Kas & Lukács, *in prep*), and the Hungarian version of the non-word repetition test (Racsmány, Lukács, Németh & Pléh, 2005). In accordance with general practice in SLI research (see, e.g., Leonard 1998/2014, Tager-Flusberg & Cooper, 1999), children who showed a normal intelligence (performance above 85 scores) and performed at least 1.5 SD below age norms on at least two out of the four tests were selected for the SLI group see the results of a screening tests in Table 1.)

For the selection of the TD group, non-verbal intelligence was measured in typically developing children who matched in age (maximum 3 months difference in date of birth between the TD child and SLI pair) and gender to the members of the SLI group. Children whose IQ was similar (the difference is not greater than 10 IQ points) to the SLI pair’s IQ were selected for the TD group.
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.

------------------------
Table 1.
------------------------

*Design and procedure*

*Fluency*

In the fluency task, children were asked to produce in 1 minute as many (a) actions or things that people do (action fluency), (b) things they can buy at a supermarket (category fluency), and (c) words starting with k (letter fluency) as they can in one minute for each condition with avoiding repetitions. The task was preceded by a practice session. We assessed three measures: the number of correct answers, number of repetitions and the number of errors for all three subtasks but only the number of correct answers was used in the analysis because most of the children did not produce any repetitions or errors.

*Rapid automatized naming*

Children were tested with four versions of the RAN task. The design and procedure was the same in all versions, only the stimuli differed. During the task (a) letters, (b) digits, (c) objects and (d) shapes in different sizes and colors were presented on the screen. Stimuli were arranged in several rows (three rows in the letters, digits and objects task and four rows in the shapes task) and in each row the same five items appeared in different orders. The same set of stimuli was presented three times in a row. The task was to name the items starting with the first one in the first row as fast as possible and in the case of the shapes
children were asked to produce first the size, then the color and finally the shape of the stimulus (e.g., big yellow square). The tasks were preceded by a short practice session. Stimuli were presented and naming times were recorded with the E-Prime software (Schneider, Eschman, & Zuccolotto, 2012). Response accuracy was coded by the experimenters on paper. In the case of the letters, digits and objects RAN children could get one point for each correct answer, yielding a maximum score of 45 (15 for each presentation of the slide) for each type of task. In the case of the shapes task children could get 3 points for each item (1 for the size, 1 for the color and one for the shape of the item), therefore the maximum score was 180 (60 for each presentation of the slide). Correctness of the sequence of the size, color and shape was also coded: if a child produced the words in the required order, (s)he gained 1 point, if the order was incorrect, (s)he got 0 point meaning that the maximum score on the sequence was 60 (20 for each presentation of the slide). The sum of the naming times for the three slides was also measured and both accuracy scores and naming times were used in the analysis as the measure of the RAN performance. The task took ca. 20 minutes to administer.

Stroop task

In the Stroop task color names (red, blue, green, yellow) were presented written with different colors (red, blue, green, yellow) and the task was to press a key (of four different keys, each representing one color) based on the color of the fonts. In the congruent condition the meaning of the words matched the font color (blue written in blue), while in the incongruent condition, they did not match (blue written in red). In the control condition a non-linguistic string was presented in different colors (xxxx in blue). Trials belonging to the incongruent, congruent and control conditions were presented in a blocked fashion with 60 items in each block and the order of the blocks was randomized. At the beginning of the task children completed a short practice session including all types of trials. We used the E-Prime
software (Schneider et al., 2012) to present stimuli and to collect data. Both reaction times and number of correct answers were collected and the average reaction times and the total number of correct answers were calculated for each block for each child. As a measure of cognitive control, we used the difference between average reaction times given for incongruent and control trials. (We did not use accuracy scores during the analysis since children’s performance did not differ in the incongruent and control conditions due to the small number of errors). The Stroop task took ca. 10 minutes to administer.

N-back task

In the n-back task, letters were presented on the screen and children were asked to monitor the letters and indicate (by pressing “ENTER”) when the same letter appeared two trials earlier. 60 trials were presented out of which 10 trials were two-back trials. Before the test trials, a short practice session was presented. Stimuli were presented and answers were collected with the E-Prime 2.0 software (Schneider et al., 2012). We calculated the number of hits (when the participant correctly pressed “ENTER” on a “two-back trial,” i.e., when the current item was identical to the target item, with a maximum of 10 hit) and the number of false alarms (the participant pressed “ENTER” on a not “two-back trial,” i.e., the actual stimulus was not identical to the one presented two before) and we used the difference between hits and false alarms (n-back score) in the analyses as a measure of cognitive control. The task took about 5 minutes to administer.

