Specificity in language: cognitive functions and impairments

Nyelvi specificitás: kognitív funkciók és zavaraik

Ágnes Lukács

Akadémiai Doktori Értekezés

2017

Table of contents

Acknowledgments

First of all I want to thank Bence Kas, Ferenc Kemény and Enikő Ladányi, my former and current PhD students for all the work they did that contributed to this dissertation and for giving me inspiration to move on with the research. Together with Anna Babarczy, they also took the time to read the manuscript and helped with their comments and suggestions in finalizing it. I am indebted to Csaba Pléh, Laurence B. Leonard and Michael Ullman, my senior co-authors who also taught me a lot during our collaborations. Lilla Zakariás, Gyula Demeter and Attila Keresztes made the research on aphasia possible and laid the grounds for further collaborations in new directions. Borbála Győri, Kata Fazekas and Zsófi Kígyóssy, together with many students and research assistants were essential in getting the research under way, and in collecting and organizing the data.

I am also very grateful to Mihály Racsmány, Szabolcs Kéri, Dezső Németh and Ildikó Király, who together with many other friends and my family, gave me constant encouragement during the writing of the dissertation. My great colleagues at the Department of Cognitive Science of Budapest University of Technology and Economics (Anna Babarczy, Mihály Racsmány, Szabolcs Kéri, Ágnes Szőllősi, Gyula Demeter, Péter Pajkossy, Kornél Németh, Márta Zimmer, Péter Simor, Bertalan Polner, Anna Torma, and Csaba Müller), provided me with a supportive and inspiring academic atmosphere.

The research presented in the dissertation would not have been possible without the financial support over the years from the Hungarian Scientific Research Fund (Országos Tudományos Kutatási Alapprogramok, OTKA TS 049840 to Csaba Pléh and OTKA K 83619 to Ágnes Lukács), from the NIH (national Institute of Health, National Institute on Deafness and Other Communication Disorders USA; R01 DC00458 to Laurence B. Leonard), and from the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

I am especially thankful to the children for their participation and to the schools for accommodating our research. I very much appreciate the help of the speech and language therapists in screening and organization at the Dr Nagy László Institute of Special Education in Kőszeg, in the ELTE Special Preschool and Early Intervention Centre, and in the Zölderdő Preschool for Speech Therapy and Nature Preservation.

List of abbreviations

AC	age control
AGL	artificial grammar learning
EF	executive functions
LIFG	left inferior frontal gyrus
MRA	Morphological Richness Account
PCL	probabilistic category learning
SLA	second language acquisition
SLI	specific language impairment
SRT	serial reaction time task
STM	short term memory
TD	typically developing, typical development
VC	vocabulary control
WM	working memory

Background and motivation

The specificity of linguistic mechanisms, i.e. the notion that these processes only play a role in the acquisition or use of language, and are not engaged in other areas of cognition, has been one of the outstanding issues in cognitive science since the cognitive revolution (see, for example, Blank et al., 2014; Chomsky, 1968/1995, 2010; Christiansen & Chater, 2008, 2015; Elman et al., 1996; Fedorenko & Thompson-Schill, 2014; Fedorenko et al., 2011; Friederici, 2002; Fodor, 1983; Grodzinsky & Santi, 2008; Hauser et al., 2002; Jackendoff, 2002; Karmiloff-Smith, 1996; 2009; MacWhinney, 1999; Müller, 1996; Pinker, 1991, 1999; Pléh, 1998, 2000, 2013; Pléh & Lukács, 2002; Tomasello, 2003; Ullman, 2001; van der Lely, 2005, among many others). This is partly due to the fact that for a long time language has been a central focus of study in discussions and research revolving around domain specificity and architectural issues of more general theoretical relevance. It also has historical reasons: influential statements by Chomsky and Fodor, prominent figures of the cognitive revolution, emphasized the specificity of linguistic abilities. As part of their legacy, specificity is often intertwined with features that are not necessarily implied by or associated with domain specificity, namely innateness, anatomical and functional modularity, and species specificity. The idea of a specific innate language acquisition device (LAD) that is dedicated to language learning and drives rapid language acquisition (Chomsky, 1968/1995) has been a strong influence in cognitive science. In the adult brain, the LAD is paralleled by a specialized language (or more specifically, grammar, or syntax) module which is autonomous and independent of other areas of cognition (see references above).

Although the proposal of a domain- and species-specific innate language module and language acquisition device had an inspirational effect not only on experimental psycholinguistics, but on the entire field of cognitive science, the idea was the target of criticism from its birth, and since then, alternative theories have been getting more and more empirical support. Converging results from typical language acquisition and developmental disorders have shown that neither innate language specific representations, nor specific language acquisition mechanisms are necessary to explain relatively rapid mastery of language (e.g. Christiansen & Chater 2008, 2015; Elman et al., 1996; Karmiloff-Smith, 1996; MacWhinney, 1999; Tomasello, 2003). Empirical arguments and their implications have been synthesized under the constructivist approach to language acquisition (Tomasello, 2003), in which grammatical development is not isolated from other areas of cognition, and is not driven by language-specific mechanisms. Constructivists argue that instead of specialized acquisition processes working only within the domain of language, language acquisition is supported by domain-general learning and processing mechanisms like statistical learning, distributional analysis, cultural learning, categorization, intention reading, structure mapping and analogical extension together with generalization skills (MacWhinney, 1999; Tomasello, 2003).

The studies presented in the dissertation fit into the line of empirical research that questions different aspects of the proposal of specificity of language. Because of their limits in scope, they mainly focus on one of the arguments supporting domain specificity of a mechanism, namely, on the selectivity of impairments of language.[1] Domain-specificity of language implies the existence of developmental or acquired disorders in which language (or part of it) is *selectively impaired* but cognitive skills outside the domain of language remain intact. The aim of the dissertation is to concentrate primarily on this area, and to present studies examining the validity of the argument of selective language impairment in the debate about the domain specificity of language. In 12 studies, we will examine issues about the specificity of language processing and linguistic representations from different aspects of language and cognition, and cast doubt on the necessity of specialized mechanisms in acquisition and processing. The questions are in most part going to be addressed by studies of specific language impairment (SLI), a developmental disorder of language, approaching the problem from two aspects: 1) through a detailed mapping of language skills, we want to test whether there is evidence of a selective grammatical deficit a) within a specific area of grammar or b) within grammar itself and 2) through the assessment of some nonlinguistic cognitive functions, we were looking for evidence for or against the specificity of impairments of language. This larger body of results is then complemented by results from aphasia, the acquired impairment of language and by a typical developmental study in which we address the question of the existence of a critical period often associated with specificity.

In the introduction to the studies, we will first explain what SLI is and why it was chosen as a target population for the majority of studies, introducing general questions on linguistic and non-linguistic abilities. We will also briefly define aphasia, and relate it to questions about the specificity of language to motivate our study on executive functions in

[1] As a consequence, important aspects of specificity and related issues are not going to be discussed in detail. We will briefly return to the problem in the discussion, in addressing the limitations of the dissertation.

this disorder, and then turn to general learning mechanisms and executive functions as potential candidates for domain-general processes supporting language acquisition. We will also address methodological issues that are relevant to the studies. The introduction is closed by exposing the questions addressed in the studies. The results are summarized (by thesis points, closed by a General discussion) in the discussion section following the individual papers.

Specific Language Impairment

Specific language impairment (SLI) is a developmental disorder that involves a significant delay and a primary deficit in the acquisition of language in the absence of any hearing deficits, neurological disorders, emotional and social problems, environmental deprivation or intellectual disability that could account for the language problems. Because of this cognitive pattern, the study of SLI is often motivated by the promise of learning something about the nature and development of the human language ability. Studies and accounts of SLI have often involved the larger perspective of cognitive architecture and sparked heated debates not just about the specificity of the language deficit in SLI but also about the specificity of language in general. Some have proposed that in SLI, language is selectively impaired in an otherwise intact cognitive system (as the original definition and the term implies); according to these statements, the main cause of SLI is a primary deficit in the abstract representation of language, that is, domain-specific language competence is impaired some way (e.g. Clahsen & Hansen, 1997; Gopnik & Crago, 1991; Rice et al., 1995; van der Lely, 2005; van der Lely & Stollwerck, 1996; for reviews, see e.g. Leonard, 1998/2014; Lukács et al., 2014). Others argue that the inability to acquire language in the typical way is a result of a general or specific cognitive processing limitation, e.g. an impairment of working memory (e.g. Adams & Gathercole, 2000; Montgomery, 2002; for a recent review, see Gillam et al., 2017) or auditory perception (Tallal & Piercy, 1973; Tallal et al., 1996). Numerous results show that SLI is often associated with impairments in several nonlinguistic domains (executive functions, motor organization, procedural learning of sequential and nonsequential information and associative learning), but the nature, extent and generality of these deficits is yet unclear, as is their relationship with language abilities (e.g. Henry et al., 2012; Kapa & Plante, 2015; Vargha-Khadem et al., 1995; Windsor, 2017; Ullman & Pierpont, 2005).

Although the initial question was whether there is evidence in SLI of a language- or grammar-specific impairment, and, correspondingly, of an autonomous language module, questions and research focuses have shifted since.[2] In the light of research over the last 25 years it is now clear that there are accompanying non-linguistic deficits in SLI (see below), the new focus is on exploring the nonlinguistic domain-general abilities that tend to be vulnerable in SLI and on how their impairments affect language: are they the primary cause of language problems and explain why they occur, or are they associative impairments that tend to co-occur with language because of shared networks or anatomical proximity that do not have a causal role in language symptoms? It is also believed that results from a developmental impairment inform theories of typical cognition, and teach us about how these mechanisms contribute to the acquisition, processing and production of typical language (even taking arguments for differences between processes in typical and atypical acquisition into account, e.g. Karmiloff-Smith, 1996).

In spite of the long line of research on SLI (of which large segments are reviewed in the introductions to specific studies), there are still many open questions both in theoretical and practical areas. One of the central problems complicating SLI research is the heterogeneity of the disorder. Suggestions have been put forward with more fine-grained classifications and subtypes of language impairment (see e.g. Leonard, 1998/2014; Bishop, 2014, and Bishop et al., 2016 for reviews and discussions). Examples of proposed subtypes include grammatical or g-SLI (van der Lely, 2005), expressive SLI (Whitehurst et al., 1992), semantic-pragmatic SLI (Bishop & Adams, 1989) or SLI with articulatory and non-verbal deficits (Vargha-Kadem et al., 1995). Attempts have been made to identify subtypes based on psychometric tools like cluster analysis (e.g. Rapin & Allen, 1983: phonological-syntactic deficit syndrome, lexical-semantic deficit syndrome and word deafness; Conti-Ramsden, Crutchley & Botting, 1997: 7 clusters; Daal, Veerhoven & van Balkom, 2004: lexical-semantic, speech production, syntactic-sequential and auditory perception subtypes), but while the overlap between the identified subtypes and profiles is clearly present, there is still no agreement on what new categories should replace SLI. The question is further complicated by observations showing that profiles also change over time in the course of development (Conti-Ramsden & Botting, 1999).

[2] The focus has similarly shifted in the study and discussions of typical language acquisition and processing as well.

Some of the suggested terminology also reflects theoretical issues: since the overwhelming majority of studies found deficits outside language, there have been attempts to discard 'specific' from the label, and refer to children with such language problems as having 'language impairment' or primary language impairment (Tomblin et al., 1996). Because of the widespread use of the term SLI by then, and also because it was common understanding among the majority of the research community that SLI is not specific to language, these suggestions to change the label has failed so far. Following convention, we will also use the term SLI in the dissertation.[3]

Just like in every area of experimental research on language, the majority of SLI studies focus on English; diagnostic protocols and tools are mostly available for English, and many accounts of SLI were informed by studies of English-speaking children with language impairment. Crosslinguistic research on SLI is growing exponentially, but we still have not identified diagnostic markers associated with different language types, and there are still many languages where profiles have not been outlined and research and diagnostic tools are missing. Typological differences between languages are also important in evaluating theoretical proposals (as we hope our studies on Hungarian will also demonstrate).

Besides the theoretical relevance of the above issues, identifying areas of vulnerability within language in different languages and identifying different subtypes requiring different focuses in therapy is important for practitioners as well, just as identifying the causes of the deficit and the mechanisms and processes that are impaired. Finding areas of impairment -- both linguistic and nonlinguistic -- that are causal in problems of children with SLI is a first step ahead on a way to devising more effective intervention procedures to help children strengthen their areas of special weakness.

Linguistic abilities in SLI

Specific language impairment is characterized by different areas of strengths and weaknesses of linguistic domains, showing significant crosslinguistic variation across different language types. Although there are significant individual differences in children with SLI in spite of all meeting the diagnostic criteria, the majority of research shows that at

[3] In our earlier studies (Studies 1-5. in the dissertation), we used the term LI instead of SLI. We made an attempt to stick with it in Study 6, but the editor convinced us otherwise based on the arguments cited above in the main text. In later studies, the term SLI is used following common practice without theoretical commitment to the specificity of the disorder.

least English-speaking children with SLI, the deficit is most severe in the use of grammatical morphemes (bound morphemes such as the past tense *-ed* and third person singular *-s* verb inflections, and function words such as *the* and *is*), accompanied by less severe problems in other areas of language. These limitations in the use of grammatical morphemes is evident not only relative to age-expectations, but also compared to TD children matched on MLU (e.g., Johnston & Kamhi, 1984; Leonard et al., 1992; Leonard et al., 2003; Oetting & Horohov, 1997; Rice & Wexler, 1996) persisting into school age (Marchman, Wulfeck, & Ellis Weismer, 1999; Norbury, Bishop, & Briscoe, 2001; Rice, Wexler, & Hershberger, 1998) or even longer (e.g., van der Lely, 1997). Children speaking typologically different languages show distinct profiles to some extent. In languages with rich verb morphology (i.e. Italian) verbal suffixes seem to be generally intact, while the omission of object clitics and articles are typical symptoms (Bortolini et al., 2002). In German children with SLI, atypical word order patterns and deficits in verbal conjugation patterns are observable (Clahsen, 1999) while the case marking system seems to be generally unaffected (Eisenbeiss et al., 2005).

While SLI is often described as a developmental language disorder with the primary deficit in the area of grammar, it has been documented that in many cases lexical deficits co-occur with grammatical problems (although they might be less severe than deficits in grammar). Late appearance of first words is one of the main risk factors for language impairment and vocabulary size often lags behind age-based expectations in preschool and school-aged children with SLI (Bishop, 1997; Watkins, Kelly, Harbers, & Hollis, 1995; Trauner, Wulfeck, Tallal, & Hesselink, 1995). These findings, together with experimental results on lexical processing and production (reviewed in detail in **Study 11**) show that the acquisition, production and processing of words are slower and more error-prone in SLI than in typical development.

Although there is agreement in the literature that problems with grammatical morphology are central in English-speaking children with SLI, the source of this difficulty is subject to debate (see Leonard, 1998/2014; Lukács et al., 2014). Accounts that propose a specific grammatical deficit in SLI highlight one area of grammar and focus on that as a potential target of a selective grammatical impairment, proposing to explain the linguistic symptoms of SLI as an impairment of universal grammar. In most of these accounts, morphology or one specific area of morphology is selectively vulnerable. Gopnik and Crago (1991), based on the first detailed examination of the famous KE family found a selective deficit[4] in the use (in spontaneous speech and elicited production with nonwords) and processing (comprehension and grammaticality judgments) of suffixes marking tense and agreement. The agreement deficit is seen as a central problem in several later accounts as well. According to Clahsen’s early proposal (1991) the primary deficit in SLI concerns agreement features without a semantic interpretation. Rice, Wexler and Cleave (1995) also emphasize the immaturity of the language system in SLI with respect to tense and agreement marking: the authors think that language acquisition in SLI is stuck at the so-called ‘optional infinitive’ stage. This is the period when typically developing children (around 2-3 years of age) consider the use of nonfinite verb forms optional in contexts that require the obligatory use of finite verb forms. In English, this optional stage is marked by the frequent omission of third person singular *-s* and past tense *-ed*.

In addition to the difficulty with agreement, van der Lely’s proposal of Representational Deficit for Dependent Relations (van der Lely, 1994; van der Lely & Stollwerck, 1997) also predicts difficulties in case marking and anaphor interpretation. It proposes that children with SLI have a broad syntactic computational deficit that leads to weaknesses in structure-dependent relationships such as 1) different forms of agreement 2) anaphor interpretation, 3) case marking due to weaknesses in feature checking for long-distance dependencies during constructing the syntactic tree. An extended more recent formulation of this account called the Computational Grammatical Complexity Hypothesis (CGC, Marshall & van der Lely, 2007) claims that children with SLI have a deficit in structural complexity that extends beyond syntax to include morphology and phonology as well. The CGC hypothesis also argues that not all kinds of dependent relations are expected to be impaired in SLI, only complex ones involving movement chain formations.

Other approaches, based on models of functional and neurocognitive dissociation of the grammar and the lexicon, argue for the selective impairment of grammar associated with the frontal brain regions (traditionally called Broca's area) and for the intactness of the lexicon associated with the posterior linguistic areas within the brain (traditionally called Wernicke's area) in SLI. Within this framework, Pinker (1991, 1999) and Clahsen (1999) assume that it is the extraction and application of productive algorithm-like rules of

[4] Selective deficit *within* language; later studies described severe problems with articulation and fine motor organization (Vargha-Khadem et al., 1995, 1998).

grammar which is difficult in SLI, and they do not expect to find problems in the area of vocabulary and lexical access.

Approaches not taking language and grammar to be isolated from the rest of cognition propose more general processing difficulties in the background of SLI (for extensive reviews, see Bishop, 2006; Leonard, 1998/2014; Lukács et al., 2014). One of the earliest proposals claimed that the primary impairment concerns the processing of rapidly changing auditory stimuli (Tallal & Piercy, 1973; Tallal et al., 1996). Another approach, the Morphological Richness account (MRA; Dromi et al., 1999; Leonard, 1998; Leonard et al., 1987; Leonard et al., 2007) assumes in the first place that children with SLI have a limited processing capacity and a general limitation in language learning ability, but it also has more specific predictions about the differential vulnerability of different areas of grammar, and also for crosslinguistic differences in that regard. The properties of the particular language that they acquire will have a significant impact on the pattern of language difficulties; as a consequence, great problems with grammatical morphology are not characteristic of all language types. The MRA argue that in English, inflections are not very frequent and are not central in grammatical marking, stems often occur in their bare form, so children devote their limited resources to processing aspects of grammar that are more important (word order in English). Therefore, less capacity remains for the learning of grammatical morphology, requiring more encounters with grammatical morphemes before they can be learned. Crosslinguistic research on many languages supports this account: in languages with a rich morphology (such as Italian or Hungarian) and a central role of inflections in grammatical marking, resources are mostly devoted to the acquisition of morphology, and children with SLI have less problems with morpheme use. Difficulties with morphology can occur in these languages as well, but instead of the selective problems with morphological marking of certain grammatical relations, we expect greater difficulties with morphemes that are difficult from a processing point of view, i.e. that are long, rare, complex, and their functions are non-transparent.

Another formulation of the reduced processing capacity approach comes from Baddeley, Gathercole and Papagno (1998)'s working memory deficits hypothesis. They argue that the reduced capacity of working memory is responsible for grammar and vocabulary problems in SLI. Children with SLI are able to hold and manipulate less verbal information in their short term memory than typically developing peers. Reduced nonword repetition span, a common measure of short term memory capacity is so common that it is considered to be a marker in SLI (although it is also often observed in dyslexia).