Backward digit span

In the backward digit span task children were presented with a set of numbers and they had to repeat the numbers in a reversed order. At the beginning, two numbers were presented and the set size was increasing during the task. There were four sets of numbers in
each level and participants could go on to the next level if they could produce the reversed order at least two times out of the four sets. Children completed a few practice trials before the test trials. The measure of the task was the last level that the child successfully completed (span). The task took ca. 5 minutes to administer.

Tasks were presented in a randomized order as part of a larger battery of tasks.

Data Analysis

First, we investigated the patterns of results in the language tasks. In the fluency task performance on the letter, action and category fluency tasks was compared across the two groups. The number of correct answers was analyzed with a 2 x 3 ANOVA with Group (SLI vs. TD) as a between-subject factor and Fluency Type (letter vs. action vs. category) as a within-subject factor. In the four RAN tasks group differences were investigated with two one-way ANOVAs: one with accuracy and one with naming time as the dependent variable.

Second, performance differences between the SLI and TD groups were investigated in cognitive control tasks – in the backward digit span task, the n-back task and the Stroop task with one-way ANOVAs. If a group difference appeared, we repeated the analysis with including non-word repetition scores as covariates to exclude the possibility that children with SLI performed worse due to their weaker verbal STM.

Associations between the language and cognitive control measures were explored with partial correlations including age as a control variable. Since both the language tasks and the cognitive control tasks require keeping information active in STM, without controlling for differences in STM, correlations might appear due to the common STM component instead of reflecting associations with cognitive control. To avoid this confound, non-word repetition scores (available for all children due to the screening process) were also included as a control variable to control for the effect of STM.
As it was mentioned before vocabulary size can have an effect on the performances on the word retrieval tasks. For investigating the relative contribution of vocabulary size scores and cognitive control to word retrieval performances, a linear regression analysis was conducted on the whole group for those word retrieval measures which showed a correlation with at least one of the cognitive control measures even after controlling for age and STM. Age and STM were also included as independent variables beyond vocabulary size and the three cognitive control measures. Therefore five independent variables were included in each of the linear regressions: n-back score, backward digit span, Stroop effect, vocabulary score and age.

Some children did not complete all of the tasks and scores for the size-color-shape RAN task were not available for each child due to an experimenter error. When data were missing for a child for a specific task, we also excluded the score of his/her match (in age, gender and IQ) in the other group on the same task. That way we had the following number of data points per measure (RAN letters RT: 52, RAN letters accuracy: 52, RAN numbers RT: 60, RAN numbers accuracy: 58, RAN pictures RT: 60, RAN pictures accuracy: 58, RAN size-color-shape RT: 60, RAN size-color-shape accuracy: 52, RAN size-color-shape sequence: 52, action/letter/category fluency: 58, backward digit span: 58, n-back: 60, Stroop: 60).

Results

Fluency tasks

The 2x3 ANOVA showed the main effect of Fluency Type (F(1,56) = 10.024, p = .003) and Group (F(1,56) = 19.189, p < .001) resulting from the significantly better performance in the TD group, while the interaction between Fluency Type and Group was not significant (F(1,56) = .213, n.s.).
We investigated further the main effect of Fluency Type with paired-sampled t-tests, which showed a significantly better performance in category fluency than in letter fluency ($t(57) = 10.438$, $p < .001$) and in action fluency ($t(57) = 3.188$, $p = .002$) and a significantly better performance in action fluency than in letter fluency ($t(57) = 7.721$, $p < .001$) (Figure 1.).

---

Figure 1.
---

RAN tasks

Accuracy of the SLI and the TD groups did not differ significantly in the RAN task when children had to name letters, digits or objects (letters: $F(1,51) = .058$, n.s.; digits: $F(1,57) = 1.224$, n.s.; objects: $F(1,57) = 1.545$, n.s.), and naming times were also similar in the two groups (letters: $F(1,51) = .357$, n.s.; digits: $F(1,59) = .683$, n.s.; objects $F(1,59) = 1.500$, n.s.). Children with SLI were significantly slower in naming shapes together with their size and color ($F(1,55) = 41.449$, $p < .001$) but accuracy on size-color-shape names ($F(1,51) = .268$, n.s.) and on their sequence ($F(1,51) = 2.907$, n.s.) did not show a significant group difference and it was, in fact, close to ceiling in both groups (Figure 2.).