Bates and colleagues (Marchman & Bates, 1994) also argue for a processing deficit in SLI, but they have a model of language in which the grammar and the lexicon are not isolated from each other: knowledge of grammatical rules emerges based on a ’critical mass’, a critical vocabulary size. Because of the general processing deficit in SLI, vocabulary acquisition proceeds slower, and it takes much longer for the lexicon to reach the critical size for grammatical generalizations to emerge. This critical size could also be larger in SLI. That is, on this account, the grammatical deficit is a consequence of the lexical deficit resulting from the processing difficulty.

Several studies in the dissertation (**Studies 1, 2, 3-6**) were designed to test and compare predictions of some of these linguistic and processing accounts (see below), since Hungarian, differing from the previously examined languages in several respects, provides exceptional opportunities for testing their predictions further.

Nonlinguistic abilities in SLI

Research in the past two decades has shown that SLI in many (if not in most) cases turns out to be not as specific to language as originally claimed. Impairments are often observed in the coordination of oral and fine finger movements, categorization, sequencing, and most evidently in working memory (in particular verbal working memory: Leonard, 2014; Lum, Conti-Ramsden, Page, & Ullman, 2012; Marton & Schwartz, 2003; Montgomery, Magimairaj, & Finney, 2010; Ullman & Pierpont, 2005) and executive functions (Bishop & Norbury 2005; Gillam et al., 2017; Im-Bolter et al., 2006; Leonard, 1998/2014; Montgomery, 2003; Windsor et al., 2017). Theories that take grammar to be a functionally and anatomically distinct module which can be selectively impaired in SLI cannot account for these difficulties and proposals of more general limitations in processing capacity (see above) also fail to cover the range of linguistic and nonlinguistic deficits observed in SLI.

One of the theories of SLI, the Procedural Deficit Hypothesis (PDH, Ullman and Pierpont, 2005) tries to incorporate linguistic and nonlinguistic impairments into a unified theory and a neurobiological model. This hypothesis is based on the Declarative/Procedural model (DP model, Ullman, 2001, 2004, 2016) of language, which claims that there is a clear within-language dissociation between the grammar and the lexicon, since they are functions of different memory systems. Grammar is a procedural function, while the lexicon is based on declarative memory. These two memory systems are not just functionally, but also anatomically dissociated: tasks charging procedural memory activate regions in the frontal lobe, the basal ganglia and the cerebellum, while declarative memory is mainly associated with the temporal lobe, the hippocampus, and certain regions in the parietal cortex (Squire et al., 1993). Importantly, while the DP model assumes a dissociation between procedural and declarative memory, and, as a consequence, between the grammar and the lexicon, it does not argue that these language systems are domain-specific and isolated from the rest of cognition within their respective memory systems.

This model has a specific prediction for language impairment formulated as the Procedural Deficit Hypothesis of Specific Language Impairment (Ullman & Pierpont, 2005). On this view, language impairment is a result of abnormal development of brain structures underlying the procedural memory system responsible for learning cognitive and motor skills, and, among them, grammar. Developmental disorders of such a system should result in deficits of skills based on procedural learning within both the linguistic and nonlinguistic domains. Theoretically any part of the network can be impaired, which explains heterogeneity of SLI. As parts of the network can be impaired to a different extent, and since they have neural connections to other areas too, SLI is not necessarily a selective impairment. The PDH also predicts that the better declarative memory abilities are, the less conspicuous SLI is, because of compensatory mechanisms like storing information in chunks and learning explicit rules (for a more recent review, see Ullman & Pullman, 2015). Indeed, in the nonverbal (mostly visual, but also auditory) domain (Baird et al., 2010; Bishop & Hsu, 2015; Dewey & Wall, 1997; Lum et al., 2012; Riccio et al., 2007), studies testing declarative memory have generally revealed largely intact performance.

Several observations are in concert with this hypothesis. As mentioned above, language disorder is often accompanied by attention deficits, motor problems or working memory impairment, which are also procedural functions. Results showing impairments in implicit sequence learning, processing of rapidly changing auditory stimuli and motor skills (especially oral movements involved in speech articulation) also fit with predictions of the PDH. Hand movements are often impaired too, although there are results showing that in some cases they remain intact. The association between language impairment and weak (sequential and nonsequential) motor skills seems to be strong (e.g. Lai et al., 2001; Ullman

& Pierpont, 2005; Ullman, 2016; Vargha-Khadem et al., 1995, 1998). The PDH predicts that the neural structural or functional differences associated with SLI are going to involve the procedural system. Neurobiological research supports this hypothesis: when compared to typical development, the greatest differences in neuroanatomical structures are found in the basal ganglia (especially in the nucleus caudatus) and in the Broca' area of the frontal lobe (Belton et al., 2002; Gauger et al., 1997; Vargha-Khadem et al., 1998). **Studies 3, 8** and **12** in the dissertation are motivated by testing predictions of the PDH.

Aphasia

Conclusions from research on SLI in the dissertation are supported by a further study on executive functions in aphasia. Although this is not the main focus of our research, and a detailed introduction to typology and different profiles, together with a review on aphasia research and its problems is beyond the scope of the dissertation (the reader is referred to e.g. Alexander, 1997; Caplan, 1987; Goodglass, 1993; Goodglass et al., 2001, Bánréti, 2014), we believe a brief definition and motivation is in place. Aphasia is the acquired impairment of language abilities caused by cerebral vascular incident or stroke, traumatic brain injury, brain tumour or other potential organic causes. It involves different modalities of language: depending on the lesion type and severity, and causes problems in the production/and or comprehension of spoken and/or written language, resulting in different patterns of deficits. Besides involving different modalities, aphasia can also target any or several levels of language: phonological, semantic, lexical, syntactic, grammatical, pragmatic linguistic deficits have been described in aphasia.

For a long time, aphasia has been addressed as an impairment of language without much focus on accompanying cognitive deficits in areas like nonverbal memory, spatial orientation, problem solving or perception. The within-language cognitive pattern was described as a mosaic of impaired and intact linguistic subfunctions (e.g. Grodzinsky & Santi, 2008). Although the focus have been on language or subdomains of language, more and more results suggest that nonlinguistic cognitive functions are also impaired in aphasia. Vulnerable areas outside language are processing of environmental sounds and gestures (Saygin et al., 2003a,b), attention, visual memory, construction abilities and executive functions (EF) are also affected in many patients with aphasia (e.g. Bonini et al., 2015; Lee et al., 2014; Novick et al., 2005, 2009, 2010; Thompson-Shill, 2005). In the case of executive functions it has also been argued by the above authors that EF problems are not only accompanying deficits in aphasia; they can also be the cause of at least some of the language symptoms. Together with the general aim of studying the specificity of language, these findings and suggestions were the motivation for testing executive functions in aphasia in one of the studies presented in the dissertation.

Nonlinguistic abilities in language acquisition, processing and production

Skill learning

Focus on skill learning in language impairment, and more generally in language acquisition is motivated by the fact that the process of language acquisition is itself a form of implicit skill learning, since it takes place incidentally, without conscious intent to learn, and without explicit knowledge of the acquired grammatical rules. The acquisition of complex motor, cognitive and social functions like language (or playing a musical instrument or mastering sports) is generally associated with implicit skill learning which relies heavily on statistical learning. Statistical learning has been proposed as a model of language acquisition in child language development, but sometimes other aspects of the same learning process like sequence learning and categorization are highlighted. Different forms of skill learning can play different roles in learning different aspects of language; statistical learning 1) of specific sequences has been argued to be important in word segmentation (e.g. Saffran et al., 1996, 1999) 2) of specific and abstract sequences is suggested to be associated with syntax (e.g. Gomez & Gerken, 1999; Marcus et al., 1999) and 3) nonsequential categorization plays a role in the learning of phoneme categories (e.g. Kuhl, 2000) and of concepts and word meanings (e.g. Yu & Smith, 2007).

Some also argue that among these non-linguistic domain-general abilities playing a role in language acquisition sequence learning has a more prominent role, since most of what we learn in language are complex sequential patterns (e.g. Cornish et al., 2017; Fisher & Scharff, 2009; Conway & Pisoni, 2008; Christiansen & Chater, 2015). Most sequence learning studies belong to the more general domain of implicit, procedural, or statistical learning, but there are also explicit claims about these general forms of learning playing a central role in language acquisition (e.g. Cleeremans et al., 1998; Conway & Christiansen, 2001; Dominey, Hoen, Blanc, & Lelekov-Boissard, 2003; Greenfield, 1991; Gupta & Dell, 1999; Gupta & Cohen, 2002; Ullman, 2004).

Beyond the theoretical proposals about the importance of statistical learning in language acquisition, there are empirical results showing a connection between implicit sequence learning and language abilities. To give a few examples, correlations have been found between implicit learning of visual sequences and speech perception in noise (Conway et al., 2007, 2010), between preschool children's implicit motor sequence learning ability on the SRT task and syntactic priming effects (Kidd, 2012). Such an association and reliance on common background mechanisms is also confirmed by Misyak and Christiansen (2012)'s results of statistical learning scores being good predictors of the individual variation in sentence processing, and by findings of a neural overlap (Christiansen et al., 2012) of processes involved in the processing of linguistic and non-linguistic sequences. The above results suggest that sequence learning shares mechanisms with language processing and language learning, implying that sequence learning may be vulnerable in SLI, and that the impairment of a domain general sequence learning or skill learning mechanism is a good candidate for explaining at least part of the deficits (Conway et al., 2009, 2010).

Many studies explicitly focus on statistical learning as a model of some aspect of language acquisition. Infant studies involving artificial language learning (e.g. Gómez & Gerken, 1999, Saffran et al., 1996) demonstrate that infants at an early age are able to learn different aspects of language from phoneme categories through word boundaries to grammatical categories and rules just relying on frequency distributions and transitional probabilities. These studies show that eight month olds are sensitive to transitional probabilities between items, and are able to segment sequences of syllables and non-linguistic stimuli into "words" as well (Saffran et al., 1996, 1999). Infants as young as 12 month old also demonstrate learning of regularities when exposed to sequences of auditory nonsense syllables strung together according to rules of a finite state grammar (Gomez & Gerken, 2000). Studies of statistical learning or artificial grammar learning also found age-related differences related to the complexity of the grammar even in infants: Gomez and Maye (2005) found that although 15-month-old children were able to learn non-adjacent dependencies, these were too complex for 12-month-olds.

Language-related research on statistical learning processes is closely tied to the debate over language-specific versus domain general mechanism in language acquisition (see e.g. the debate on Marcus et al., 1999, and also Marcus et al., 2007; Saffran et al., 2007). Although statistical learning has been proposed as a domain-general mechanism effective across many domains and modalities and species, some studies suggest that there can be important differences in the effectiveness and other parameters of learning depending on modality and domain (Conway & Christiansen, 2005, 2006; for a review see Frost et al., 2015).

Since statistical learning is a core process in learning complex skills, and many argue that the acquisition of complex skills, and especially of language has a critical period, developmental differences in the effectiveness of statistical learning mechanisms are also relevant to the proposal of a critical period for language learning. This proposal is based on the observation that language (and perhaps other complex skills like playing a musical instrument or mastering sports) is best acquired if learning starts early in childhood, and after a certain age, the effectiveness of learning and the quality of the acquired representations decline. The proposal found empirical support from observations showing that recovery from brain damage is less likely to lead to permanent problems with language if incident happens before the end of the critical period (before teenage years; Lenneberg, 1967, 1974) and from earlier results form second language learning showing gradual decline with age in the ability to acquire language at a native-like level. The existence of a critical period for language acquisition is part of the arguments cited in support of an innate and domain-specific language acquisition mechanisms (Johnson & Newport, 1989; Lenneberg, 1967, 1974).

While the existence of an early critical period in the first year of life in which the quality and quantity of input is crucial is well established (e.g. Shafer et al., 2011; Peña et al., 2012; Werker & Hensch, 2015), critical period proposals for the effectiveness of the language learning mechanisms are a lot more debated. Although critical periods have been linked to biologically pre-programmed maturational processes of specific innate mechanisms, observations of age-related differences in the effectiveness of language learning have been given other explanations too: age-related decline in the overall plasticity of the brain can be part of the explanation (Hiscock & Kinsbourne, 1995). Others argue (partly based on reanalysing earlier results and on meta-analyses of second language acquisition (SLA) end-state studies) that the existence of a critical period implies steep decline in ability, but actual results show a gradual linear decline (Bialystok & Hakuta, 1999; Birdsong & Molis 2001; Elman et al., 1996, Flege et al., 1999,). Some propose that while there are important age-effects in language acquisition, there is no critical period for second language learning, and critical period effects are in fact age-related effects due to entrenchment and competition, and social commitment (e.g. Hernandez & MacWhinney, 2005).

Evidence is also accumulating for native-like representations in late language learners showing similar ERP components and brain activations based on fMRI in processing first and second language, and even in processing an artificial language (e.g., Friederici et al., 2002; Rossi et al. 2006). The quantity and quality of exposure seems to be more important, than the timing: Morgan-short et al. (2012) demonstrated that Adult SLA may rely on the same brain mechanisms as first language acquisition, depending on the method of learning (immersion instead of classes) and on the language proficiency of learners.

Systematic studies of both language acquisition and statistical learning across the lifespan would be important contributions to this debate, but so far, they are very scarce. There is one study that argues for the existence of a dedicated period for skill learning based on results from performance changes across the lifespan (Janacsek, Fiser, & Németh, 2012), and in one of the studies presented in the dissertation (**Study 9**) we will also provide data on age related changes in three different forms of skill learning arguing against an early dedicated period of skill learning.

Executive functions

As executive functions (EF) are essential in coordinating all higher-order behaviour, they are also prominent among candidates for non language-specific cognitive processes that play a significant role in language production and comprehension. Executive functions is an umbrella term for a set of higher order processes that are responsible for the strategic organization of information for behaviour: the generation of new responses (fluency), planning, concurrent storage, updating and manipulation of working memory representations of context-relevant information, inhibiting irrelevant stimuli and inappropriate responses, switching between different tasks, mental sets or actions and resolving conflict between competing pieces of information that are also often referred to as processes of cognitive control[5] (e.g., Anderson, 2002; Baddeley, 1996; Burgess, 2000; d'Esposito et al., 1995; Engle & Kane, 2004; Friedmann et al., 2006; Huizinga et al., 2006;

[5] Both the terms 'executive functions' and 'cognitive control' are going to be used int he papers and the dissertation; the specific tasks and subfunctions will make it clear what functions are addressed.

Miller & Cohen, 2001; Miyake et al., 2000; Norman and Shallice, 1986; Smith & Jonides, 1999).

The role of updating and working memory in various linguistic tasks (e.g. holding the subject of a sentence in the mind until the verb of the sentence appears and we can form the argument structure of the sentence) is well-established (e.g. Caplan & Waters, 1999; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Just & Carpenter, 1992, 1996; Miyake, Just & Carpenter, 1994; Lewis et al., 2006; Németh, 2006; Németh et al., 2014; Waters & Caplan 1996), although there are still many questions to be answered about this relationship. Much less is known about the role of other aspects of cognitive control, such as inhibition, which have not been the focus of intensive research for a long time. Recently it has been suggested that cognitive control abilities responsible for coordinating sensory and motor information to perform goal-directed actions might play a larger role in language functions as well, especially in tasks involving competing representations either in processing, production or in acquisition of words or sentences (Kan & Thompson-Schill, 2004; Schnur et al., 2006; 2009). According to this proposal, selection between competing candidates and inhibition of irrelevant or incorrect representations or responses is managed by cognitive control functions in the language domain just as in other domains. Outstanding examples of conflict are different types of ambiguities in language: reading a homonym in a sentence, interpreting an ambiguous or garden path sentence or understanding irony or metaphor are all situations where representations compete, and the language system has to select among them.

Cases of conflict are a lot more common in language and are not restricted to exceptional cases of ambiguities. As an example, when we retrieve a word several other competing words sharing part of their representations (sound form or meaning) with the target word are also activated to a certain degree. To successfully produce the intended word its activation has to be higher than the activation of the competing words (Kan & Thompson-Schill, 2004; Schnur et al., 2006; 2009; Novick et al., 2005, 2009, 2010; Thompson-Shill, 2005). Cognitive control processes may be recruited in the inhibition of the activation of competing words and facilitating the activation of the target word, but exact details of these mechanisms are yet unknown.

It has also been pointed out that language and cognitive control processes share some of their supporting structures in the prefrontal cortex of the brain, and are associated more closely with ventrolateral prefrontal cortex (VLPFC) and, more specifically, with left inferior frontal gyrus (LIFG; Miller & Cohen, 2001; Novick et al., 2005, 2009, 2010; although see Hsu, Naeggi & Novick, 2017 for a more fine-grained analysis). The LIFG has been shown to be strongly activated during performance of domain-general conflict resolution tasks like in incongruent trials of the Stroop task, Flanker test (e.g. Milham, Banich, & Barad, 2003) and the n-back task on one hand, and on the other, it has also been associated (especially Broca's area which is part of LIFG) with several language-specific functions such as articulation and syntax, and verbal working memory (e.g. Caplan & Waters, 1999; Friederici, 2002; Goodglass, 1993; Grodzinsky, 2000; Grodzinsky & Santi, 2008).

The traditional role of Broca's area as a language-specific area has been continuously challenged by both neuropsychological and imaging findings. Many patients with lesions in this area do not show symptoms or deficits in functions associated with Broca's area, while patients with lesions in other areas can also show them. Imaging results also draw a diverse picture: Broca's area is not always activated in tasks involving functions proposed to be associated with it (e.g. Barde & Thompson-Schill, 2002; d'Esposito & Postle, 1999; Friederici, 2002, Kaan & Swaab 2002). These diverse and controversial findings are resolved in Novick et al. (2005)'s suggestions that LIFG has a unique and non-language specific role in resolving conflict among competing representations, and it will only be involved in language processes when representations compete, and conflict resolution is called for.

Against this background, and motivated by controversial sets of findings on EFs in SLI and aphasia (see detailed reviews in **Study 7** and **Study 10**), we examined in two studies domain-general abilities of cognitive control to see whether impairments in cognitive control are associated with language disorder in a developmental (SLI) and in an acquired (aphasia) form of the impairment. Besides testing whether impairments in language and cognitive control tend to co-occur, we also wanted to examine the contribution of cognitive control language production in **Study 11**.

Methodological issues

In most of our studies, we used well-established and widely used paradigms of experimental psychology and language acquisition adapted to our research questions. In testing cognitive functions outside language, we relied on adaptations of well-known tasks. Skill learning in both SLI and in TD was tested by 3 paradigms (SRT, PCL and AGL) that have been used by many authors to test implicit/procedural/skill learning, and Stroop tasks, n-back and fluency tasks or the stop-signal task employed in our studies on executive functions in both SLI and aphasia are also frequently used in the literature.

Although SLI has been part of language research in cognitive psychology for more than 25 years, at the time of launching our research, there has been very few systematic and theory-driven experimental studies on language impairment in Hungarian. We have developed a number of language tests which allow the detailed mapping of linguistic problems of Hungarian-speaking children with SLI on the one hand, and at the same time, allow us to test competing accounts of SLI. In the selection of the phenomena to be tested and in the construction of the tests we were motivated by three more general aspects: (1) data obtained from the linguistic phenomenon selected for testing should be potentially able to support or refute a specific theory of specific language impairment; for this reason, areas of testing include agreement, regular-irregular inflection, or examination of procedural and declarative learning (2) we wanted to examine areas that are specific to Hungarian, or at least that are not found in languages that have been extensively tested before; these areas include verb-object agreement, or the rich system of case markers appearing as part of the verb's argument structure belong here (3) we also wanted to provide suggestions or areas of testing for practitioners that allow screening and detailed examination of different subtypes of developmental disorders of language.

In the case of testing different areas of language, most of our methods were widely used procedures (picture naming, sentence-picture verification, sentence repetition, grammaticality judgments, etc.), but it was always necessary to develop a new task with Hungarian stimuli and a design adapted to language-specific features. Part of the reason was that Hungarian tests were not available for most linguistic phenomena, but it was also necessary because many of the structures we tested were specific to Hungarian, so simple translations of existing test for English would not have worked. Together with using familiar task designs and methods with carefully compiled Hungarian stimuli, we also used a novel task which we believe is an innovative tool of elicited production in both children and adults. Because this procedure has not been used for purposes of testing productive use of grammatical morphemes in the literature, we will introduce it in more detail below.

Sentence repetition with masked inflections: a new method of elicited production

A generally known major drawback of elicited production studies relying on pictures/and or targeted questions is that it is difficult to create a context in which it is impossible to give an acceptable answer without using the target structure, and children can be extremely creative in avoiding structures the experimenter would like to hear. The need for a new method of elicited production was urged by our lack of success with traditional ways of eliciting some of the suffixes in a study on agreement in Hungarian. Of the 24 suffixes we wanted to elicit (Person (3) x Number (2) x definiteness (2) x Tense (2)), it was very difficult, if not impossible, to create a pragmatically appropriate context using pictures and descriptions for some (like the second person plural forms).

To get around this problem, we devised a new method of eliciting suffixes and testing morphological productivity which gets around the above difficulties in a procedure in which participants are asked to repeat digitized sentences heard through a loudspeaker. Critically, the target inflections in each sentence are masked by a cough inserted to prevent hearing the inflection but not the stem or the remaining portions of the sentence. The method was adapted from Warren's (1970) phoneme restoration procedure; the restoration effect has been demonstrated at the morpheme level as well for affixes in Hungarian (Dankovics & Pléh, 2001), but the effect has not yet been exploited in developmental studies as an elicited production technique.

Examples from our study on agreement in SLI (**Study 1**) are given in (a) and (b); the *XXX* marks a cough. Although the target inflections in the middle of the sentence were replaced by a cough, the remainder of the sentence contained all the source features for unequivocal identification of the missing inflection. Pre-testing also ensured that there were no anticipatory coarticulation cues left in the sentence: adults guessed correctly on 5.6% of the items (having to select from the 24 possible forms). Children occasionally commented that the speaker was coughing, but they never noticed that something was missing from the sentence.

(a) *Mi olvas*XXX *egy mesét*

Target: *olvas**unk*** ["read" 1PlPresIndef]

"We are reading a story"

(b) *Tegnap ti tol*XXX *a biciklit*

Target: *tol**tátok*** [“push” 2PlPastDef]

“Yesterday you pushed the bike”

This method was successfully used in testing knowledge of agreement morphemes and case-marking in Hungarian-speaking children with SLI and typically developing control children between 4 and 12 years (Lukács et al. 2009; Lukács et al., 2013). In combination with the novel word paradigm, it also proved useful in investigating early productivity of morpheme use in Hungarian children aged between 2;1 and 5;3, where results suggested that Hungarian nominal and verbal suffixes can be used productively before the age of three, highlighting the viability of the method of sentence repetition with masked inflections in studying young children’s morphological productivity (Gábor and Lukács, 2012). Using this technique has several methodological advantages: 1) the fully audible portions of the sentence make unequivocally clear which inflection is the appropriate one to use, 2) it tests productivity implicitly, disguised in the form of sentence repetition, making experimental investigation of morphological productivity possible in young children as well 3) it can be used to test knowledge of morphemes that are otherwise (by using pictures or prompts) difficult or impossible to elicit. The low computational load of the task reduces the biasing effect of attentional, memory or other performance demands of the experiment, and makes this method a real alternative to traditional elicitation procedures used in child language research, which can be especially useful in testing productivity in morphologically rich languages containing large paradigms with many suffixes.

Participants in the SLI studies

Participants in the large scale developmental study were typically developing children and adults; they are introduced in detail in the paper. Since the dissertation contains only one study on patients with aphasia, we also refer the reader to the paper on EF in aphasia for a detailed description. In the majority of our studies, participants were children with SLI and typically developing children, who formed the control groups. Although full descriptions of participants with SLI and their TD controls are also given in the individual papers, a general introduction to the selection and matching procedures is perhaps useful.

Children with SLI were recruited from special preschools and schools. They 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. All children met inclusive and exclusive criteria for SLI that are standardly used in selecting SLI children in research (see e.g. Dollaghan, 2007; Leonard 1998/2014, Tager-Flusberg and Cooper, 1999). Each child scored above 85 on the Raven Coloured Progressive Matrices (Raven, Court, & Raven, 1987), a measure of nonverbal intelligence. No child had a hearing impairment or a history of neurological impairment.

Language impairment is often associated with other developmental disorders, most frequently with ADHD, dyslexia and ASD (for a review, see Leonard, 1998/2014). No children in our SLI group had any known comorbidities. Comorbidities were identified and excluded on the basis of detailed case history in the anamneses by the speech and language therapists at the schools of children with SLI and based on teachers' report in the case of TD children.

Each child scored at least 1.25 SDs (in some studies, 1.5 SDs) below age norms on at least two of four language tests administered. The four tests included two receptive tests: the Hungarian version of the Peabody Picture Vocabulary Test (PPTV, Dunn & Dunn, 1981; Csányi, 1974) and the Test for Reception of Grammar (TROG, Bishop, 1983/2012, Lukács, Győri, Rózsa, 2012) and two expressive tests: the Hungarian Sentence Repetition Test (Magyar Mondatutánmondási Teszt, MAMUT, Kas & Lukács, in preparation), and a nonword repetition test (Racsmány, Lukács, Németh, & Pléh, 2005). Other than English, very few languages have properly standardized language tests for screening and diagnosing language impairment; unfortunately Hungarian is not yet among them. The Hungarian version of the TROG has Hungarian standards, and the nonword repetition test also has age norms. As for the other 2 tests, although we do have data from large samples of children from different age ranges, there are no official standard scores for the Hungarian version of the PPVT or the sentence repetition test.

Children in the control groups were typically developing children (TD) matched on i) receptive vocabulary level (PPVT raw scores) in our language studies ii) on chronological age (each child in the TD group was within 3 month 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) in our studies on non-linguistic abilities in SLI. TD children were recruited from schools and preschools with no special selection processes for children. All children were tested with the informed consent of their parents.

Questions and aims of the studies

In what follows, we first review and motivate questions and methods that address grammatical functions, that we turn to studies focusing on nonlinguistic cognitive functions in SLI, and the larger-scale developmental study of skill learning abilities in TD, which is followed by the brief description of the study on EFs in aphasia.

Grammatical functions in SLI

During our research, we first aimed at a detailed assessment of grammatical functions in SLI in several studies, together with examining their relationships with vocabulary size and with different indices of processing capacity. We were curious to see if we find evidence of selective grammatical deficits in any area of grammar.

We formulated different predictions based on grammar-specific versus processing accounts of SLI introduced above. Based on predictions of grammatical theories of SLI we expected to see

a) a selective impairment in grammar, together with an intact lexicon

b) a selective impairment in one of the following areas of grammar: agreement, tense/aspect marking, case marking

c) no impairments in non-linguistic cognitive functions.

Based on processing accounts, we expected that

a) problems in processing will affect grammar and lexicon alike, we will see deficits in both areas

b) no selective impairments occur in any specific area or representation of grammar; vulnerable areas will be defined by factors affecting processing difficulty

c) performance on grammatical tasks will be associated with indices of processing capacity

d) impairments will occur in cognitive functions outside the language domain as well.

As discussed above, children with SLI often show weaknesses in grammatical morphology and Hungarian offers new opportunities for testing contrasting accounts of morphological difficulties in SLI together with offering new phenomena to examine. It is a non-configurational language with a very rich system of inflections. Since word order is relatively free, inflections convey the core grammatical information and they play a more important role in grammar than they do in configurational languages like English. Inflections can combine, and many combinations are possible, but the order in which inflections appear at the end of a word is fixed. There are also other complex morphophonological patterns governing inflection use. The verb inflection system contains both agglutinating and fusional elements, and has agreement between the verb and object as well as between the verb and subject. Noun morphology also offers new avenues of testing: the case marking system is very complex, above nominative and accusative cases, there are 16 more case marking suffixes. These characteristics make this language a good test case for several accounts of the grammatical deficits of children with language impairment (SLI).

As problems with finite verb morphology (agreement and tense marking) are mentioned among the most problematic areas and have been suggests to be targets of a selective impairment (e.g. Clahsen, 1991; Rice et al., 1995), we studied grammatical morphology of subject-verb and verb-object agreement phenomena in 25 Hungarian children with SLI and 25 control children matched individually on receptive vocabulary scores (VC). We devised a structured method of eliciting responses in the guise of a sentence repetition task. Target inflections in each (digitized) sentence were masked by a cough that prevented the child from hearing the inflection but not the stem or the remaining portions of the sentence which contained all the source features for unequivocal identification of the missing inflection. The verb inflections marked distinctions according to tense, person, number, and definiteness of the object (**Study 1**). We also asked for grammaticality judgments for sentences containing errors along these dimensions of agreement and tense (**Study 4**).

We examined the comprehension and production of aspect as well as tense with perfect and imperfect verb forms in children with SLI. SLI performance was compared to performance of two groups: a group of typically developing children matched for age, and another TD group matched for receptive vocabulary scores. Earlier studies of English-speaking children with SLI showed that impairment in sensitivity to aspect can also disrupt the acquisition of tense in this group, and is a problem in itself in SLI (e.g. Leonard and Deevy, 2010). Here again, Hungarian provides a unique opportunity for testing tense and aspect independently of each other (**Study 5**).

We also examined the use of agglutinating noun inflections by Hungarian-speaking children with SLI with multiple suffixes and with morphophonologically regular and irregular stem classes. We focused on two noun inflections – plural and accusative case singly and in combination. When these two inflection types appear together, the plural precedes the accusative. For example, the bare stem for ’room’ is *szoba*, the plural form ’rooms’ is *szobák*, the accusative form is *szobát*, and the plural + accusative form is *szobákat*. Hungarian nouns can also be classified as regular or irregular as a function of the regularity of the phonological changes that the stem must undergo when an inflection is added. The great majority of stems belong to productive regular classes, and the minority of irregulars form closed classes with small type frequencies. Importantly, however, the inflection is invariable and identifiable with all stem types, regular and irregular (for example, the plural form of every noun, whether regular or irregular, ends in *–k*). Taken together, these factors allow examination of grammatical marking (of plural and accusative) and the extraction of morphophonological patterns at the same time. We tested 5-7-year-old and 8-10-year-old children with SLI and two control groups matched on vocabulary size (VC older: 4.5-8 years, VC younger 3.25-7 years) on an elicited production task using pictures. Altogether 60 children participated. Response accuracy was analyzed both from a morphosyntactic (examining whether they marked the plural and the accusative correctly) and a morphophonological point of view (examining whether they used the correct allomorph) (**Study 2**).

To further investigate productivity with noun morphology, we also tested production of case markers in spontaneous speech samples and in a sentence repetition task with masked inflections to see whether case marking poses a selective problem (as suggested with grammar in a language with rich morphology. To test predictions of processing accounts, we also tested if there is a difference between the difficulty of case markers depending on whether they appear in the sentence in their (a) regular spatial meaning or in their (b) lexically specified nontransparent nonspatial meaning.

a) *Az oroszlán megszökött a ketrec**ből.***
The lion escaped the cage-FROM.
‘The lion escaped from the cage’.

b) *Pisti tanult a baleset**ből***
Pisti learnt the accident-FROM.

‘Pisti learnt from the accident’.

Just like in the noun morphology task, we tested SLI children in two age groups: the older group had 29 children with SLI between 7;11 and 11;4 and the younger group had 17 children with SLI between 4;10-7;2. Performance was compared to two verbal control (VC) groups of typically developing children matched on receptive vocabulary scores (**Study 6**).

Nonlinguistic abilities in SLI, TD and aphasia

It is clear from the literature that despite the fact that the definition of SLI excludes major nonlinguistic impairments, specific language impairment cannot be regarded as a purely linguistic deficit. Motivated by these findings, we wanted to study several nonlinguistic abilities that are good candidates for being involved in the process of language acquisition, but their relationship with language abilities is not clear. Besides testing general learning abilities contributing to language acquisition, our aim was also to test the procedural deficit hypothesis of SLI (Ullman & Pierpont, 2005). As the procedural system is responsible for learning not only motor skills, but also for other cognitive skills like probabilistic category learning or rule learning, and SLI children are claimed to have a more general procedural deficit, we expected that they will show impaired performance in several linguistic and non-linguistic tasks that rely on the functioning of the procedural system. To test not just the presence, but also the selectivity of the procedural deficit, declarative functions were also tested in both the linguistic and non-linguistic domains.

Language specificity of impairments in SLI was first tested by examining general skill learning functions. We tested three different forms of skill learning in 29 children with SLI and age matched TD children using three paradigms: 1) the Serial Reaction Time Task (learning of motor sequences), 2) Artificial Grammar Learning (the extraction of regularities from auditory sequences) 3) and Probabilistic Category Learning in the Weather prediction task (a non-sequential categorization task) (**Study 3.** and **Study 8**).

To test the prediction of the PDH concerning the intactness of declarative functions in SLI, we examined learning and overnight retention in declarative memory in Hungarian children with (n = 21) and without (n = 21) SLI. Since we wanted to minimize the influence of working memory and free recall (functions that in themselves tend to be impaired in SLI) we tested declarative memory with a recognition memory task, following incidental encoding.

To test initial learning we examined recognition memory 10 minutes after encoding. Both nonverbal items (pictures of real and novel objects) and verbal items (auditorily presented real and novel words) were examined. To test for retention and the effects of consolidation we examined recognition memory of the same items 24 hours later (**Study 12**).

As reviewed above, numerous results show that SLI is often associated with impairments in working memory (WM) and in executive functions (EF), but the nature, extent and generality of these deficits is yet unclear, as is their relationship with language abilities. With this background, we used linguistic and non-linguistic tasks examining WM and EF in Hungarian-speaking children with SLI and their age- and nonverbal IQ-matched typically developing peers (TD). We wanted to examine whether verbal and non-verbal executive and WM functions are similarly affected in SLI with verbal and nonverbal versions of the following tasks: simple and complex span (digit span, Corsi-blocks, listening span, odd-ball-task), n-back tasks (with letters and with images of fractals that are difficult to verbalize) fluency (category and design fluency) and Stroop tasks (verbal and nonverbal) (**Study 10**).

As there is also a growing body of research demonstrating that executive functions, more specifically control abilities play an important role in production and comprehension at both the word and the sentence level where representations compete (also see above), we were also interested to see whether these potential impairments in cognitive control (WM and EF) contribute to language deficits or are only associated with them. We measured naming latencies in a naming task (based on Schnur et al., 2006, 2009) in which the level of conflict (and so the involvement of cognitive control) was manipulated. Participants had to name pictures that had more vs. just one possible name (low versus high naming agreement) and that appeared in semantically homogeneous vs. mixed blocks. In both manipulations, the first condition is the one that results in greater conflict charging cognitive control abilities more. For items with low naming agreement, available alternative names have to be inhibited, and in homogenous blocks, names from similar competing category members have to be controlled. We tested the influence of such manipulations in typically developing children and in children with Specific Language Impairment (SLI) (**Study 11**).

This detailed study of linguistic and nonlinguistic functions in SLI was complemented by two further studies that approach the question of specificity in language from different angles. We examined general skill learning abilities that can be associated with the acquisition of different aspects of language in typical development. To test whether these skill learning abilities that potentially contribute to language learning show critical period effects, we tracked age-related changes in different forms of implicit learning between 7 and 80 years using 3 paradigms: 1) the Serial Reaction Time Task (learning of motor sequences), 2) Artificial Grammar Learning (the extraction of regularities from auditory sequences) 3) and Probabilistic Category Learning in the Weather prediction task (a non-sequential categorization task). Although these three tasks are standard paradigms of skill learning that have been extensively used in typical populations and in neuropsychological studies, systematic developmental tracking of performance throughout the lifespan are missing (**Study 9**).

We also examined cognitive control abilities in aphasia, the acquired impairment of language abilities. Aphasia has been defined as the impairment of language specifically, but there is an increasing set of results showing nonlinguistic impairments as well (see above). In the case of executive functions it has also been argued by the above authors that EF problems are not only accompanying deficits in aphasia, the can also be the cause of at least some of the language symptoms. We tested five individuals with transcortical motor aphasia (TMA), five patients with conduction aphasia and ten healthy controls participated on four EF tasks. A visual and an auditory *n*-back task was used to assess updating of working memory representations. Effectiveness of inhibition in the resolution of response conflict was tested with a Stop-signal task, and resolution of representational conflict was examined with a nonverbal Stroop task (**Study 7**).

The research presented in the dissertation hopes to offer a unique contribution to the question of the specificity of language by profiling several areas of language and language-relevant non-linguistic functions together along with language abilities. By studying different areas within grammar, we can test whether there are linguistic symptoms that are good candidates for language-specific clinical markers of SLI in Hungarian. By testing non-linguistic functions in SLI and in aphasia, we have the opportunity to examine how non-linguistic mechanisms back up grammar and vocabulary acquisition, and to test whether it is necessary to postulate specific mechanisms for language acquisition and processing. It can also provide important data for testing the Procedural Deficit Hypothesis of SLI (PDH, Ullman and Pierpont, 2005) by showing whether the deficit selectively impairs the procedural system, or it also involves mechanisms of declarative memory. The study on age-related changes in different forms of skill learning promises to determine whether developmental changes show a critical period effect for these non-linguistic learning mechanisms that might be important in language acquisition.

Empirical studies presented in the dissertation in support of the 11 thesis points

1. Lukács, Á., Leonard, L. B., Kas, B. & Csaba Pléh (2009) The Use of Tense and Agreement by Hungarian-Speaking Children with Language Impairment. *Journal of Speech, Language and Hearing Research*. 52/1, 1-22. doi: 10.1044/1092-4388(2008/07-0183)
2. Lukács, Á., Leonard, L. B., & Kas, B. (2010). The Use of Noun Morphology by Children with Language Impairment: The Case of Hungarian. *International Journal of Language & Communication Disorders*, 45, 145-161. doi: 10.3109/13682820902781060.
3. Kemény F., & Lukács Á. (2010). Impaired procedural learning in language impairment: results from probabilistic categorization *Journal of Clinical and Experimental Neuropsychology,* 32:3, 249–258. doi: 10.1080/13803390902971131
4. Lukács, Á., Leonard, L. B., & Kas, B. (2011). "The dog chase the cat": grammaticality judgments by Hungarian-speaking children with language impairment. *Acta Linguistica Hungarica,* 58:1–2, 24–38. doi : 10.1556/ALing.58.2011.1–2.2
5. Leonard, L. B, Lukács, Á., & Kas, B. (2012). Tense and Aspect in Childhood Language Impairment: Contributions from Hungarian. *Applied Psycholinguistics, 33,* 305–328. doi: 10.1017/S0142716411000361
6. Lukács, Á., Kas, B. & Leonard, L. B. (2013). Case marking in Hungarian children with specific language impairment. *First Language*, 33/4. 331-353. doi: 10.1177/0142723713490601
7. Zakariás, L., Keresztes, A., Demeter, Gy. & Lukács, Á. (2013). A specific pattern of executive dysfunctions in transcortical motor aphasia. *Aphasiology*, 27:12, 1426-1439, doi: 10.1080/02687038.2013.835783
8. Lukács Á., & Kemény F., (2014). Domain general sequence learning deficit in Specific Language Impairment. *Neuropsychology*, 28(3), 472-483. doi: 10.1037/neu0000052
9. Lukács Á., & Kemény F., (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science*, 39(2) 383-404. doi: 10.1111/cogs.12143
10. Lukács Á., Ladányi, E., Fazekas, K. & Kemény F. (2016) Executive Functions and the Contribution of Short-Term Memory Span in Children With Specific Language Impairment. *Neuropsychology*, 30(3):296-303. doi: 10.1037/neu0000232
11. Ladányi, E. & Lukács Á. (2016). Lexical Conflict Resolution in Children with Specific Language Impairment. *Journal of Communication Disorders*, 61, 119-130. doi: 10.1016/j.jcomdis.2016.04.004
12. Lukács Á., Kemény F., Lum, J.A.G. & Ullman, M.T. (2017). Learning and overnight retention in declarative memory in specific language impairment. *PLoS ONE* 12(1): e0169474. doi: 10.1371/journal.pone.0169474

Thesis points and empirical papers supporting them

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6.)