---

Figure 2.
---

Cognitive control tasks

The SLI group had shorter backward digit span than the TD group ($F(1,57) = 12.323$, $p = .001$) and their performance was lower on the n-back task as well ($F(1,59) = 9.804$, $p = .003$) even after adding non-word repetition scores as covariates (backward digit span:
F(2,57) = 7.124, p = .002; n-back: F(2,59) = 7.591, p = .001). The size of the Stroop effect did not differ between the two groups (F(1,59) = 3.371, n.s.) (Figure 3.).

Figure 3.

Associations between tasks

Correlations were conducted both in the whole group of children and in the SLI and TD groups separately to test whether the pattern of correlations differ in the two groups. Since almost all of the correlations disappeared when tested in the two groups separately, we only report the results of the analysis performed on the two groups collapsed. We conducted correlation analyses only with the results of those word production tasks (and with those measures) in which children with SLI showed a weaker performance as our main aim was the investigation of the source of difficulties of children with SLI in word production.

The partial correlation analysis with age as a control variable showed a significant correlation between letter fluency and backward digit span (r = .464, p < .001) as well as between letter fluency and n-back performance (r = .307, p = .023). A significant correlation appeared between action fluency and backward digit span (r = .283, p = .036) as well as between action fluency and the Stroop effect (r = .352, p = .007). Category fluency scores did not correlate with any of the cognitive control measures. After controlling for individual differences in non-word repetition span, the statistically significant correlation between letter fluency and n-back disappeared (r = .108, n.s.) while the correlation between the letter fluency and the backward digit span decreased but remained significant (r = .316, p = .020). The correlation between action fluency and backward digit span disappeared (r = .133, n.s.)
but action fluency and the Stroop effect still showed a significant, and in fact, stronger
correlation ($r = .503, p < .001$) (see the summary of significant correlations in Table 2.)

Table 2.

To investigate the contribution of cognitive control to word retrieval and separate it
from the effect of age, STM and vocabulary size, we conducted stepwise linear regressions
for those measures which showed a significant correlation with at least one cognitive control
task even after controlling for age and vocabulary size, namely the size-color shape RAN and
the letter fluency tasks. The dependent variable was the performance on the word retrieval
task, while the independent variables were the results on cognitive control tasks (Stroop task,
backward digit span, n-back), on the non-word repetition tasks, on the PPVT and age.

In the case of the letter fluency task the regression showed that non-word repetition
scores and backward digit span scores explained 37.5% of the performance and Stroop scores,
n-back scores, PPVT scores and age were excluded by the analysis ($R^2 = .375, p < .001$,
standardized $b_{\text{non-word repetition}} = .416$, standardized $b_{\text{backward digit span}} = .304$).

In the case of the size-color-shape RAN task the regression also showed that the best
model is the one combining non-word repetition scores and backward digit span scores which
explained 48.2% of the variance and Stroop scores, n-back scores, PPVT scores and age
were excluded by the analysis ($R^2 = .482, p < .001$, standardized $b_{\text{non-word repetition}}$

Since 12 correlations were tested in both correlation analyses, the chance of finding
false positives is relatively high. Therefore correction for multiple comparisons could be
used which would, however, increase the chance of finding false negatives. Since there
is no general practice in the literature about correction for multiple comparisons in the
case of correlations, we decided to report significant values based on both the original
and corrected alpha level in Table 2 but results will be discussed only based on the
original alpha level.
Discussion

The main aim of our study was to investigate whether the impairment of cognitive control – the ability responsible for the inhibition of irrelevant and enhancement of relevant representations – contributes to difficulties with word production under conflict in children with SLI. To explore the question we assessed word production using the fluency task and the RAN task – which require conflict resolution – together with cognitive control tasks in children with SLI and age and IQ-matched TD children. We compared performance on the word production and cognitive control tasks in the two groups and investigated associations between these performance measures. We found weaker word retrieval performance in the SLI group than in TD peers on the fluency tasks and on the size-color-shape RAN task and deficits on two cognitive control tasks: the backward digit span task and the n-back task. Weaker performance on the letter fluency task was associated with shorter backward digit span and longer naming times in the size-color-shape RAN task were associated with shorter backward digit span and n-back scores. Individual performance differences on both word retrieval tasks were better accounted for by a model combining scores on the non-word repetition and backward digit span tasks. Taken together, our results support the hypothesis that cognitive control is impaired in children with SLI and it contributes to word retrieval problems in RAN and fluency tasks, although impaired STM also has a strong effect.