1. Lukács, Á., Leonard, L. B., Kas, B. & Csaba Pléh (2009). The Use of Tense and Agreement by Hungarian-Speaking Children with Language Impairment. *Journal of Speech, Language and Hearing Research*. 52/1, 1-22. doi: 10.1044/1092-4388(2008/07-0183)
2. Lukács, Á., Leonard, L. B., & Kas, B. (2010). The Use of Noun Morphology by Children with Language Impairment: The Case of Hungarian. *International Journal of Language & Communication Disorders*, 45, 145-161. doi: 10.3109/13682820902781060.
4. Lukács, Á., Leonard, L. B., & Kas, B. (2011). "The dog chase the cat": grammaticality judgments by Hungarian-speaking children with language impairment. *Acta Linguistica Hungarica, 58:1–2, 24–38.* doi: 10.1556/ALing.58.2011.1–2.2
5. Leonard, L. B, Lukács, Á., & Kas, B. (*2012*). Tense and Aspect in Childhood Language Impairment: Contributions from Hungarian. *Applied Psycholinguistics, 33,* 305–328. doi:10.1017/S0142716411000361
6. Lukács, Á., Kas, B. & Leonard, L. B. (2013). Case marking in Hungarian children with specific language impairment. *First Language*, 33/4. 331-353. doi: 10.1177/0142723713490601

Thesis 2. Agreement deficits in SLI are better explained by processing difficulties than by a selective grammatical impairment targeting agreement. (1, 4)

1. Lukács, Á., Leonard, L. B., Kas, B. & Csaba Pléh (2009) The Use of Tense and Agreement by Hungarian-Speaking Children with Language Impairment. *Journal of Speech, Language and Hearing Research*. 52/1, 1-22.
4. Lukács, Á., Leonard, L. B., Kas, B. (2011) "The dog chase the cat": grammaticality judgments by Hungarian-speaking children with language impairment. *Acta Linguistica Hungarica,* 58:1–2, 24–38. doi: 10.1556/ALing.58.2011.1–2.2

Thesis 3. Difficulties with multiple suffixation and with morphophonologically irregular forms suggest lexical and processing problems instead of a grammar-specific deficit in SLI (2)

2. Lukács, Á., Leonard, L. B., Kas, B. (2010) The Use of Noun Morphology by Children with Language Impairment: The Case of Hungarian. *International Journal of Language & Communication Disorders*, 45, 145-161. doi: 10.3109/13682820902781060.

Thesis 4. Difficulties in aspect marking in production but not in comprehension in past tense forms suggest a processing problem instead of a selective impairment of aspect marking (5)

5. Leonard, L. B, Lukács, Á., Kas, B. (*2012*) Tense and Aspect in Childhood Language Impairment: Contributions from Hungarian. *Applied Psycholinguistics, 33,* 305–328. doi:10.1017/S0142716411000361

Thesis 5. Problems with case marking in SLI suggest lexical and processing deficits instead of a selective case marking impairment within grammar (6)

6. Lukács, Á., Kas, B. & Leonard, L. B. (2013). Case marking in Hungarian children with specific language impairment. *First Language*, 33/4. 331-353. doi: 10.1177/0142723713490601

Thesis 6. Cognitive impairments in SLI do not selectively target language; deficits also occur in skill learning outside the language domain, most prominently for sequentially organized stimuli (3, 8.)

3. Kemény F., Lukács Á. (2010) Impaired procedural learning in language impairment: results from probabilistic categorization *Journal of Clinical and Experimental Neuropsychology.* 32:3, 249–258. doi: 10.1080/13803390902971131
8. Lukács Á., Kemény F., (2014). Domain general sequence learning deficit in Specific Language Impairment. *Neuropsychology*, 28(3), 472-483. doi: 10.1037/neu0000052

Thesis 7. In concert with the Procedural Deficit Hypothesis, procedural learning is vulnerable in SLI, while processes of declarative learning and retention are relatively intact (3, 8, 12).

3. Kemény F., Lukács Á. (2010) Impaired procedural learning in language impairment: results from probabilistic categorization *Journal of Clinical and Experimental Neuropsychology.* 32:3, 249–258. doi: 10.1080/13803390902971131
8. Lukács Á., Kemény F., (2014). Domain general sequence learning deficit in Specific Language Impairment. *Neuropsychology*, 28(3), 472-483. http://dx.doi.org/10.1037/neu0000052
12. Lukács Á., Kemény F., Lum, J.A.G.& Ullman, M.T. (2017). Learning and overnight retention in declarative memory in specific language impairment. *PLoS ONE* 12(1): e0169474. doi:10.1371/journal.pone.0169474

Thesis 8. SLI difficulties in executive functions are mainly present on verbal versions of EF tasks, and are eliminated by controlling for verbal short term memory span. (10).

10. Lukács Á., Ladányi, E., Fazekas, K. & Kemény F. (2016) Executive Functions and the Contribution of Short-Term Memory Span in Children With Specific Language Impairment. *Neuropsychology*, 30(3):296-303. doi: 10.1037/neu0000232

Thesis 9. Lexical inhibition is effective in SLI (11).

11. Ladányi, E. & Lukács Á. (2016). Lexical Conflict Resolution in Children with Specific Language Impairment. *Journal of Communication Disorders*, 61, 119-130. doi: 10.1016/j.jcomdis.2016.04.004

Thesis 10. Age-related changes in different forms of skill learning with potential roles in language acquisition argue against the existence of a critical period for these learning mechamisms (9).

9. Lukács Á., Kemény F., (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science*, 39(2) 383-404. doi: 10.1111/cogs.12143

Thesis 11. In aphasia, the acquired language impairment is not specific to language: it is often accompanied by the impairment of nonverbal executive functions (7).

7. Zakariás, L., Keresztes, A., Demeter, Gy. & Lukács, Á. (2013). A specific pattern of executive dysfunctions in transcortical motor aphasia. *Aphasiology*, 27:12, 1426-1439, doi: 10.1080/02687038.2013.835783

Study 1.

Lukács, Á., Leonard, L. B., Kas, B. and Csaba Pléh (2009) The Use of Tense and Agreement by Hungarian-Speaking Children with Language Impairment. *Journal of Speech, Language and Hearing Research*. 52/1, 1-22. doi: 10.1044/1092-4388(2008/07-0183)

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6)

Thesis 2. Agreement deficits in SLI are better explained by processing difficulties than by a selective grammatical impairment targeting agreement. (1, 4)

dc_1365_16

The Use of Tense and Agreement by Hungarian-Speaking Children With Language Impairment

Ágnes Lukács
HAS-BME Cognitive Science Research Group, Budapest, Hungary, and Research Institute of Linguistics, Hungarian Academy of Sciences, Budapest

Laurence B. Leonard
Purdue University, West Lafayette, IN

Bence Kas
Budapest University of Technology and Economics, Hungary, and Eötvös Loránd University of Sciences, Budapest, Hungary

Csaba Pléh
Budapest University of Technology, Hungary

Purpose: Hungarian is a null-subject language with both agglutinating and fusional elements in its verb inflection system, and agreement between the verb and object as well as between the verb and subject. These characteristics make this language a good test case for alternative accounts of the grammatical deficits of children with language impairment (LI).
Method: Twenty-five children with LI and 25 younger children serving as vocabulary controls (VC) repeated sentences whose verb inflections were masked by a cough. The verb inflections marked distinctions according to tense, person, number, and definiteness of the object.
Results: The children with LI were significantly less accurate than the VC children but generally showed the same performance profile across the inflection types. For both groups of children, the frequency of occurrence of the inflection in the language was a significant predictor of accuracy level. The two groups of children were also similar in their pattern of errors. Inflections produced in place of the correct inflection usually differed from the correct form on a single dimension (e.g., tense or definiteness), though no single dimension was consistently problematic.
Conclusions: Accounts that assume problems specific to agreement do not provide an explanation for the observed pattern of findings. The findings are generally compatible with accounts that assume processing limitations in children with LI, such as the morphological richness account. One nonmorphosyntactic factor (the retention of sequences of sounds) appeared to be functionally related to inflection accuracy and may prove to be important in a language with numerous inflections such as Hungarian.

KEY WORDS: Hungarian, language impairment, morphosyntax, language disorders

Children with language impairment (LI) show significant deficits in language ability without accompanying deficits such as hearing impairment, neurological damage, or mental retardation. Although children with LI represent a heterogeneous population, common profiles can be identified. In English, for example, a very common profile is a mild to moderate deficit in lexical skills and a more serious deficit in morphosyntax. Within the area of morphosyntax, the use of tense and agreement morphemes seems to be especially problematic.

One complicating factor in the study of LI is that a common profile in one language is uncommon or even absent in another language. For example, word order errors are common in Swedish and German but not in English. In Italian, verb inflections that express agreement with the subject are not among the areas of special difficulty, unlike the case for English.

Proposals for these cross-linguistic differences are beginning to emerge in the literature. Following a brief review of these proposals, we will describe a study employing Hungarian, a language that represents an excellent test case for the suitability of these alternative proposals. Hungarian differs from other languages studied by LI researchers in key respects. One characteristic is the agglutinating morphology with respect to tense and agreement, where an inflection marking tense is followed by an inflection marking agreement, both attached to the verb stem. A second important characteristic of Hungarian is the fact that verb inflections agree with both the subject (in person and number) and the object (in definiteness). As will be seen below, these characteristics have implications for current accounts of the morphosyntactic difficulties seen in LI.

Recent Accounts of Morphosyntactic Deficits in LI

Morphological Richness

The morphological richness account has evolved from the findings of Leonard and his colleagues (Leonard, 1998, pp. 255–257; Leonard, Sabbadini, Leonard, & Volterra, 1987; Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999). According to this account, extraordinary difficulties with tense and agreement morphemes are the result of an interaction between a more general limitation in language ability and the properties of the particular system of grammar that must be learned. Key details of the morphological richness account were inspired by the competition model (e.g., Bates & MacWhinney, 1989; MacWhinney, 1987), such as the views that languages differ in the details of grammar that have the greatest cue validity, that the discovery and use of these cues are probabilistic in nature, and that some cues have greater processing cost than others.

An important assumption of the morphological richness account is that children with LI have a limited processing capacity. For languages such as English, this limitation can be problematic for the learning of grammatical morphology. Inflections are sparse in English, and bare stems are frequent. Faced with a limited processing capacity, then, children with LI might devote their limited resources to the more prevalent information conveyed by word order. Fewer resources would remain for the learning of grammatical morphology, requiring more encounters with grammatical morphemes before they can be learned. In contrast, children with LI acquiring languages with a rich inflectional morphology are expected to devote their limited resources to this area of the grammar. Thus, differences in the use of grammatical morphology between these children and their typically developing peers will be smaller than in a language such as English. It is for this reason that the account gets its name—*morphological richness.*

However, if the inflections themselves reflect a complex combination of grammatical dimensions (e.g., tense, number, person, gender), problems can arise even in the area of inflections in a language with a rich morphology. The more dimensions children must consider simultaneously, the greater the demands on their limited processing capacity. These demands can result in incomplete processing, requiring more encounters with the inflection before it can become a stable part of the children's grammar. Based on findings from Italian and Hebrew, Leonard (1998) proposed that children with LI may approach their processing limitations when four dimensions must be considered simultaneously. According to Leonard, incompletely processed inflections are the functional equivalent of inflections with low frequency of occurrence because they are not registered consistently and therefore do not achieve sufficient strength in the child's grammar to be retrieved as reliably as can be accomplished by typically developing children. Given that children with LI must have a greater number of encounters with each inflection before it is sufficiently established to be retrieved for production with facility, the frequency of occurrence of the inflection in the input is an important factor in the morphological richness account. It is predicted that accuracy will be greater for inflections that are encountered more frequently in the input.

The morphological richness account's focus on the number of dimensions in an inflection system differs from an approach such as the competition model in that the latter places an emphasis on cue validity. Thus, an inflection that reflects a complex combination of four dimensions would be expected to be challenging for children with LI according to the morphological richness account, but if that inflection has high cue validity, the number of dimensions would play a much smaller role according to the competition model.

Another assumption of the morphological richness account is that if errors occur, the substitute inflection is expected to share most features with the inflection that it replaces. In many instances, this will be a "near-miss" error—an inflection that possesses most but not all features reflected in the correct form (e.g., Bedore & Leonard, 2001; Dromi et al., 1999). For example, a third person plural form in the past might be replaced by a third person plural form in the present or a third person singular form in the past. Children with LI are not expected to resort to a default form. Furthermore, if an inflection used as a substitute is found to differ from the correct inflection on multiple dimensions (e.g., tense, person, and number), the substitute should prove to have high frequency of occurrence in the language (leading to greater strength in the paradigm). Only forms of high frequency should serve as competitors to inflections that constitute near misses, as retrieval is assumed to be driven initially by shared features and only highly frequently occurring forms should have enough strength to alter the retrieval process. The morphological richness account grants no special status to any given dimension. Thus, if the correct inflection is not retrieved, the substitute should differ only minimally from the correct form, but no single dimension will dominate. Thus, although all dimensions are operative, they are not hierarchically arranged.

Agreement Deficit

Clahsen and his colleagues (Clahsen, Bartke, & Göllner, 1997; Clahsen & Dalalakis, 1999; Clahsen & Hansen, 1997; Eisenbeiss, Bartke, & Clahsen, 2005) have proposed that children with LI have a selective syntactic deficit that affects agreement in particular. These investigators adopted Chomsky's (1995) distinction between interpretable and noninterpretable features and posited that in LI, the verb's noninterpretable features are not properly acquired. Even in null-subject languages, subject–verb agreement is posited to be problematic (Clahsen & Dalalakis, 1999). Errors are expected to be productions of default forms, such as the production of a present third person singular inflection in contexts that obligate a different inflection. The agreement deficit account does not predict difficulties with tense.

Nonmorphosyntactic Language Processing Factors

The morphological richness account is concerned with processing limitations within the scope of morphosyntactic learning and use. This emphasis is well placed, of course, given the striking limitations that children with LI exhibit in this area of language. However, other important areas are important in LI, and these may have at least an indirect, negative impact on morphosyntactic ability. Bishop, Adams, and Norbury (2006) have identified two fundamental impairments in children with LI that are both heritable yet show minimal etiological overlap (see also Conti-Ramsden, 2003). Not surprisingly, one of these is a reduced ability to carry out grammatical computations. The behavioral measure most frequently used to identify this limitation is a test of morphosyntactic ability, including the use of tense and agreement morphemes (e.g., Rice & Wexler, 2001). The other fundamental impairment is a deficit in the ability to retain sequences of speech sounds for brief periods of time. Nonword repetition tasks constitute the most frequent measures for this type of problem (e.g., Gathercole, Willis, Baddeley, & Emslie, 1994).

Although an ability to retain sequences of sounds is often associated with word learning (e.g., Gathercole & Baddeley, 1993), it should be clear how limitations in the ability to retain sound sequences could also play havoc with the learning of inflections. If a child cannot retain a sequence that represents an inflection that marks tense and agreement, it is likely that the acquisition of this inflection will be delayed. To the degree to which the inflection system of a language contains many different sequences, the detrimental effect of this retention problem could be considerable. This influence could occur even though retention of sound sequences and grammatical computation are genetically and etiologically distinct. First, as noted by Bishop et al. (2006), many children with LI have a double deficit—a deficit in both of these areas. Second, although poor retention of sound sequences appears to be a deficit distinct from a deficit in grammatical computation, if the inflection system of a language involves many different sequences, each of which must be detected and retained by the child, the functional relationship between these two areas may be stronger than in a language such as English.

The Contribution of Hungarian

Hungarian possesses characteristics that make it extremely useful for evaluating the morphological richness and agreement deficit accounts. Research on LI in this language, then, might not only contribute to the development of clinical assessment and treatment methods for Hungarian-speaking children with LI but also to theory development or refinement. We provide a more detailed description of the structure of Hungarian tense and agreement morphology in the next section. However, some of the highlights of Hungarian and its relevance to these accounts of LI can be stated here. Hungarian is a null-subject language with inflections for tense and inflections that simultaneously mark agreement with the subject in person and number and agreement with the object (if any) in definiteness. The agreement deficit account assumes that the difficulty with agreement resides in the agreement features of the verb. Therefore, even in a null-subject language such as Hungarian, agreement inflections will be difficult for children with LI. This may be especially so given that agreement is of two different types—agreement between the subject and verb, and agreement between the verb and the object. Errors of agreement are expected to be default forms such as present third person singular. However, tense features are not affected; for this reason, errors on the tense marking of inflections are not predicted.

According to the morphological richness account, children with LI acquiring a language such as Hungarian, in which inflectional morphology plays a central role, will

Table 1. Inflections and their allomorphs for the four paradigms tested in the study.

Tense	Person	Definite (e.g., Én tolom a dobozt "I am pushing the box") Singular	Plural	Indefinite (e.g., Én tolok egy dobozt "I am pushing a box") Singular	Plural
Present	1st	*-om/em/öm*	*-juk/jük*	*-ok/ek/ök*	*-unk/ünk*
	2nd	*-od/ed/öd*	*-játok/itek*	*-sz/ol/el/öl*	*-tok/tek/tök*
	3rd	*-ja/i*	*-ják/ik*	0	*-nak/nek*
Past	1st	*-tam/tem*	*-tuk/tük*	*-tam/tem*	*-tunk/tünk*
	2nd	*-tad/ted*	*-tátok/tétek*	*-tál/tél*	*-tatok/tetek*
	3rd	*-ta/te*	*-tták/ték*	*-t/ott/ett/ött*	*-tak/tek*

differ from typical peers to a lesser extent than in a language such as English. However, this account explicitly predicts that the processing capacity of children with SLI will begin to reach its limits when four dimensions must be considered simultaneously as in Hungarian, in which tense, person, number, and definiteness play a role in the verb inflection system. Errors should not be default forms; rather, inflections that differ from the correct inflection on only a single dimension (e.g., present first person singular indefinite in place of present first person plural indefinite) should be the most likely. Accuracy will be greater for inflections with higher frequency of occurrence in the language.

Hungarian is also a highly suitable language to evaluate the role that limitations in the ability to retain sound sequences might play in the use of tense and agreement inflections by children with LI. Although problems in nonword repetition are notorious in this population, their effects on tense and agreement inflection use has not yet been put to a stringent test, as the languages studied have relatively sparse inflection systems. In contrast, the verb inflections of Hungarian make 24 different distinctions, with all but one of these involving two or more different allomorphs. Problems in the retention of sound sequences might well slow the development of inflections in this language. If problems of this type are playing a role, the children's accuracy with inflections should be related to factors such as inflection length and nonword repetition ability.