As predicted by our first hypothesis, children with SLI show weaker performance on the fluency task and on one condition of the RAN task than their TD peers. The SLI group was significantly less fluent on all three conditions of the fluency task, and on the size-color-shape version of the RAN task mirrored by longer naming times. On the digits, letters and objects RAN task, SLI performance was comparable to the TD group’s. This latter finding is in accordance with earlier studies finding no group differences between children with SLI
and TD in the case of the digits, letters and objects RAN task (Bishop et al., 2009; Vandewalle et al., 2012). Children with SLI might have difficulties with the size-color-shape RAN task because three words have to be produced for one item which burdens cognitive control more than the retrieval of one word in the other versions of the task. Furthermore, the three words have to be produced in a certain order which requires sequencing skills – and cognitive control is probably necessary for that too (see more details below).

According to our second hypothesis, cognitive control is weaker in children with SLI than in TD children. Results partly supported our hypothesis: the SLI group showed a weaker performance on the backward digit span task and on the n-back task but not on the Stroop task. A possible explanation for the lack of group difference in the Stroop task is that the SLI group was affected by the conflicting information originating from the meaning of the word to a smaller degree than TD children because of poor reading skills. In other words, even if children with SLI have weaker cognitive control, the Stroop effect won’t be increased because their reading is less automatic than their TD peers’ (although this explanation is contradicted by findings from dyslexia showing bigger Stroop effects associated with poor reading skills; Faccioli, C., Peru, A., Rubini, E., & Tassinari, 2008). Our study did not target reading skills, therefore the question needs further investigation.

After finding weaker performance both on word production tasks and on cognitive control tasks in children with SLI than in TD children, we explored relationships between these impairments. We hypothesized that weaker performance on the word production tasks will be associated with weaker cognitive control measures. First, we will discuss our findings related to the three conditions of fluency task followed by the discussion of the results connecting to the size-color-shape RAN task – the only RAN task in which children with SLI performed weaker than their TD peers.
In the case of the fluency task we assumed that cognitive control is involved in the resolution of conflict originating from the activation of irrelevant words, already produced words and to be produced words. Among the three fluency conditions, only letter fluency was associated with digit span and n-back, no other fluency measures showed a relation with any of the cognitive control tasks in a meaningful way. Although letter fluency scores were associated both with backward digit span and n-back performance, the latter association disappeared when the potential confound of verbal STM was eliminated by controlling for differences in STM capacity suggesting that the relationship between letter fluency and n-back performances appeared only because both tasks require STM. The association between the letter fluency and backward digit span performance was weaker after controlling for differences in STM but it was still present suggesting the involvement of cognitive control – beyond STM – in the letter fluency task.

Associations with letter fluency and lack of associations with category and action fluency task are in accordance with neuropsychological results cited in the Introduction (Rogers et al., 1998; Robinson et al., 2012) showing impaired performance on letter fluency task but not on category fluency task in patients injured in the LIFG area responsible for cognitive control. These findings together with our results suggest that conflict is higher in the letter fluency task than in the category and action fluency tasks presumably due to the fact that most of the activated items are irrelevant in the case of the letter fluency task while several activated words are correct responses in the case of the category and action fluency tasks (See the Introduction and Robinson et al., 2012).

In the RAN task cognitive control can be recruited for the resolution of conflict originating from the activation of already produced word representations. Conflict was assumed to be higher in the case of the size-color-shape RAN task since shapes are semantically closer to each other than objects, numbers or letters. Furthermore the three
words which have to be produced for one item can also increase conflict if one of the two others is activated instead of the one which should be produced (e.g., the color is activated first when the size has to be produced). The result that children with SLI did not show weaker performance than their TD peers in the case of the letters, digits and objects RAN task but a deficit was observed on the size-color-shape task support the hypothesis that cognitive control impairment leads to impaired word production performance under conflict in children with SLI. Associations appearing both with the backward digit span and with the n-back performance confirm this assumption. After excluding the effect of STM, the associations were still present but they were weaker showing that STM is required for the size-color-shape RAN task on the one hand and corroborates the conclusion that cognitive control is involved in word production under conflict on the other hand.

Stroop performance was not strongly associated with any of the word production measures. The lack of associations between the Stroop task and the letter fluency or the size-color-shape RAN task are especially surprising since these tasks showed correlations with the other two cognitive control tasks. This result suggests that cognitive control required by the Stroop task and word retrieval tasks are not overlapping. Another explanation for the lack of correlations with the Stroop effect could be that in this task, cognitive control is confounded with reading skills. If a child’s reading is slower and less automatic then activation of the meaning of the color word might be weaker which leads to a smaller level of conflict. That way, the Stroop effect will be small even if the child’s cognitive control is weak. The question should be investigated further with using cognitive control tasks that do not require reading.

An association appeared in the unexpected direction between the Stroop task and action fluency scores: higher action fluency scores were associated with greater Stroop effects. This result suggests that weaker cognitive control led to higher action fluency scores.
Although this result is in contrast with our predictions, one could argue that strong cognitive control can have a negative effect on fluency performance when cognitive control requirements of the task are low. If most of the activated word representations are relevant, but the system inhibits them since they are not the current target words, then it will take more time to retrieve these word representations when the participant wants to produce them. With weaker cognitive control, inhibition is not that efficient, and it takes less time to retrieve the words later which leads to a better fluency performance. It is not clear, however, that if that is the case, why the same positive correlation was not observed between the Stroop effect and category fluency where cognitive control demand is also low according to neuropsychological results and the lack of correlations with the n-back and backward digit span task in our study. The potential negative consequences of effective cognitive control on fluency performance and on word production in general is another area which would be important to investigate further.

The investigation of relative contribution of cognitive control, STM, vocabulary and age to word retrieval under conflict showed that the model including the backward digit span and the non-word repetition scores explain the performance both on the letter fluency task and the size-color-shape RAN task the best. The fact that non-word repetition is present in the model suggests that individual differences in STM explain a significant part of individual differences in the word retrieval performances. The presence of backward digit span in the model support our hypothesis that cognitive control contribute to word retrieval under conflict. A significant difference between the backward digit span task and the other two cognitive control tasks is that the backward digit span has a stronger STM demand, therefore the result that this measure was included in the model suggests that the ability to apply cognitive control when the STM load is high is especially important for word retrieval under conflict. Furthermore, we can infer from these results that weaknesses in STM and cognitive
control in children with SLI contribute to their problems in word retrieval under conflict. Vocabulary size and age as well as cognitive control measured by the n-back and Stroop tasks seem to play a less important role.

Conclusions

In summary, our results are in accordance with the hypothesis that cognitive control is involved in the resolution of conflict between irrelevant word representations and the target word representation during word retrieval. They also support the view that cognitive control is impaired in children with SLI and this impairment contributes to word production difficulties.

Nonetheless, several questions remained unanswered in the current study which need further investigation. First, we did not address the question whether cognitive control involved in word retrieval is a domain-general process or is specific to linguistic stimuli. Future studies should test cognitive control with nonlinguistic stimuli together with language performance to answer this question. The lack of associations between the Stroop task and word production tasks has to be explored with using a modified Stroop task that does not require reading skills. Furthermore, motivated by the positive relationship between the size of the Stroop effect and performance on the action fluency task, the negative role of strong cognitive control on word production has to be investigated directly. Although our results generally support the hypothesis that cognitive control is impaired in SLI and word production problems can be partly accounted for by this impairment, some of our results – e.g., the lack of group differences in the Stroop task, lack of associations between word production and the Stroop tasks and the finding that some associations between word production and cognitive control disappeared after including STM scores in the analysis – support the opposite hypothesis. There is no consensus about the issue in the literature either:
beyond results showing a cognitive control impairment in SLI, there are several studies suggesting that cognitive control is intact in SLI – including two of our previous studies (Lukács et al., 2016; Ladányi & Lukács, 2016). Contradictory results in the current study and in Lukács et al. (2016) are especially surprising since participants in the two studies partly overlapped and some of the tasks were used in both studies therefore there was an overlap also between the scores included in the analyses. Together with contradictory findings in the literature in general, the fact that in our studies a group difference is present with a group of children but disappears with another partly overlapping group of children suggest that cognitive control might be impaired in some children with SLI and it might contribute to their language problems but it does not seem to be a general impairment and leading cause of SLI. Word retrieval is probably impaired due to another or several other reasons in SLI, for instance the impairment of STM is probably an important factor which is supported also by the current study but the impairment could be more pronounced when conflict is present due to weaker cognitive control in some children.

Despite these controversies, the results show that cognitive control impairment might contribute to lexical impairments in (at least some) children with SLI, suggesting that training targeting cognitive control might be helpful in improving lexical skills in SLI.