A Sketch of Hungarian Tense and Agreement Morphology

In Hungarian, verb inflections mark tense and mode, agreement with the subject in person and number, and agreement with the object in definiteness. (Of these dimensions, distinctions according to mode are not examined in the present study; all inflections assessed are in the indicative.) Although Hungarian is often referred to as an agglutinating language, the dimensions of person and number are clearly fusional, and there is a complex relationship between agglutinating and fusional elements. We will return to this issue after introducing the verb inflections under investigation.

Table 1 provides the tense and agreement inflections with their allomorphs. Table 2 shows the tense and agreement inflections applied to the verb *tol* "push."[1] Inflections appear in bold for ease of illustration. In these tables, we divide the inflections into four "paradigms." However, this division is primarily for illustrative purposes, as the inflections for tense, person, number, and definiteness can be viewed as a single paradigm.

Several details can be noted from an inspection of the tables. First, Hungarian's use of agreement between the verb and the object (in definiteness) as well as between the subject and the verb (in person and number) effectively doubles the size of the paradigm. The number of inflections that must be learned by Hungarian-speaking children, then, is quite large indeed. Verb–object agreement is typologically much less common than subject–verb agreement. In fact, many languages show subject–verb agreement without verb–object agreement, but the reverse does not seem to occur. Note from the tables that any difficulty that is specific to verb–object agreement should be detectable. For example, in contexts requiring a present first person singular form, a child might produce *tolok* instead of *tolom* (or vice versa).

The indefinite conjugation is regarded as unmarked. It is used with intransitive verbs as well as with transitive verbs with indefinite objects. It is also employed when the object is a first or second person

[1]For ease of exposition, we use standard Hungarian orthography and do not give phonetic transcriptions. Hungarian orthography is fairly transparent, geminates are marked by double consonants (also by doubling the first letter in a consonant digraph), and accents above vowels mark length. However, not every accented vowel is phonetically equivalent to their short counterpart, so we present the phonetic symbols for Hungarian vowels and nontransparent consonantal letters here. Vowels: a [ɔ], á [a:], o [o], ó [o:], u [u], ú [u:], e [ɜ], é [e:], i [i], í [i:], ö [ø], ő [ø:], ü [y], ű [y:]; consonants: c [ts], cs [ʧ], dzs [dʒ], g [g], gy [ɟ], j [j], ly [j], ny [ɲ], r [r], s [ʃ], sz [s], ty [ç], zs [ʒ].

Table 2. Inflected forms for *tol* "push" in the four paradigms tested in the study.

Tense	Person	Definite (e.g., Én tolom a dobozt "I am pushing the box") Singular	Plural	Indefinite (e.g., Én tolok egy dobozt "I am pushing a box") Singular	Plural
Present	1st	*tol**om***	*tol**juk***	*tol**ok***	*tol**unk***
	2nd	*tol**od***	*tol**játok***	*tol**sz***	*tol**tok***
	3rd	*tol**ja***	*tol**ják***	*tol*	*tol**nak***
Past	1st	*tol**tam***	*tol**tuk***	*tol**tam***	*tol**tunk***
	2nd	*tol**tad***	*tol**tátok***	*tol**tál***	*tol**tatok***
	3rd	*tol**ta***	*tol**ták***	*tol**t***	*tol**tak***

pronoun.[2] The definite conjugation is chosen when the object noun phrase (NP) is clearly marked with a definite article (*a* or *az* "the") and when the object is a possessively modified noun. Proper names as object NPs also take the definite conjugation. There are additional factors that are associated with the choice of a definite or indefinite inflection that go beyond the scope of the present study. For a more detailed description, see Bartos (1997) and MacWhinney and Pléh (1997).

A second notable detail that is evident in Tables 1 and 2 is the relatively large number of allomorphs. Most of the variation in the form of the inflection is a function of the vowel harmony rules of Hungarian. These rules seem to be acquired at a rather young age by Hungarian-speaking children (e.g., MacWhinney, 1985), even if they render the relationship between agreement inflections in present and past tense less clear. Other allomorphs are a product of phonological conditioning. Chief among these is the present indefinite second person singular allomorph, *-sz,* whose form is determined by the particular consonant appearing at the end of the verb stem.

Many languages with rich inflectional paradigms do not permit bare verb stems. Hungarian is an exception, in that the present indefinite third person singular inflection is a "zero" form, as in *tol*. The existence of a finite bare stem form in Hungarian means that, in principle, a child could employ such a form as a default whenever the appropriate inflected form is not known or is difficult to retrieve in the moment. Finally, it can be seen in Tables 1 and 2 that there is minimal syncretism (MacWhinney & Pléh, 1997); the only neutralization occurs in the past first person singular forms where the same inflection is used for both definite and indefinite objects (thus, *toltam* is used for both "I was pushing the box" and "I was pushing a box").[3] Hereafter, we employ the following abbreviations: "1", "2", and "3" for first, second, and third person, respectively; "Sg" for singular and "Pl" for plural; "Pres" for present and "Past" for past; and "Indef" for indefinite and "Def" for definite.

The subject–verb agreement (for person and number) reflected in Tables 1 and 2 corresponds to that seen in many other languages (apart from its fusion with definiteness marking). However, Hungarian subject–verb agreement operates somewhat differently because quantified nouns do not formally agree in number with their quantifiers. For example, *ten bottles* is expressed with a singular noun *tíz üveg* "ten bottle" rather than a plural noun **tíz üvegek* "ten bottles." The same is true for nouns preceded by terms corresponding to "many," "some," and "all." This characteristic has implications for subject–verb agreement because agreement is based on formal marking and not conceptual plurality. Thus, a subject such as "ten bottle" would require a verb inflected for singular.

The relationship between agglutinating and fusional elements of the inflection system is very complex. When (past) tense is overtly marked, this element precedes elements reflecting person and number. Thus, in Table 2 it can be seen that in the indefinite past third person plural, past tense *–t-* precedes third person plural *–unk*; the present tense counterpart has no overt tense element preceding *–unk*. However, for inflections marked for definite, position is less transparent. For example, whereas definite past third person plural has the sequence *–t-uk,* definite present third person plural has the sequence *–j-uk,* with *–j-* representing an element marking definiteness, not tense. This complexity has led to proposals (e.g., Rebrus, 2005) that the same position can serve more than one grammatical function, depending on the particular tense, definiteness, and person and number features involved. Phonologically conditioned allomorphy in Hungarian can also reduce the transparency of the agglutinating elements of the inflections. For example, whereas *tolom* is the form for definite present first person singular "I am pushing," the form *toltam* is used for definite past first person singular "I was pushing," not **toltom,* due to lowering of mid-vowels after past tense *–t-*.

[2]There is also a special inflection in the indefinite conjugation when the subject is first person singular and the object is in the second person, expressing both persons in a single inflection, as in *tol-lak* "I push you."
[3]Here we are constraining our description to the section of the verbal paradigm under investigation in our study.

Hungarian-Speaking Children: Previous Findings

Although no systematic experimental examination has been done thus far on the development of agreement marking by typically developing Hungarian-speaking children, two case studies (Lengyel, 1981, data from a boy between 1;0 and 3;0 [years;months]; Meggyes, 1971, data from a girl between 1;8 and 2;2) and a more extensive analysis of data from 3 Hungarian children between 1;8 and 2;9 from the CHILDES database (Babarczy, 2005) report errors in agreement or other inflection details. According to these studies, the very first verb forms are usually either imperative forms or third person singular declarative forms that are sometimes applied to non-third-person referents. In early verb usage, Hungarian children generally use all three singular forms together with Pl1 to express Sg1 meanings. For example, in contexts requiring *tolok* "I am pushing [indefinite]," a child might produce *tolok, tolsz, tol, or tolunk* (see Table 2). Because these utterances usually lack a subject, there is no overt error of subject–verb agreement in such utterances. Based on these three studies, there seems to be individual variation in the extent children use Sg2 as a substitute for Sg1, but for some children such errors are more common in the beginning than Sg3 substitutions, which frequently occur with all children and for a longer period. Pl2 first appears in imperative form, and even when it does appear in declarative form, it is fairly uncommon. There are very few errors in marking Pl3 from the beginning, but these forms are also not frequent. Past tense forms also appear toward the end of the second year, and at first they are generally used to express completed actions.

Babarczy (2005, 2007), in her analysis of CHILDES data from 6 Hungarian children between 1;8 and 2;10, found many errors in definiteness agreement, revealing the children's preference for using the default indefinite form with a definite object (she was focusing on imperative forms) and fewer errors in subject–verb agreement. Based on a comparative analysis of early verb forms, she found that subject–verb agreement is delayed in English relative to Hungarian. Interestingly, she also observes that there is no sentence length effect on the agreement errors that young Hungarian-speaking children make. Lengyel (1981) points out that although mixing up first and third person is common in the indefinite conjugation, it is very rare in the definite conjugation. In summary, typically developing children first mainly use singular forms, they most often to refer to first person, and they make many errors of using Sg3, Sg2, and Pl1 forms for Sg1 meanings. Indefinite verb forms are sometimes used in place of definite forms.

Systematic studies of Hungarian-speaking children with LI have also been few in number. Vinkler and Pléh (1995) reported on a child with LI who had difficulty with noun as well as verb morphology. This child often resorted to a more frequently occurring inflection as a substitute for the required form. Marton, Schwartz, Farkas, and Katsnelson (2006) compared the working memory performance of Hungarian-speaking and English-speaking children with specific language impairment. They found that, for the Hungarian-speaking children, morphological complexity played a larger role than sentence length, whereas syntactic complexity was the most influential factor for the English-speaking children.

Hypotheses

Given the details of tense and agreement inflections in Hungarian, several hypotheses can be advanced. First, according to the agreement deficit account, children with LI should be significantly less accurate than their typically developing peers in the agreement details of the inflections. Errors are likely to be default forms such as third person singular forms. Tense should be correctly marked. According to the morphological richness account, the rich inflectional morphology and null-subject character of Hungarian will lead children with LI to make much more use of tense and agreement inflections than is the case for children with LI in English. However, the four dimensions of tense, definiteness, person, and number that are required in Hungarian inflections (rather than the more commonly encountered three dimensions seen in other languages studied) will place demands on these children's limited processing capacity, leading to small but statistically reliable differences between children with LI and typically developing children. When errors are observed, a disproportionate number should constitute near misses, with no single dimension consistently serving as the source of error. Substitute inflections that are exceptions to the near-miss pattern will tend to have higher frequency of occurrence in the language. Default forms should not be seen. If nonmorphosyntactic language processing factors such as poor retention of sound sequences are involved, errors not clearly attributable to the number of dimensions involved in the inflections should be found, and the children's use of inflections should prove to be related to factors such as the length of the inflection and the children's ability in nonword repetition.

Method

Participants

Fifty children participated in the study. Twenty-five children were selected for the LI group from two special schools for children with language impairments. All of these children met the criteria for LI. Each child scored above 85 on the Raven's Coloured Progressive Matrices (Raven, Court, & Raven, 1987), a measure of nonverbal intelligence. All children passed a hearing screening, and no child had a history of neurological impairment. Each child scored at least 1.5 *SD*s below age norms on at least two of four language tests administered. These four tests included two receptive tests and two expressive tests. The receptive tests were the Hungarian standardizations of the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1981; Csányi, 1974) and the Test for Reception of Grammar (TROG; Bishop, 1983). The expressive tests were the Hungarian Sentence Repetition Test (Magyar Mondatutánmondási Teszt [MAMUT]; Kas & Lukács, 2008) and the Hungarian Nonword Repetition Test (Racsmány, Lukács, Németh, & Pléh, 2005). The rationale for including a nonword repetition test (described below) in the assessment battery is that the ability to repeat nonwords has proven to be one of the most accurate means of identifying children with LI (e.g., Dollaghan & Campbell, 1998; Tager-Flusberg & Cooper, 1999), demonstrating excellent sensitivity and specificity, and seems to be one of the fundamental and heritable weaknesses seen in this type of disorder (Bishop et al., 2006).

Although the PPVT was used as one of the language tests in our selection battery, it was also used as the basis for matching participant groups, as seen in a subsequent section. The Hungarian adaptation of the original TROG is being standardized on children from 4 to 12 years of age.[4] Items assess the children's comprehension of increasingly difficult grammatical structures. The test consists of 20 blocks, each with four sentences of the same construction (such as sentences with comparatives, postmodified subjects, and embedded clauses). The test has an 80-page booklet, each with four pictures, and on each page the child must point to the picture that matches the sentence spoken by the experimenter. A block is considered completed if the child responds correctly to all four pictures in the block. Performance is measured in terms of number of blocks correctly completed.

The Hungarian Sentence Repetition Test (MAMUT; Kas & Lukács, 2008) manipulates length and structural complexity independently. Its 40 sentences are distributed evenly across five types of grammatical constructions: (a) simple subject-verb-object (SVO), (b) simple OVS sentences, (c) complex sentences containing SS relative clauses, (d) SO relative clauses, and (e) OS relative clauses. Sentence length varies between 8 and 15 syllables within each type. The task of the participant is to immediately and accurately repeat the sentences presented by the experimenter. Performance is measured in terms of the number of correctly repeated sentences, which can be evaluated based on grouping by syllable number and by grammatical construction.

The Hungarian Nonword Repetition Test (Racsmány et al., 2005) requires the repetition of meaningless but phonotactically licit strings of Hungarian phonemes. The test contains 36 nonwords between one and nine syllables in length. Each length is represented by four nonwords. The phonological structure of the nonwords does not reflect frequency distributions of Hungarian phoneme sequences, but the test avoids sequences that would be articulatorily difficult for speakers. The span of the participant is the highest syllable number for which he or she could correctly repeat at least two out of the four nonwords.

The remaining 25 children were typically developing. These children scored above –1 *SD* on each of the four language tests that were administered to the children with LI. These children were matched with the LI group on the basis of their raw scores on the PPVT. Because the children with LI scored below age level on the PPVT, the typically developing children matched on this measure were younger. A typically developing child was considered a match if his or her PPVT score was within 3 points of the PPVT score of a child in the LI group. Hereafter, this group will be referred to as the vocabulary control (VC) group. The use of younger typically developing children matched on a nongrammatical language measure was designed to detect whether the difficulties of the children with LI on tense and agreement morphology exceeded their more general limitations in language. If so, group differences favoring the VC group should be seen. Of course, differences in the two groups' pattern of use across the different tense and agreement morphemes was also of interest. Means for age (in years; months) and raw scores on each of the tests together with ranges for both groups are given in Table 3.

Method

Given the large number of tense and agreement inflections in Hungarian, we devised a structured method of eliciting responses that ensured multiple opportunities

Table 3. Means (and ranges in parentheses) for the language impairment (LI) and vocabulary controls (VC) groups for age in years; months and in raw scores on the Peabody Picture Vocabulary Test (PPVT), the Test for the Reception of Grammar (TROG), the Nonword Repetition Test, and the Sentence Repetition Test.

Variable	LI	VC
Age	9;10 (7;6–11;10)	7;1 (5;2–8;5)
PPVT	91.3 (61–114)	92.1 (62–115)
TROG (blocks correct)	12.30 (8–18)	13.76 (6–20)
Nonword Repetition Test	3.5 (1–5)	5.8 (3–8)
Sentence Repetition Test	22.0 (0–39)	33.6 (18–40)

[4]We thank Dorothy Bishop for providing us with the Test for the Reception of Grammar (TROG) for this purpose. Thus far, 600 typically developing children have been seen as part of the norming process; the scores for the children with LI were compared against the values obtained for the typically developing children.

for the child to produce each inflection of interest. The children were asked to repeat sentences; however, the target inflections in each sentence were actually masked by a carefully inserted cough that prevented the child from hearing the inflection but not the stem or the remaining portions of the sentence. This method was adapted from Warren's (1970) phoneme restoration procedure. The restoration effect has been demonstrated at the morpheme level as well, such as for affixes in Hungarian (Dankovics & Pléh, 2001), but the effect has not yet been exploited in developmental studies as an elicited production method. Importantly, in our study the fully audible portions of the sentence (notably, the temporal adverbial, the person and number of the subject, and the definiteness of the object) made it clear (to a mature speaker of Hungarian) which verb inflection was the appropriate one to use. The child was only asked to repeat the sentences and was not told that information was missing.

Specifically, children were instructed to repeat sentences they heard through a loudspeaker. The sentences were recorded by a female speaker and digitized with coughs inserted to replace the inflections only (see subsequent elaboration). All sentences were normalized for a length between 8 and 14 syllables. Although the target inflections in the middle of the sentence were replaced by a cough, the remainder of the sentence contained all the source features for unequivocal identification of the missing inflection. Children occasionally commented that the speaker was coughing a lot; in these cases, we told them that she had a cold, and that they should just disregard the coughs.

Six verbs were used in both present and past tense; definite and indefinite conjugations; singular and plural; and in first, second, and third person. Thus, 144 sentences (6 × 2 × 2 × 2 × 3 = 144) were created. The sentences were blocked according to tense and definiteness paradigm. That is, all 36 sentences marked for present definite were presented together, as were the 36 sentences marked for present indefinite, past definite, and past indefinite. Children were tested in at least two different sessions, with the order of the four blocks counterbalanced across children.

Given the vowel harmony involved in the allomorph used for the inflection, we selected three verbs whose stems had front vowels and three that had stems with back vowels. The six verb stems selected for the task were: *tol* "push," *olvas* "read," *simogat* "stroke (pet)," *kerget* "chase," *épít* "build," and *fésül* "comb."

All sentences were simple SVO sentences. Past tense sentences were systematically longer than present tense sentences because they contained the temporal adverbial *tegnap* "yesterday," used to make the past time of the described event clear. (Hungarian does not possess a temporal adverbial that is unique to present tense.) The subsequent examples illustrate the types of sentences used for each tense and definiteness combination. The location of the inflection masked by a cough is indicated by "XXX."

1. *Mi olvas*XXX *egy mesét.*
   Target: *olvas**unk*** ["read" 1PlPresIndef]
   "We are reading a story."
2. *A gyerekek simogat*XXX *a malacot.*
   Target: *simogat**ják*** ["stroke" 3PlPresDef]
   "The children are petting the pig."
3. Tegnap én építXXX egy tornyot.
   Target: *épít**ettem*** ["build" 1SgPastIndef]
   "Yesterday I built a tower."
4. *Tegnap te tol*XXX *a biciklit.*
   Target: *tol**tad*** ["push" 2SgPastDef]
   "Yesterday you pushed the bike."

It was important to ensure that the inserted coughs were sufficient to obscure the inflection and that there were no anticipatory coarticulatory cues in the verb stem that might have provided the children with an indication of the inflection that was masked. Accordingly, we extracted the verb stem plus cough from each recorded sentence and presented them to 15 adult listeners. The listeners were asked to guess which inflection was used with the stem in each case (for all 144 verb forms). For every item, they had to select from 24 possible forms, and they guessed correctly on 5.6% of the items, which, as will be seen, is significantly below the performance level for either group of children (LI = 62%, χ^2 test, $p < .001$; VC = 83%, χ^2 test, $p < .001$). These findings indicated that our stimuli probably did not contain unintended cues that could lead to correct performance without knowing the appropriate inflection. In fact, the adult listeners' guessing behavior suggested that other factors were influencing their choices. The log frequency of allomorphs in Hungarian based on the Hungarian Webcorpus (Halácsy et al., 2004; Kornai, Halácsy, Nagy, Trón, & Varga, 2006) was a significant predictor of the frequency of the listeners' specific choices (R^2 = .132, B = 0.363, $p < .001$). Not surprisingly, the items whose inflections happened to correspond to the listeners' most frequent choices were most likely to be guessed correctly. However, even the inflection type that was most frequently guessed correctly was associated with only 14% accuracy.