Acknowledgments

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

References


Jaeggi, S. M., Buschkuehl, M., Perrig, W. & Meier, B. (2010). The concurrent


Kas, B., & Lukács, Á. (in preparation). Magyar Mondatutánmondási Teszt [Hungarian Sentence Repetition Test].


Weckerly, J., Wulfeck, B., & Reilly, J. (2001). Verbal fluency deficits in children with specific language impairment: Slow rapid naming or slow to name? *Child*


Tables, figures

**Table 1.** Demographical data and screening results of the SLI and TD groups (PPVT: Peabody Picture Vocabulary Test, TROG: Test for Reception of Grammar)

<table>
<thead>
<tr>
<th></th>
<th>TD mean (SD)</th>
<th>SLI mean (SD)</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>107 mo (13 mo)</td>
<td>107 mo (14 mo)</td>
<td>F(1,60) = 0.008, n.s.</td>
</tr>
<tr>
<td>IQ (Raven)</td>
<td>106 (9.86)</td>
<td>102 (9.98)</td>
<td>F(1,60) = 2.944, n.s.</td>
</tr>
<tr>
<td>PPVT</td>
<td>124 (2.2)</td>
<td>94 (3.3)</td>
<td>F(1,60) = 42.96 p &lt; .001</td>
</tr>
<tr>
<td>TROG</td>
<td>77 (0.4)</td>
<td>69 (0.9)</td>
<td>F(1,60) = 65.70 p &lt; .001</td>
</tr>
<tr>
<td>sentence repetition</td>
<td>37 (0.8)</td>
<td>21 (1.4)</td>
<td>F(1,60) = 94.78 p &lt; .001</td>
</tr>
<tr>
<td>non-word repetition</td>
<td>6.4 (0.2)</td>
<td>3.4 (0.2)</td>
<td>F(1,60) = 120.24 p &lt; .001</td>
</tr>
</tbody>
</table>
Figure 1. Means of total number of correct answers in the action, letter and category fluency tasks in the TD and SLI groups.
Figure 2a-2b. Means of total naming times (2a) and means of correct answers (2b) in the size-color-shape RAN task in the TD and SLI groups.
Figure 3a-3b-3c. Performance on the a) backward digit span task, b) n-back task and c) the Stroop task in TD children and children with SLI.
Table 2. Summary of significant correlations between word production tasks and cognitive control tasks (correlations marked with an * were significant also based on the corrected alpha level (.004))

<table>
<thead>
<tr>
<th></th>
<th>Correlations with age as a control variable</th>
<th>Correlations with age and non-word repetition as a control variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>letter fluency – backward digit span</td>
<td>( r = .464, p &lt; .001^* )</td>
<td>( r = .316 p = .020 )</td>
</tr>
<tr>
<td>letter fluency and n-back</td>
<td>( r = .307, p = .023 )</td>
<td>( r = .108, \text{n.s.} )</td>
</tr>
<tr>
<td>action fluency – backward digit span</td>
<td>( r = .283, p = .036 )</td>
<td>( r = .133, \text{n.s.} )</td>
</tr>
<tr>
<td>action fluency – Stroop effect</td>
<td>( r = .352, p = .007 )</td>
<td>( r = .503 p &lt; .001^* )</td>
</tr>
<tr>
<td>size-color-shape RAN – backward digit span</td>
<td>( r = -.571, p &lt; .001^* )</td>
<td>( r = -.436 p = .002^* )</td>
</tr>
<tr>
<td>size-color-shape RAN – n-back</td>
<td>( r = -.481, p &lt; .001^* )</td>
<td>( r = -.301 p = .030 )</td>
</tr>
</tbody>
</table>
6. General discussion

The dissertation investigated cognitive control in children with SLI. More specifically, we aimed to find out whether children with SLI show weaker performance on cognitive control tasks and on language tasks potentially requiring cognitive control than their TD peers, and whether measures of cognitive control are associated with language measures in children with SLI as well as in TD children. To explore these questions we conducted four experiments with different, partly overlapping groups of children with SLI between 5 and 11 years of age in which three main hypotheses were tested.

According to the first hypothesis we expected cognitive control to be weaker in children with SLI than in TD children. Study 1 showed the expected group difference in some of the cognitive control tasks but as subsequent analysis showed, the difference was not the result of weaker cognitive control but weaker short-term memory in children with SLI than TD children. Results from Study 3 and Study 4, however, partly supported, the first hypothesis because children with SLI showed weaker cognitive control than TD children in some of the cognitive control tasks even if short-term memory abilities were taken into account.\(^5\)

The second hypothesis stated that children with SLI will show weaker performance on language tasks requiring cognitive control. Study 3 and Study 4 supported this hypothesis with indicating problems in the comprehension of sentences with anaphors and in word production under conflict. Importantly, performance patterns of children with SLI on the sentence comprehension task were very similar to those of younger children in earlier studies.

\(^5\) Note that group differences with including short-term memory scores as covariates are reported only in Study 4. However, since the same children participated in the two studies – only one child with SLI and his TD pair who participated in Study 4 did not participate in Study 3 – group differences are presumeably present after including short-term memory scores also in Study 3.
suggesting a typical developmental route with a delay in SLI. In contrast with Study 3 and Study 4, results of Study 2 suggest that word production under conflict is not impaired in children with SLI.

Finally, we hypothesized that cognitive control and language measures will be associated with each other. Study 3 and Study 4 partly supported this hypothesis with showing better comprehension of sentences with anaphors and better word production under conflict in children with stronger cognitive control. The associations appeared, though, only with some of the cognitive control measures.

Controversial results of our studies about cognitive control in children with SLI are in accordance with earlier findings in the field. As it was discussed in the Introduction, several studies showed impairments on various paradigms involving cognitive control in children with SLI (e.g., Lum et al., 2012; Vugs et al., 2014; Ellis Weismer et al., 1999; Archibald & Gathercole 2006; Mainela-Arnold & Evans, 2005; Marton et al., 2006; Marton et al., 2007; Montgomery & Evans, 2009; Evans & Pollack, 2011; Im-Bolter et al., 2006; Marton et al., 2014), while others showed as good cognitive control as in TD children (e.g., Archibald & Gathercole, 2006; Engel de Abreu et al., 2014, Reichenbach, et al., 2016). One possible source of discrepancies between the results of different studies is the fact that research about cognitive control in general is limited and it is not clear whether different tasks measuring representational conflict resolution abilities require the same type of conflict resolution process or there are different processes for the different situations/stimuli. In Study 1, Study 3 and Study 4 children with SLI showed as good performance as TD children on Stroop tasks similarly to a previous study using a Stroop paradigm in children with SLI (Reichenbach et al., 2016). Group differences appeared, though, in the case of n-back paradigms and backward digit span tasks. These tasks require the manipulation of several items in short-term memory and it might require a different conflict resolution process than the Stroop task.
which does not have a short-term memory demand. Importantly, group differences were present in the case of the n-back and backward digit span tasks even after controlling for short-term memory in Study 3 and Study 4. This result supports the idea that different types of cognitive control processes are recruited for the manipulation of several items in the short-term memory than for situations when short-term memory demand is low. Future research should explore the potential types of processes recruited for the resolution of representational conflict appearing in various situations between different types of stimuli.

Beyond the potential variance in the type of conflict resolution required for the tasks used to assess cognitive control, the variability of cognitive control in children with SLI might also contribute to contradictory results between different studies in the dissertation and in the literature. While we found a group difference in the backward digit span task and in the n-back task, some of the earlier studies did not which suggests that even if there is a cognitive control impairment in some children with SLI, it is not general. This heterogeneity is not surprising, as the symptoms of children with SLI are highly heterogeneous in general.

The relationship between cognitive control and certain language processes is also only partly supported by our results. Previous studies in adults and neuropsychological patients showed that cognitive control plays an important role in various language processes. Associations between some of the cognitive control measures and the comprehension performance of sentences with anaphors in Study 3 as well as some of the word production measures in Study 4 are in accordance with that view – although these relationships were moderate. In addition, we did not find any meaningful relationships between the Stroop measures and word production tasks in Study 4 and some of the associations with other cognitive control tasks were, in fact, the result of their shared short-term memory demand with the language task. One explanation for the lack of strong relationships between cognitive control and language measures is the methodological difference between our
studies and adult studies supporting the role of cognitive control in language. While those works were lesion studies or brain imaging studies, we investigated associations between individual differences with behavioral measures of cognitive control and language. Since cognitive control is presumably less impaired in children with SLI than in the case of people with a frontal brain injury, variance might be too low for finding associations with language measures.

These results suggest that weaker cognitive control might be one factor which can lead to language problems in some areas if it co-occurs with other linguistic and nonlinguistic impairments. One of these other potential impairments is low verbal short-term memory capacity which seems to be a more general characteristic of children with SLI shown by earlier results discussed in the Introduction and also by our studies. In Study 1 lower performance on cognitive tasks was the result of lower verbal short-term memory span in children with SLI than in TD children and in Study 3 and Study 4 several associations between cognitive control and language measures became weaker or even disappeared when variation in verbal short-term memory span was taken into account. These results support the view that verbal short-term memory capacity is reduced in children with SLI and it has a role in their language problems in various domains.