Scoring

Our scoring method emphasized accuracy of tense and agreement marking rather than accuracy of the sentence as a whole. That is, we allowed for differences between the child's response and the stimulus sentence

Table 4. Examples of different types of errors or deviations from the target sentence for the stimulus sentence *Tegnap ti fésül**tétek** az oroszlánt* "Yesterday you (Pl2) were combing (comb PastDefPl2) the lion."

Response type	Child's response	Translation
Person error	Tegnap ti fésül**tük** az oroszlánt.	Yesterday you (Pl2) were combing (PastDefPl1) the lion.
Number error	Tegnap ti fésül**ted** az oroszlánt.	Yesterday you (Pl2) were combing (PastDefSg2) the lion.
Tense error	Tegnap ti fésülitek az oroszlánt.	Yesterday you (Pl2) are combing (PresDefPl2) the lion.
Definiteness error	Tegnap ti fésül**tetek** az oroszlánt.	Yesterday you (Pl2) were combing (PastIndefPl2) a lion.
Nontarget verb with correct agreement	Tegnap ti fésülköd**tétek** az oroszlánt.	Yesterday you (Pl2) were combing (reflexive, PastDefPl2) the lion.
Nontarget subject or object with correct agreement	Tegnap ti fésültetek egy oroszlánt.	Yesterday you (Pl2) were combing (PastIndefPl2) a lion.

provided that the child's response showed internally accurate agreement as well as tense marking. This scoring method was selected to reduce the effects of recall errors and to provide as clear a view of inflection use as possible to evaluate the agreement deficit and morphological richness accounts—two accounts expressly developed to explain the tense and agreement inflection problems of children with LI.

According to this scoring method, if the child used a nontarget verb with correct inflection or if the child used a different subject or object but the verb inflections were appropriate for this change, the response was scored as correct. In addition, if a child produced a past tense form when the stimulus sentence was in present tense (without any other change), the child was credited with a correct response. Although in such cases it is more customary to assume such sentences are in present tense, recall that there is no adverbial that is unique to present tense. (To use the closest English equivalent, whereas we must use past tense with "yesterday," either past or present tense might be appropriate with "today.") As Hungarian has somewhat flexible word order, variations in word order were also permitted, provided that all of the above details were included. Using this method, the following errors could occur: (a) person error; (b) number error; (c) tense error; (d) definiteness error; or (e) other error, such as a sentence that bore no resemblance to the stimulus sentence. If errors (a)–(e) or any of their combinations occurred, the answer was scored 0. The children's use of the wrong allomorph in otherwise correct responses was also noted but was not scored as an error.[5] Examples of error types and deviations from the stimulus sentence that were counted as correct are shown in Table 4.

To assess interjudge reliability, the responses of five children in each group were selected at random and were scored by an independent judge. Percentage agreement ranged from 97.2 to 100.0, with similar percentages of agreement for the LI (M = 98.75) and VC (M = 99.60) groups.

Data Analysis

The data were examined in several ways. First, we examined the children's percentages of correct responses for each inflection type, using a general linear model analysis of variance (ANOVA) with Group as a between-subjects factor and Tense, Definiteness, Number, and Person as within-subjects factors. Second, given the predictions of the morphological richness account, we determined whether the children's scores were related to frequency of occurrence factors. For each inflected verb form, we calculated the following: (a) inflected word frequency (the frequency of the exact inflected verb form), (b) inflection frequency (e.g., the frequency of all PresDefSg3 allomorphs combined), and (c) allomorph frequency (mostly conditioned by stem category for vowel harmony; e.g., the frequency of the *-ja* allomorph of PresDefSg3). The source of frequency data was the Hungarian Webcorpus (Halácsy et al., 2004; Kornai et al., 2006). Calculations employed the logarithm of frequency. Finally, we performed an analysis of the children's errors to test the prediction of the morphological richness account that near-miss errors would be disproportionately high relative to errors differing from the correct form on more than one dimension.

Results

Accuracy According to Group and Inflection Type

The ANOVA on accuracy revealed Group as a significant main effect, $F(1, 48) = 10.02$, $\eta^2 = .173$, $p < .01$. With the exception of Definiteness, $F(1, 48) = 0.09$, *ns,* all within-subjects factors proved to be significant main effects: Tense, $F(1, 48) = 13.91$, $\eta^2 = .225$, $p < .01$; Number,

[5] We also used a second scoring method that was more stringent. This method required that the target verb (in correctly inflected form) be used in the child's response, and no changes were allowed in the person and number of the subject or the definiteness of the object. The pattern of results seen for this scoring method matched those seen for our first scoring method, except that the group effects were even stronger.

$F(1, 48) = 8.91$, $\eta^2 = .157$, $p < .01$; and Person, $F(1, 48) = 27.19$, $\eta^2 = .362$, $p < .001$. The significant interactions were Tense × Definiteness × Person, $F(2, 96) = 7.22$, $\eta^2 = .131$, $p < .01$; Number × Person, $F(2, 96) = 10.05$, $\eta^2 = .180$, $p < .001$; Definiteness × Number × Person, $F(2, 96) = 8.85$, $\eta^2 = .156$, $p < .001$; and Tense × Definiteness × Number × Person, $F(2, 96) = 4.81$, $\eta^2 = .156$, $p < .05$. Pairwise comparisons (LSD tests) at the .05 level revealed that past, plural, and first person were significantly more difficult than present, singular, and first and third person, respectively (first and third person did not differ). Figure 1 provides an illustration of the findings.

It can be seen that overall performance of the LI group was significantly lower than that of the VC group, but no interactions with Group were significant, suggesting that the two groups basically showed the same pattern of performance across the dimensions examined. The interactions involving Person and Number were due to low scores of Second Person and, especially, of Pl2 forms. These difficulties are evident from Figure 1.

Relationship With Frequency

We examined the relationship between several frequency factors and the children's use of the tense and agreement inflections. According to the morphological richness account, children should have greater success producing more frequently occurring inflections than less frequently occurring inflections. However, it is also true that other details such as the frequency of the words themselves could also influence the children's success. To determine if these factors could predict performance on the experimental task, we included them in stepwise regression analyses. We tested the effects of log-inflected word frequency, log inflection frequency, and log allomorph frequency on the total number of correct responses for each test item, separately for the LI and VC groups. Only variables that showed a significant correlation ($p < .05$) with the target variable were entered into the analysis.

For both groups, the factor that best contributed to predicting performance levels was log inflection frequency. As can be seen in Table 5, the LI data are somewhat better predicted by this factor, where it explains 31% of variance, as opposed to 20% explained in the VC group.

Error Analysis

Both groups of children produced many errors on the task. Out of the 3,600 responses from each group, the VC group produced 371 errors (10.0%), and the LI group erred on 905 (25.1%) responses. It is notable that the number of inappropriate productions of the present

Figure 1. Mean percentage correct for each inflection type for the language impairment (LI) and vocabulary controls (VC) groups. Standard errors are also shown.

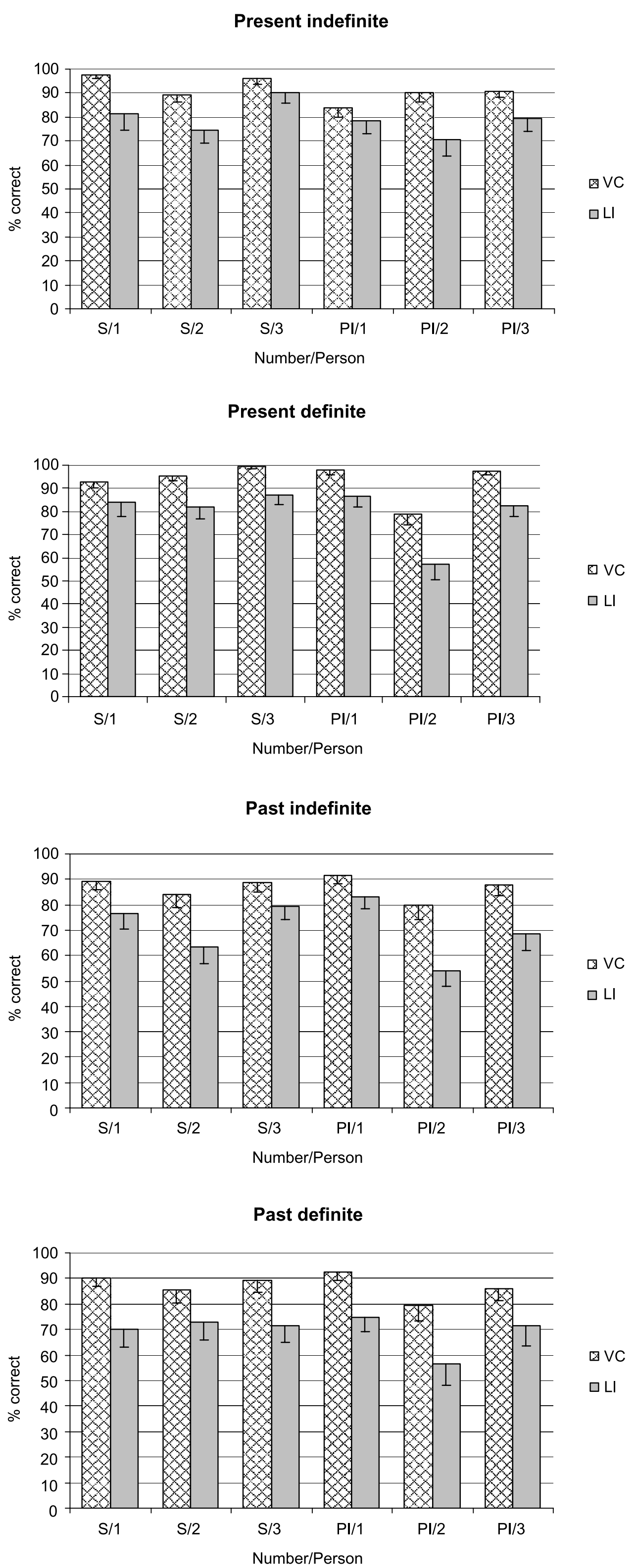


Table 5. Log inflection frequency as a predictor of the performance of the LI and VC groups.

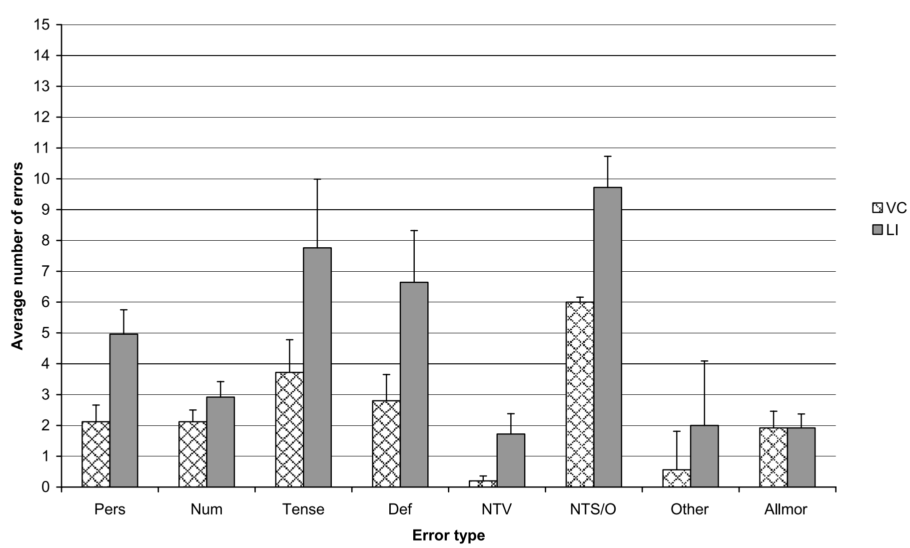
Group	Predictor	B	p	R^2
VC	Log inflection frequency	0.45	< .001	.20
LI	Log inflection frequency	0.56	< .001	.31

Note. R^2 shows the amount of variance in the data explained by the predictor.

third person singular indefinite—the zero-marked form—was not especially high, suggesting that this form was not used as a default. This zero-marked form constituted only 6.8% of the errors in the VC group and 5.2% of the errors in the LI group. Inappropriate productions of these zero-marked forms were outnumbered by the inappropriate production of inflected forms. For example, the incorrect production of present third person singular definite forms represented 8.2% of the errors for each group, and inappropriate productions of present first person plural definite forms constituted 9.4% of the errors for the VC group and 13.0% of the errors for the LI group.

Figure 2 provides the mean number of errors according to error type. Numbers for each error type represent errors that constituted an error only on that single dimension. Along with the responses treated as errors in the preceding analyses, we include in Figure 2 nontarget responses that were scored as correct in those analyses, namely, the use of a nontarget verb with correct tense and agreement (NTV), the use of a nontarget subject or object with correct agreement (NTS/O), and the use of an incorrect allomorph (Allmor) even though agreement was correct. Figure 2 illustrates several group differences, but not all of them are confirmed by statistical analysis. The LI group made more single-dimension errors overall, $F(1, 49) = 9.2$, $\eta^2 = .21$, $p < .01$. ANOVAs were also performed for each error type separately. The difference reached significance for Person, $F(1, 49) = 8.8$, $\eta^2 = .155$, $p < .01$, and Definiteness, $F(1, 49) = 4.16$, $\eta^2 = .08$, $p < .05$, but not for Number, $F(1, 49) = 1.6$, *ns,* or Tense, $F(1, 49) = 2.68$, *ns*. More detailed comparison of dimension errors across groups shows that among person errors, the LI group only made significantly more errors than VC children in using third person forms, $F(1, 49) = 8.75$, $\eta^2 = .154$, $p < .01$. In definiteness errors, the difference was only significant with using indefinite forms when the target was definite, $F(1, 49) = 7.98$, $\eta^2 = .143$, $p < .01$. The remaining response type treated as an error in the earlier analyses, Other, also revealed a difference between the two groups of children, $F(1, 49) = 4.93$, $\eta^2 = .093$, $p < .05$. None of the deviations from the target originally scored as correct showed a group difference, such as NTV, $F(1, 49) = 1.97$, *ns,* and NTS/O, $F(1, 49) = 2.34$, *ns*. Finally, although use of the wrong allomorph (Allmor) was not considered an error, it can be seen from Figure 2 that the two groups were highly similar in this regard, suggesting that rules of vowel harmony were well established and did not seem to be an area of particular difficulty for the LI group. An inspection of Figure 2 reveals that although the children with LI made a greater number of errors than the VC children, the pattern of errors across error types was highly similar in the two groups.

Figure 2. Mean number of errors on different error types in the two groups. Only errors in a single dimension are counted. Standard errors are also shown. Pers = person; Num = number; Def = definiteness; NTV = nontarget verb with correct tense and agreement; NTS/O = nontarget subject or object with correct agreement; Allmor = incorrect allomorph.



The morphological richness account predicts that single-dimension or "near-miss" errors will be especially frequent. To test this prediction, we compared the children's near-miss errors to productions that constituted an error on more than one dimension (e.g., an error of tense plus number). Of the 23 inflections that could be used as an incorrect substitute for the correct inflection, 5 differed from the target on only one dimension, 9 differed from the target on two dimensions, 7 differed on three dimensions, and only 2 differed on all four dimensions. This was true for all 24 target inflections. In Table 6, we provide the number of substitution errors for each target inflection. Given that the four types of errors had different probabilities (the most probable were two-dimension errors, the least probable were four-dimension errors), we created adjusted scores by dividing the total number of errors of each type by the number of different inflections that could have created each error type. From Table 6, it can be seen that for 23 of the 24 target inflections for the LI group, a higher total number of errors was seen for one-dimension errors than for each of the other error types. When adjusted scores are considered, the differences are even more dramatic, with all 24 inflections having more one-dimension errors than errors of the other types. This was confirmed by a repeated measures ANOVA by target inflection type performed for each participant group. The analysis for the LI group revealed a highly significant difference, $F(2, 46) = 93.12$, $\eta^2 = .802$, $p < .001$. Post hoc testing at the .05 level revealed that one-dimension errors ($M = 4.81$, $SD = 2.22$) were significantly more frequent than two-dimension errors ($M = 0.86$, $SD = 0.59$), which, in turn, were more frequent than three-dimension errors ($M = 0.28$, $SD = 0.34$). Four-dimension errors were not included in the ANOVA because, as can be seen in Table 6, no errors of this type were found in the data. Nearly identical findings emerged for the VC group although, as noted earlier, these children committed fewer errors than the LI group. Specifically, a significant difference was found for error type, $F(2, 46) = 43.25$, $\eta^2 = .653$, $p < .001$, with one-dimension errors ($M = 1.88$, $SD = 1.25$) being more frequent that two-dimension errors ($M = 0.45$, $SD = 0.43$), which, in turn, were more frequent than three-dimension errors ($M = 0.09$, $SD = 0.20$). Again, four-dimension errors were not seen in the data.

This type of analysis also permitted us to assess a prediction of the agreement deficit account. One-dimension errors could have been an error in tense only, person only, number only, or definiteness only. According to the agreement deficit account, errors in tense are not expected. In fact, we found that problems with tense were concentrated in past tense items. The results indicated that the number of one-dimension errors of tense in past tense items represented 34% ($SD = 18\%$) of the total one-dimension errors by the children with LI. Given the four dimensions possible, this value is clearly in line with the expectation of 25% if difficulty with past tense were comparable to difficulty with each of the other three dimensions. Similar results were seen for the VC group; 37% ($SD = 26\%$) of their one-dimension errors in past tense items involved an error of tense.

Whereas Table 6 provides the types of errors according to the target inflection, in Table 7 we provide the types of errors according to the inflection used as a substitute. As can be seen in this table, all 24 inflections were used as a substitute by the LI group. Furthermore, all 24 were more likely to be used as a substitute when it differed from the target on one dimension than when it differed from the target on two, three, or four dimensions. This was true for total number of errors as well as for adjusted scores. A repeated measures ANOVA by substitute inflection type confirmed this difference for the LI group, $F(2, 46) = 88.35$, $\eta^2 = .793$, $p < .001$. Post hoc testing at the .05 level revealed that one-dimension errors ($M = 4.76$, $SD = 2.69$) were significantly more abundant than two-dimension errors ($M = 0.85$, $SD = 0.99$), which, in turn, were more frequent than three-dimension errors ($M = 0.28$, $SD = 0.58$). Four-dimension errors were not included in the analysis, as this type of error did not occur in the data. The findings for the VC group mirrored those seen for the children with LI. A difference according to error type was seen, $F(2, 46) = 37.38$, $\eta^2 = .619$, $p < .001$. Post hoc testing indicated that one-dimension errors ($M = 1.90$, $SD = 1.40$) occurred more frequently than two-dimension errors ($M = 0.45$, $SD = 0.62$), which, in turn, were more frequent than three-dimension errors ($M = 0.09$, $SD = 0.18$). One of the 24 inflections, third person singular definite in past tense was never used as a substitute. The remaining 23 inflections showed the same pattern evident for the group data, with greater tendency for the inflection to serve as a substitute when it differed from the correct inflection on a single dimension.

Although Table 7 clearly shows that the number of substitutions differing from the target by a single dimension was disproportionately high in the data, as predicated by the morphological richness account, these data do not provide an indication of the role of the substitute inflections' frequency of occurrence. According to the morphological richness account, substitute inflections that differ from the target on two or more dimensions are likely to have relatively strong representations, as estimated by frequency of occurrence in the language. We examined this issue by performing a regression analysis to determine if log inflection frequency served as a significant predictor of the children's tendency to use an inflection as a substitute when it differed from the correct form on two or more dimensions. Indeed, this prediction was borne out for the LI group; log inflection frequency explained 20% of the variance associated with substitutions differing from the target on two or more

Table 6. The number of times the target inflection was replaced by a substitute inflection that differed from the target on one, two, three, or four dimensions, and the adjusted score (Adj Score), computed by dividing the total by the number of different inflections that had the potential to differ from the target on the same number of dimensions.

TARGET	#1-Dimen Err and Adj Score		#2-Dimen Err and Adj Score		#3-Dimen Err and Adj Score		#4-Dimen Err and Adj Score
			LI GROUP				
PRIDSG1	20	4.00	6	0.67	0	0	0
PRIDSG2	21	4.20	11	1.22	3	0.43	0
PRIDSG3	8	1.60	1	0.11	3	0.43	0
PRIDPL1	24	4.80	1	0.11	1	0.14	0
PRIDPL2	24	4.80	10	1.11	4	0.57	0
PRIDPL3	26	5.20	4	0.44	0	0	0
PRDSG1	15	3.00	3	0.33	0	0	0
PRDSG2	8	1.60	10	1.11	2	0.29	0
PRDSG3	12	2.40	3	0.33	1	0.14	0
PRDPL1	13	2.60	2	0.22	1	0.14	0
PRDPL2	40	8.00	11	1.22	1	0.14	0
PRDPL3	22	4.40	0	0	0	0	0
PAIDSG1	21	4.20	9	1.00	0	0	0
PAIDSG2	34	6.80	13	1.44	8	1.14	0
PAIDSG3	21	4.20	5	0.56	1	0.14	0
PAIDPL1	16	3.20	6	0.67	1	0.14	0
PAIDPL2	54	10.80	11	1.22	2	0.29	0
PAIDPL3	32	6.40	9	1.00	2	0.29	0
PADSG1	23	4.60	17	1.89	1	0.14	0
PADSG2	18	3.60	11	1.22	9	1.29	0
PADSG3	20	4.00	11	1.22	5	0.71	0
PADPL1	31	6.20	1	0.11	0	0	0
PADPL2	44	8.80	20	2.22	1	0.14	0
PADPL3	30	6.00	12	1.33	1	0.14	0
M	4.81		0.86		0.28		0
SD	2.22		0.59		0.34		
			VC GROUP				
PRIDSG1	3	0.60	1	0.11	0	0	0
PRIDSG2	10	2.00	6	0.67	0	0.43	0
PRIDSG3	5	1.00	1	0.11	0	0.43	0
PRIDPL1	4	8.00	0	0.11	0	0.14	0
PRIDPL2	7	1.40	8	0.89	0	0.57	0
PRIDPL3	13	2.60	1	0.11	0	0	0
PRDSG1	6	1.20	0	0	0	0	0
PRDSG2	2	0.40	3	0.33	0	0	0
PRDSG3	1	0.20	0	0	0	0	0
PRDPL1	0	0	3	0.33	0	0	0
PRDPL2	15	3.00	13	1.44	0	0	0
PRDPL3	4	0.80	0	0	0	0	0
PAIDSG1	12	2.40	4	0.44	0	0	0
PAIDSG2	13	2.60	11	1.22	4	0.57	0
PAIDSG3	9	1.80	8	0.89	0	0	0
PAIDPL1	10	2.00	3	0.33	0	0	0
PAIDPL2	26	5.20	4	0.44	0	0	0
PAIDPL3	9	1.80	2	0.22	1	0.14	0
PADSG1	10	2.00	3	0.33	2	0.29	0
PADSG2	9	1.80	9	1.00	5	0.71	0
PADSG3	10	2.00	3	0.33	3	0.43	0

(Continued on the following page)

Table 6 *Continued.* The number of times the target inflection was replaced by a substitute inflection that differed from the target on one, two, three, or four dimensions, and the adjusted score (Adj Score), computed by dividing the total by the number of different inflections that had the potential to differ from the target on the same number of dimensions.

TARGET	#1-Dimen Err and Adj Score		#2-Dimen Err and Adj Score		#3-Dimen Err and Adj Score		#4-Dimen Err and Adj Score
PADPL1	9	1.80	2	0.22	0	0	0
PADPL2	20	4.00	11	1.22	0	0	0
PADPL3	19	3.80	2	0.22	0	0	0
M	1.88		0.45		0.09		0
SD	1.25		0.43		0.20		

Note. PR = present tense; PA = past tense; ID = indefinite; D = definite; SG = singular; PL = plural; 1 = first person; 2 = second person; 3 = third person. Adjusted (Adj) scores are not presented for four-dimension errors (Err) because such errors did not occur. Dimen Err = dimension error.

dimensions ($B = 0.45$, $p < .05$, $R^2 = .20$). In contrast, log inflection frequency was not a significant predictor of the total number of times that an inflection served as a substitute when the number of dimensions on which it differed from the target was ignored. Clearly, the frequency of occurrence effect was limited to multi-dimension substitutions in the LI group. Identical analyses using the VC group data indicated that, as expected, log inflection frequency was not a significant predictor of the total number of times that an inflection served as a substitute when the number of dimensions was disregarded. However, log inflection frequency was also not a predictor of the number of times an inflection served as a substitute when it differed from the target on two or more dimensions. This finding differed from that observed for the LI group. As can be seen in Table 7, the number of two- and three-dimension errors was extremely low for the VC group, raising the possibility that floor effects obviated the detection of log frequency effects.

Nonmorphosyntactic Language Processing Factors

The agreement deficit account and the morphological richness account predict difficulties according to the nature of the dimension involved (e.g., agreement) or the number of dimensions involved (e.g., four) in the inflections. However, if the children's use of inflections is also influenced by factors pertaining to the retention of sound sequences, factors other than the specific nature or number of dimensions involved should be observable. One such factor is the length of the verb plus inflection, measured in number of phonemes. Accordingly, we determined whether length in number of phonemes could serve as a significant predictor of the children's accuracy of inflection use, as measured by the total number of accurate responses for each inflection. This proved true for each group. For the VC group, this factor accounted for 20% of the variance in the children's inflection accuracy scores ($B = 0.45$, $p < .001$, $R^2 = .20$); for the LI group, 31% of the variance was explained by this factor ($B = 0.55$, $p < .001$, $R^2 = .31$).

Recall, however, that log inflection frequency also proved to be a predictor of the children's accuracy of inflection use. Some inflections that were relatively low in frequency such as the second person plural inflections (e.g., *játok*, *tatok*) are also among the longest inflections. Therefore, we performed a regression analysis to determine if length in phonemes contributed to the prediction of the children's inflection accuracy even when log inflection frequency is taken into account. The results appear in Table 8. As can be seen, for each group, length in number of phonemes proved significantly related to the children's inflection accuracy along with log inflection frequency; together, these factors explained 27% of the variance in the VC data and 41% of the variance in the LI data.

The data in Table 8 address the degree to which length of the verbs with inflections related to the children's inflection accuracy, but this factor cannot be divorced from the dimensions (e.g., person, number) reflected in the inflections. To gain an impression of the role of length independent of tense and agreement, we used the children's scores on the nonword repetition test as a covariate and again compared the VC and LI groups. Although low (LI group) or age-appropriate (VC group) Nonword Repetition Test scores constituted one of the bases on which the children were selected, the typically developing comparison group (mean age = 7;1) was, on average, more than 2 years younger than the LI group (mean age = 9;10). Nevertheless, the two groups differed on this measure: LI, $M = 3.5$, $SD = 1.5$; VC, $M = 5.8$, $SD = 1.3$, $t(48) = 6.14$, $p < .001$. When nonword repetition was entered as a covariate, the group difference in inflection accuracy disappeared, $F(1, 47) = 0.68$, *ns*.

Table 7. The number of times an inflection (INFLECT) was used as an incorrect substitute (SUBST) when it differed from the target on one, two, three, or four dimensions, and the adjusted score, computed by dividing the total by the number of different inflections that had the potential to differ from the target on the same number of dimensions.

INFLECT USED AS SUBST	# Times Differ by 1 Dimen and Adj Score		# Times Differ by 2 Dimen and Adj Score		# Times Differ by 3 Dimen and Adj Score		# Times Differ by 4 Dimen and Adj Score
			LI GROUP				
PRIDSG1	15	3.00	2	0.22	2	0.29	0
PRIDSG2	8	1.60	0	0	1	0.14	0
PRIDSG3	31	6.20	10	1.11	4	0.57	0
PRIDPL1	30	6.00	22	2.44	11	1.57	0
PRIDPL2	28	5.60	3	0.33	0	0	0
PRIDPL3	23	4.60	8	0.89	1	0.14	0
PRDSG1	25	5.00	2	0.22	0	0	0
PRDSG2	16	3.20	3	0.33	0	0	0
PRDSG3	38	7.60	17	1.89	6	0.86	0
PRDPL1	61	12.20	39	4.33	17	2.43	0
PRDPL2	16	3.20	3	0.33	1	0.14	0
PRDPL3	48	9.60	8	0.89	1	0.14	0
PAIDSG1	5	1.00	0	0	0	0	0
PAIDSG2	7	1.40	2	0.22	0	0	0
PAIDSG3	23	4.60	10	1.11	0	0	0
PAIDPL1	36	7.20	17	1.89	0	0	0
PAIDPL2	23	4.60	3	0.33	0	0	0
PAIDPL3	20	4.00	6	0.67	2	0.29	0
PADSG1	8	1.60	2	0.22	0	0	0
PADSG2	10	2.00	0	0	0	0	0
PADSG3	23	4.60	7	0.77	1	0.14	0
PADPL1	35	7.00	13	1.44	0	0	0
PADPL2	17	3.40	3	0.33	0	0	0
PADPL3	26	5.20	5	0.55	0	0	0
M	4.76		0.85		0.28		0
SD	2.69		0.99		0.58		0
			VC GROUP				
PRIDSG1	7	1.40	2	0.22	0	0	0
PRIDSG2	1	0.20	2	0.22	0	0	0
PRIDSG3	11	2.20	7	0.78	0	0	0
PRIDPL1	7	1.40	1	0.11	1	0.14	0
PRIDPL2	7	1.40	11	1.22	4	0.57	0
PRIDPL3	4	0.80	2	0.22	0	0	0
PRDSG1	3	0.60	1	0.11	0	0	0
PRDSG2	11	2.20	0	0	0	0	0
PRDSG3	17	3.40	19	2.11	0	0	0
PRDPL1	12	2.40	9	1.00	3	0.43	0
PRDPL2	10	2.00	3	0.33	1	0.14	0
PRDPL3	25	5.00	2	0.22	2	0.29	0
PAIDSG1	1	0.20	0	0	0	0	0
PAIDSG2	3	0.60	0	0	0	0	0
PAIDSG3	2	0.40	1	0.11	0	0	0
PAIDPL1	20	4.00	17	1.89	4	0.57	0
PAIDPL2	12	2.40	0	0	0	0	0
PAIDPL3	6	1.20	0	0	0	0	0
PADSG1	0	0	0	0	0	0	0
PADSG2	10	2.00	1	0.11	0	0	0
PADSG3	7	1.40	4	0.44	0	0	0

(Continued on the following page)

Table 7 *Continued.* The number of times an inflection (INFLECT) was used as an incorrect substitute (SUBST) when it differed from the target on one, two, three, or four dimensions, and the adjusted score, computed by dividing the total by the number of different inflections that had the potential to differ from the target on the same number of dimensions.

INFLECT USED AS SUBST	# Times Differ by 1 Dimen and Adj Score		# Times Differ by 2 Dimen and Adj Score		# Times Differ by 3 Dimen and Adj Score		# Times Differ by 4 Dimen and Adj Score
PADPL1	23	4.60	13	1.44	0	0	0
PADPL2	18	3.60	0	0	0	0	0
PADPL3	11	2.20	2	0.22	0	0	0
M	1.90		0.45		0.09		0
SD	1.40		0.62		0.18		0

Note. Adjusted scores are not presented for four-dimension errors because such errors did not occur.

The effect of nonword repetition was significant, $F(1, 47) = 4.75$, $\eta^2 = .096$, $p < .05$. These findings suggest that factors such as ability to retain sequences of sounds may have had a bearing on the children's use of inflections on our experimental task.

Conclusions

In this study, we found that a group of Hungarian-speaking children with LI performed significantly below the level of younger VC children in a task in which the children had to repeat sentences and supply the appropriate tense and agreement inflections. Although the two groups differed in accuracy, their patterns of performance across inflection types—both in terms of inflections with greatest and least accuracy and in terms of error types—were highly similar. Before discussing the implications of these findings, we discuss some potential limitations of the study.

One potential limitation is that we cannot be certain that our task yielded results that were representative of the children's actual abilities. Studies of children with LI in other languages have typically employed spontaneous speech samples and/or sentence completion tasks. We believe our choice of tasks was highly appropriate given the characteristics of Hungarian. For example, the distinction between agreement inflections as a function of the definiteness of the object is not one that can be easily manipulated through sentence completion tasks. Despite the novel nature of our task, the higher scores by the younger VC children compared to the children with LI suggest that it was developmentally appropriate.

Table 8. Length in number of phonemes and log inflection frequency as predictors of the performance of the LI and VC groups.

Group	Predictor	*B*	*p*	R^2
VC	Number of phonemes	–0.32	< .001	.27
	Log inflection frequency	0.28	< .01	
LI	Log inflection frequency	0.38	< .001	.41
	Number of phonemes	–0.36	< .001	

Note. R^2 shows the amount of variance in the data explained by the predictor.

Another potential limitation is our use of younger typically developing children matched with the LI group according to receptive vocabulary rather than according to an expressive measure such as MLU. However, for a language with a rich morphology such as Hungarian, MLU matching would carry the risk of matching two groups on the very ability that we were wishing to compare. Nevertheless, matching on the basis of receptive vocabulary was a more stringent test of the status of tense and agreement morphology in Hungarian LI than would be the case if chronological age controls had been used. As can be seen in Table 3, the children with LI were nearly 3 years older than the VC children, yet they did not perform as well as these younger typically developing children.

Another potential criticism of the study is that given our use of a nonword repetition test and a sentence repetition test as two of the four tests in our diagnostic battery, it might be argued that we selected only or primarily those children with LI with limitations in working memory. However, all of the children with LI earned low scores on the PPVT—a receptive vocabulary measure that seems to place fewer working memory demands on the children than all of our other measures. In addition, the children's enrollment in special schools for children with language impairments required a diagnosis made by professionals prior to the children's participation in this study. Thus, although these children may have had limitations in working memory, they were not clearly different from the more general population of children with LI in having working memory limitations along with problems with language itself.

Hungarian is a language with agreement required between both the subject and the verb and between the verb and the object. According to the agreement deficit account, children with LI should have more difficulty than the VC children in the marking of agreement. To evaluate the predictions of this account, it is important to examine the children's accuracy with regard to tense separately from their accuracy with regard to agreement. As can be seen in Figure 2, the children with LI made a greater number of tense errors than the VC children, but this difference did not achieve statistical significance. On the other hand, one-dimension errors involving past tense implicating past tense difficulty were no less frequent than would be expected if all four dimensions (tense, person, number, and definiteness) were equally vulnerable to error. As would be predicted by this account, agreement errors were clearly evident in the responses of the LI group. Yet, the group difference for number errors was not significant. These errors were relatively infrequent by the LI group. In addition, considering that 24 different inflections were required in our task, all involving agreement of some type, the LI group's mean percentage of correct use of 60% suggests that these children were clearly not producing inflections at random. Furthermore, these children were clearly not relying on a default form when responding to the items. These findings suggest that if the agreement deficit account is generally correct, provisions must be made in the account to explain how children with LI can use all person, number, and definiteness forms with some degree of accuracy, and not differ from VC children in the use of number. In addition, the agreement deficit account provides no reason for the special difficulty with Pl2 forms experienced by the children with LI.

Hungarian differs from languages with a rich inflectional morphology such as Italian and Spanish in that distinctions in four dimensions—tense, person, number, and definiteness—are required rather than the distinctions in three dimensions required in these other languages. According to the morphological richness account, rich inflectional morphology is beneficial to children with LI up to a point; however, four dimensions have been proposed as the number of dimensions that begin to tax these children's limited capacities. For this reason, Hungarian-speaking children with LI are expected to perform below the level of typically developing peers even though their levels of inflection use should be considerably higher than the levels reported for children acquiring English.

The findings were in keeping with this prediction. Furthermore, this account predicts that the inflections with the greatest likelihood of accuracy in the speech of children with LI will be those of higher frequency of occurrence. Our results were also consistent with this expectation.

An additional finding in line with the morphological richness account was the disproportionate number of one-dimension errors relative to errors of two, three, or four dimensions. For the LI group, this finding held true for all 24 target inflections and all 24 inflections used as substitutes. One might argue that even the differences between two-dimension errors and three- and four-dimension errors also support this account, as the likelihood of a substitution was found to decrease as the number of dimensions differing from the target increased. In fact, it is noteworthy that across the 24 target inflections, there were 288 opportunities for a four-dimension error to occur in the data for each child (2 different inflections could have differed from the target by four dimensions, each with 6 items, for each of 24 target inflections, thus $2 \times 6 \times 24 = 288$). Yet, not a single error of this type was seen—a striking finding considering that there were 25 children in each group. The absence of these errors was not due to the children's avoidance of particular inflections. For each child, all 24 inflections had two opportunities (for a total of 12 items) to be used in place of a target that differed by four dimensions, and all of these inflections were used correctly to some degree, and in substitutions in which the inflection replaced the target inflection when it differed on one dimension.

These findings show that even though the children with LI were less proficient than the VC children, their production of inflections—even when in error—reflected some degree of knowledge of the target. This pattern of performance is consistent with an assumption that processing limitations contributed to the children's performance. All inflections were used correctly to some extent, with greater accuracy seen for inflections that occur more frequently in the language, and errors usually approximated the target by differing on relatively few dimensions.

Another prediction of the morphological richness account is that if a substitute inflection differs from the target on two or more dimensions, the substitute should have relatively high frequency of occurrence in the language because only such inflections are assumed to have sufficient strength in the paradigm to alter the tendency for a near miss to be retrieved when an error occurs. The regression analyses confirmed this prediction; log inflection frequency was a significant predictor of the number of times an inflection was a substitute that differed from the target on two or more dimensions. This frequency effect was quite specific. Log frequency of the inflection did not predict the total number of times it was used as a substitute when distance from the target was ignored.

Although the data were consistent with several predictions of the morphological richness account, there are details in the data that this account does not explain in its current formulation. As a case in point, we noted that children with LI produced a greater number of definiteness errors than the VC group but did not differ from the VC group in committing errors involving number. Both definiteness and number require agreement, both have contrasts of two features (definite vs. indefinite, singular vs. plural), and both are crossed with tense and person distinctions in the same way in the sentence stimuli. Therefore, the fact that the LI and VC groups differed in the number of errors on one of these dimensions and not the other suggests that factors beyond the number of dimensions are probably relevant.

Given the gaps that remain in explaining the data, other proposals should be considered and a determination should be made as to whether they might supplement or even fully replace the morphological richness account. For example, Rispoli (1991) noted that transitive verb inflections in Hungarian may be difficult for children because they require a "global case marking" system, given that agreement with both the subject (in person and number) and the object (in definiteness) is necessary. We believe that such global agreement might well increase processing demands, yet the morphological richness account in its current formulation captures this fact only in terms of the number of dimensions that must be considered, not in terms of whether agreement must occur with both the subject and the object. Thus, in the present formulation, the morphological richness account makes no distinction between, for example, the Hebrew verb paradigm that involves four dimensions with all three agreement dimensions (person, number, gender) involved in subject–verb agreement and the Hungarian verb paradigm that involves four dimensions with two of the agreement dimensions (person, number) involved in subject–verb agreement and the third (definiteness) involved in verb–object agreement.