Although the main focus of the dissertation was cognitive control in SLI, Study 3 and Study 4 also aimed to contribute to the investigation of the role of cognitive control in the comprehension of anaphors and to word production in general. The result that performance on the anaphor resolution task is associated with cognitive control measures suggests that cognitive control is involved in anaphor comprehension and the relationships between cognitive control and word production tasks confirm previous findings that it might have a role in word production under conflict too. Furthermore, with Study 2 we were able to extend the research on word production in TD children with a novel finding, beyond investigating
children with SLI. We found that the effect of presence of conflict in word production is very similar in primary school children as in adults in previous studies suggesting relatively efficient conflict resolution during word production at this age.

Studies presented in the dissertation aimed to contribute both to the SLI literature and to the line of research investigating relationships between cognitive control and language with targeting unexplored questions and using paradigms which have been rarely used in the literature. In Study 3 we decided to investigate the role cognitive control in the comprehension of sentences with anaphors instead of using temporally ambiguous sentences which are the typically used stimuli for investigating this question. In Study 4 we aimed to study the role of cognitive control in word production with using methods beyond the picture naming paradigms with conflict manipulations. This way we could extend the scope of language processes and tasks in which cognitive control might play a role. At the same time, using novel tasks made our research more challenging due to the lack of earlier results about these paradigms, especially in SLI research.

Future research should investigate the controversial issues mentioned above further. First, it would be important to test many different cognitive control paradigms on the same group of participants to determine overlapping and diverse processes recruited by them in typical populations. The involvement of brain imaging methods would probably contribute a lot to this line of research. Using several cognitive control paradigms in children with SLI would be also important to find out whether there are cognitive control tasks which are more difficult for these children than others. Second, although recruiting children with SLI can be difficult, it would be crucial to test large groups of children with SLI with the same paradigms to be able to explore the possibility that a subgroup of children with SLI has a cognitive control impairment, while in others cognitive control is intact.
Despite of the controversial findings, the studies in this dissertation are important contributions to research targeting SLI. Although our results suggest that even if cognitive control is impaired in some children with SLI, and this impairment is probably mild, cognitive control seems to contribute to language problems both in the sentence comprehension and word production domain. Therefore if future research will confirm that – at least in the case of some children with SLI – impairments in cognitive control contributes to language problems, training of cognitive control should be considered as an option beyond the training other non-linguistic and linguistic abilities in clinical practice.

Conclusion

The main finding of studies in the dissertation is that even when cognitive control is weaker in children with SLI than in TD children, the impairment is mild. As our results also showed, weaker cognitive control is associated to a moderate degree with weaker performance in sentence comprehension and word production, arguing that cognitive control might be one of the factors which contribute to weaker language abilities in some children with SLI. Together with previous studies showing impairments in different linguistic and non-linguistic domains, our results suggest that symptoms of SLI appear when impairments in several linguistic and non-linguistic abilities co-occur. They are, therefore, compatible with all the theories proposing the impairment of a cognitive ability in children with SLI but contradict with grammar-specific theories assuming the impairment of a certain part of the grammatical system leads to the symptoms of SLI. This multifactorial view of SLI is in accordance with Bishop and colleagues’ (e.g., Newbury, Bishop, & Monaco, 2005) theory according to which the combination of several genetic and environmental factors leads to symptoms of SLI, although we assume that there are much more potential factors than those proposed by the authors – the impairment of abilities described in the 1.2.2. section: working
memory, processing of rapidly changing stimuli or procedural learning, as well as cognitive control and other abilities which haven’t been recognized yet can also contribute to SLI. The multifactorial nature of SLI implies that causes of SLI can show great variation across children, therefore underlying causes should be discovered on an individual basis with testing all the abilities which can potentially contribute to the symptoms and impaired linguistic and non-linguistic processes should be targeted during speech therapy.

7. References


Kas B., Lukács Á., & Szentkuti-Kiss K. (in preparation). A szórend feldolgozása specifikus nyelvfejlődési zavarban [The processing od word order in specific language impairment].


napkins, and Broca’s area. Poster presented at the Seventeenth Annual CUNY Conference on Human Sentence Processing, College Park, MD.