Contributions may also come from work conducted within the framework of other processing-related accounts. For example, in an application of the competition model to Hungarian, MacWhinney and Pléh (1997) noted that adults' interpretations of sentences relied less on definiteness agreement between the verb and the object than on other cues. These investigators suggested that definiteness agreement in Hungarian has relatively low "contrast availability." That is, because in Hungarian both the subject and the object may be definite, or both may be indefinite, definiteness is often noncontrastive and, as a result, adults seem to depend less on this type of cue than on other types of cues. It is possible that factors such as contrast availability influence production as well, and perhaps especially so in the case of children with LI. An application of the competition model to the study of inflection use in children with LI might prove quite informative in this regard.

Along with their well-documented problems in the area of morphosyntax, children with LI often have considerable difficulty retaining sequences of sounds, as measured by tasks such as nonword repetition (see Graf Estes, Evans, & Else-Quest, 2007, for a recent meta-analysis). Although these two deficits are separable (Bishop et al., 2006), many children with LI have both of these deficits. An assumption of the present study is that in a language with a multitude of inflections and allomorphic variations such as Hungarian, children's ability to retain sequences of sounds may have a greater influence on their ability to learn the inflection system than is seen in a language such as English.

Our findings seem consistent with this assumption. The length of the verb with inflection proved related to the children's inflection accuracy even when log inflection frequency was taken into account. More importantly, the very clear differences between the two groups in inflection accuracy were no longer evident when the children's nonword repetition scores were used as a covariate.

Collectively, our findings lend support to the notion that processing-related factors play a role in the inflection limitations of children with LI in a language such as Hungarian. However, it is likely that we have not identified all of the factors related to processing that were at play in this study. Earlier, we noted that factors considered in the competition model such as contrast availability may prove important. In addition, other types of processing factors might be identified. For example, the children sometimes changed the verb or a subject or object in the stimulus sentence. It is true that even when such changes were allowed (provided that the verb inflection was correct), group differences favoring the VC children were seen in inflection accuracy. Nevertheless, it seems important to determine why such substitutions of verbs, subjects, and objects were relatively frequent in the data.

In summary, the findings of this investigation indicate that models assuming processing limitations on the part of children with LI are more compatible with the pattern of verb inflection use seen in Hungarian-speaking children with LI than are accounts based on an assumption of deficits specific to agreement. One processing-related approach, the morphological richness account, seems to predict a substantial portion of the findings, though unexplained gaps remain. Nonmorphosyntactic language processing factors such as the retention of sequences of sounds may well account for additional details in the findings. We suspect that this factor may play a larger than usual role in a language laden with inflections such as Hungarian. Yet, it seems likely that other factors will prove important as well. Additional research is clearly warranted.

Acknowledgments

This research was supported by Grant R01 DC00458 from the National Institute on Deafness and Other Communication

Disorders to Laurence B. Leonard and by Grant OTKA TS 049840 from the Hungarian National Science Foundation to Csaba Pléh. We would like to thank Anna Babarczy, Huba Bartos, and Péter Rebrus for their valuable help and suggestions on the article.

References

Babarczy, A. (2005, May). *Implicit subjects in early child language*. Paper presented at the 7th International Conference on the Structure of Hungarian, Veszprém, Hungary.

Babarczy, A. (2007, April). *Measures of complexity in language acquisition*. Paper presented at the Workshop on Language Complexity as an Evolving Variable, Leipzig, Germany.

Bartos, H. (1997). The nature of object agreement in Hungarian. In *Proceedings of the 21st Annual Penn Linguistics Colloquium* (pp. 19–34). Philadelphia: University of Pennsylvania.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 3–73). Cambridge, England: Cambridge University Press.

Bedore, L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44,* 905–924.

Bishop, D. V. M. (1983). *Test for the Reception of Grammar*. Unpublished manuscript.

Bishop, D. V. M., Adams, C., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: Evidence from 6-year-old twins. *Genes, Brain and Behavior, 5,* 158–169.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics, 10,* 151–171.

Clahsen, H., & Dalalakis, J. (1999). Tense and agreement in Greek SLI: A case study. *Essex Research Reports in Linguistics, 24,* 1–25.

Clahsen, H., & Hansen, D. (1997). The grammatical agreement deficit in specific language impairment: Evidence from therapy experiments. In M. Gopnik (Ed.), *The inheritance and innateness of grammars* (pp. 141–160). Oxford, England: Oxford University Press.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46,* 1029–1037.

Csányi, F. I. (1974). *Peabody Szókincs-Teszt*. Budapest, Hungary: Bárczi Gusztáv Gyógypedagógiai Föiskola.

Dankovics, N., & Pléh, C. (2001). Hangrestaurációs jelenségek és alaktani feldolgozás a magyarban: Azt halljuk-e, amit várunk? [Restoration phenomena and morphological processing in Hungarian: Do we hear what we expect to hear?]. In C. Pléh & Á. Lukács (Eds.), *A magyar morfológia pszicholingvisztikája* [*The psycholinguistics of Hungarian morphology*] (pp. 55–83). Budapest, Hungary: Osiris.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Dromi, E., Leonard, L., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1414–1431.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test–Revised*. Circle Pines, MN: AGS.

Eisenbeiss, S., Bartke, S., & Clahsen, H. (2005). Structural and lexical case in child German: Evidence from language-impaired and typically-developing children. *Language Acquisition, 13,* 3–32.

Gathercole, S., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.

Gathercole, S., Willis, C., Baddeley, A., & Emslie, H. (1994). The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory, 2,* 103–127.

Graf Estes, K., Evans, J., & Else-Quest, N. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50,* 177–195.

Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I., & Trón, V. (2004). Creating open language resources for Hungarian. In *Proceedings of Language Resources and Evaluation Conference,* Lisboa (pp. 203–210). Paris: European Language Resources Association.

Kas, B., & Lukács, Á. (2008). *Magyar Mondatutánmondási Teszt* [Hungarian Sentence Repetition Test]. Unpublished manuscript, Budapest University of Technology and Economics.

Kornai, A., Halácsy, P., Nagy, V., Trón, V., & Varga, D. (2006). Web-based frequency dictionaries for medium density languages. In A. Kilgarriff & M. Baroni (Eds.), *Proceedings of the 2nd International Workshop on Web as Corpus* (pp. 1–9).

Lengyel, Zs. (1981). *A gyereknyelv* [Child language]. Budapest, Hungary: Gondolat.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L., Sabbadini, L., Leonard, J., & Volterra, V. (1987). Specific language impairment in children: A cross-linguistic study. *Brain and Language, 32,* 233–252.

MacWhinney, B. (1985). Hungarian language acquisition as an exemplification of a general model of grammatical development. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition, Volume 2: Theoretical issues* (pp. 1069–1155). Hillsdale, NJ: Erlbaum.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Erlbaum.

MacWhinney, B., & Pléh, C. (1997). Double agreement: Role identification in Hungarian. *Language and Cognitive Processes, 12,* 67–102.

Marton, K., Schwartz, R. G., Farkas, L., & Katsnelson, V. (2006). The effect of sentence length and complexity on working memory performance in Hungarian children with specific language impairment (SLI): A cross-linguistic comparison. *International Journal of Language and Communication Disorders, 41,* 653–673.

Meggyes, S. K. (1971). Egy kétéves gyermek nyelvi rendszere [The language system of a 2-year-old child]. *Nyelvtudományi Értekezések, 71.* Budapest, Hungary: Akadémiai Kiadó.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai [Hungarian tools for examining verbal working memory]. *Magyar Pszichológiai Szemle, 4,* 479–505.

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

Rebrus, P. (2005). Hogyan inflektál a magyar? [How does Hungarian inflect?]. In J. Gervain, K. Kovács, A. Lukács, & M. Racsmány (Eds.), *Az ezerarcú elme. Tanulmányok Pléh Csaba 60. születésnapjára* [A mind with a thousand faces. Papers in honor of Csaba Pleh for his 60th birthday] (pp. 56–69). Budapest, Hungary: Akadémiai Kiadó.

Rice, M., & Wexler, K. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. San Antonio, TX: Psychological Corporation.

Rispoli, M. (1991). The mosaic acquisition of grammatical relations. *Journal of Child Language, 18,* 517–551.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities of defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1001–1004.

Vinkler, Z., & Pléh, C. (1995). A case of a specific language impaired child in Hungarian. In M. Kovačević (Ed.), *Language and language communication barriers* (pp. 131–158). Hrvatska Sveucilisna Naklada, Zagreb: Hrvatska Sveucilisna Naklada.

Warren, R. M. (1970, January 23). Perceptual restorations of missing speech sounds. *Science, 167,* 392–393.

Received August 8, 2007

Revision received December 18, 2007

Accepted June 23, 2008

DOI: 10.1044/1092-4388(2008/07-0183)

Contact author: Ágnes Lukács, who is now with the Department of Cognitive Science, Budapest University of Technology and Economics, H-1111 Budapest, Hungary, Stoczek 2. Ágnes Lukács is also with the Research Institute of Linguistics, Hungarian Academy of Sciences, Budapest, Hungary. E-mail: alukacs@cogsci.bme.hu.

Study 2.

Lukács, Á., Leonard, L. B., Kas, B. (2010) The Use of Noun Morphology by Children with Language Impairment: The Case of Hungarian. *International Journal of Language & Communication Disorders*, 45, 145-161. doi: 10.3109/13682820902781060.

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6)

Thesis 3. Difficulties with multiple suffixation and with morphophonologically irregular forms suggest lexical and processing problems instead of a grammar-specific deficit in SLI (2)

INT. J. LANG. COMM. DIS.
2009, iFirst Article, 1–26

informa
healthcare

Research Report

Use of noun morphology by children with language impairment: the case of Hungarian

Ágnes Lukács†, Laurence B. Leonard‡ and Bence Kas§

†Department of Cognitive Science, Budapest University of Technology, Budapest, Hungary
‡Purdue University, West Lafayette, IN, USA
§HAS Research Institute for Linguistics, Budapest, Hungary

(*Received 5 September 2008; accepted 27 January 2009*)

Abstract

Background: Children with language impairment often exhibit significant difficulty in the use of grammatical morphology. Although English-speaking children with language impairment have special difficulties with verb morphology, noun morphology can also be problematic in languages of a different typology.
Aims: Hungarian is an agglutinating language with multiple suffixation, in which both regular-class and irregular-class nouns contain the same recognizable grammatical markers, but the two classes differ in their morphophonology and productivity. Such typological characteristics provide a good basis for evaluating processing accounts of language impairment such as the morphological richness account.
Methods & Procedures: We examined the production of Hungarian irregular and regular noun morphology through elicited production of nouns with plural, accusative case and plural plus accusative case suffixes in an older (8–10 years) and a younger (4–7 years) group of children with language impairment and two verbal control groups matched on vocabulary size. The children's accuracy was scored both in terms of grammatical function (whether plural and/or accusative case was appropriately marked) and morphophonology (whether the production reflected the phonotactic form required for the stem plus suffix).
Outcomes & Results: The younger children with language impairment were less accurate than the younger verbal control children when two suffixes (marking plural and accusative case) were required, at least when noun stem classes were regular. All groups showed significant overgeneralization of stem forms with correct selection of suffixes. However, there were strong word frequency effects in the language impairment, but not in the verbal control groups.

Address correspondence to: Ágnes Lukács, Department of Cognitive Science, Budapest University of Technology, Stoczek. u. 2., Budapest H-1111, Hungary; e-mail: alukacs@cogsci.bme.hu

International Journal of Language & Communication Disorders
ISSN 1368-2822 print/ISSN 1460-6984 online
http://www.informahealthcare.com
DOI: 10.1080/13682820902781060

Conclusions & Implications: Much of the data were consistent with predictions of the morphological richness account. However, there was also evidence suggestive of differences between the language impairment and verbal control groups in their representations. In particular, the children with language impairment seemed to rely more (though not exclusively) on memorized items in the lexicon.

Keywords: morphology, Hungarian, language impairment, frequency.

What this paper adds

Children with language impairment often exhibit significant difficulty in the use of grammatical morphology. However, relatively little is known about these children's use of noun morphology, especially in agglutinating languages in which there is multiple suffixation. Hungarian provides an opportunity to examine the use of noun morphology by children with language impairment. This language makes use of sequences of suffixes, and the manner in which noun stems are altered to accommodate suffixes permits separate assessment of the children's ability to express grammatical functions (such as accusative case) and their morphophonological ability.

Based on comparisons with typically developing children, we found slight weaknesses in Hungarian-speaking children with language impairment, and differences in the types of representations reflected in their patterns of use. The children with language impairment seemed to rely more (though not exclusively) on memorized items in the lexicon.

Introduction

Children with language impairment (LI) often exhibit significant difficulty in the area of morphosyntax. In many languages, this difficulty includes a weakness in the use of grammatical morphology. English-speaking children with LI often have extraordinary limitations in the use of verb morphology, especially those morphemes that express tense and agreement, and less noteworthy problems with noun morphology (for example, Rice and Wexler 1996, 2001, and Bedore and Leonard 1998). However, this particular profile is likely to be influenced by the typology of English. For example, verb morphology in languages with a rich inflectional morphology — such as Spanish and Italian — is not as troublesome for children with LI as verb morphology in English (for example, Bortolini *et al.* 1997, and Bedore and Leonard 2001). It is likely that the difficulty posed by noun morphology will also vary according to the type of language being acquired. For example, whereas Swedish-speaking children with LI make greater use of past tense inflections than their English-speaking counterparts (Leonard *et al.* 2004), these children have special difficulty with noun phrase morphemes involved in article plus adjective plus noun sequences, where gender, definiteness, and number must be considered (Leonard *et al.* 2001).

Differences in the cross-linguistic profiles of LI have implications for the alternative theoretical accounts that have been proposed to explain this disorder. For example, the extended optional infinitive (EOI) account of Rice and Wexler (for example, Rice and Wexler 1996, 2001, and Rice 2003) and the related extended unique checking constraint (EUCC) account of Wexler (1998, 2003) were designed to explain why children with LI who are acquiring English and other Germanic languages show such a protracted period of using tense and agreement morphology inconsistently. These accounts were not designed to address problems in grammatical morphology that are unrelated to tense and agreement, and were not intended to explain problems in tense/agreement morphology in languages that differ significantly from Germanic languages in typology. In fact, our recent analysis of Hungarian LI tense/agreement use (Lukács *et al.* forthcoming) supports the contention that Hungarian is not a language that shows an extended period of inconsistent use of tense/agreement.

In this paper, we depart from both tense and agreement morphology and from Germanic languages by focussing on noun morphology as used by Hungarian-speaking children with LI. In Hungarian, noun morphology exhibits considerable complexity and differs from Germanic languages in important ways. For an examination of this type, we employ a theoretical framework that places importance on the processing demands involved in language learning.

Morphological richness account

The framework adopted for this study is that of the morphological richness account used by Leonard and his colleagues in their investigations of children with LI who speak inflectionally rich languages such as Italian, Spanish, and Hebrew (for example, Leonard *et al.* 1987, and Leonard 1998: 255–257). This account borrows certain assumptions from the Competition Model (for example, MacWhinney 1987, and Bates and MacWhinney 1989) such as the assumption that languages differ in the details of grammar that have the greatest cue validity, and that children's detection and application of these cues are probabilistic rather than all or none.

According to the morphological richness account, children with LI have a limited processing capacity. This limitation compels these children to devote their limited resources to the most prevalent grammatical information that the language offers. In the case of English — a language in which inflections are sparse and bare stems are frequent — children with LI will devote their limited resources to word order, leaving fewer resources for the learning of grammatical morphology. As a result, a greater number of encounters with each grammatical morpheme will be needed before it is learned adequately. In contrast, children with LI who are learning languages with a rich inflection system are likely to devote their limited resources to this area of grammar. Consequently, differences in inflection use between children with LI and typically developing children in these languages will be smaller than in a language such as English. The label applied to this account — 'morphological richness' — receives its name because of the relative advantage of rich-inflection languages that is assumed.

This relative advantage notwithstanding, the limited processing capacity of children with LI will often prevent them from achieving the level of inflection use seen in their typically developing same-age peers. Inflection errors will be found in their speech. However, these errors will nevertheless reflect considerable knowledge of the grammatical functions of the inflections. According to this account, when children encounter an inflected word, it is stored along with its presumed grammatical functions (for example, first person, singular). Subsequent experiences with the same inflected word that seem to conform to the presumed grammatical functions will strengthen the representation of the inflected word in the child's grammar (Leonard 1989). Inflected words that share meaning and much of the phonetic content but differ slightly in grammatical function (for example, first person, plural) will be stored in the same 'paradigm' as the first inflected word. This inflected word, too, will be strengthened in the child's grammar with repeated encounters. Over time, when a large number of word-specific paradigms have been built, the child will recognize common material across paradigms (for example, that the first person singular inflection is the same across all words), and a general paradigm will be built. This general paradigm can be used to fill in missing cells in any word-specific paradigm that is not yet complete. For example, a child could hazard a guess at the first person singular form of a new word upon hearing only the first person plural form of the word. This process accounts for the productivity (creativity) of children's inflection use.

However, given the limited processing capacity of children with LI, some encounters with inflected words will not be fully processed. This will be more likely in the case of inflections that require consideration of multiple grammatical functions (for example, person, number, tense). Inflected words that are incompletely processed will require more exposures before they are adequately represented in the children's grammars. Until that point, their representations will be relatively weak, rendering them more prone to error. Similarly, the inflections in the general paradigm will be somewhat weak, as these depend on the accumulation of fully processed inflected words. It can be seen, then, that frequency of occurrence is an important factor in this account.

According to the morphological richness account, when errors occur, they will usually be 'near-miss' errors — that is, errors that resemble the correct form in most but not all grammatical functions. For example, a child might produce a first person singular inflection in a context requiring a first person plural inflection. Children are not expected to show haphazard use of inflections or rely on a single default form whenever the correct form is not selected.

These assumptions have been supported by studies on the use of tense/agreement verb morphology by Italian- Hebrew-, and, more recently, Hungarian-speaking children with LI conducted by Leonard and his colleagues (for example, Leonard *et al.* 1987, Dromi *et al.* 1999, and Lukács *et al.* 2009). In these studies, the children with LI did not differ from their typically developing compatriots as dramatically as has been found in English, and the incorrect inflections used by the children were very similar to the correct inflections in their grammatical function. Occasional instances of creative use were also seen, suggesting that the grammars of the children with LI included general paradigms that enabled the children to supply an inflection to a word that was not familiar.

In the present study, we evaluate the assumptions of the morphological richness account using noun morphology in Hungarian as our focal point. As will be seen, this language constitutes an excellent test case, given its particular grammatical properties.

The contribution of Hungarian

Hungarian possesses a rich set of noun suffixes. These suffixes can be attached to the noun stem in a sequence, and many combinations of suffixes can result. For example, the word for 'dogs' in the Hungarian sentence 'The girl is watching the dogs' consists of the stem followed by a plural suffix followed by an accusative case suffix. Absence of the plural suffix signals a singular noun of accusative case, and absence of the accusative case suffix indicates a plural noun of nominative (subject) case. This 'agglutinating' aspect of noun morphology — suffixes appearing in sequence, each marking a different grammatical function — offers an advantage for evaluating accounts that assume processing limitations because a sequence of a stem plus a plural suffix plus an accusative case suffix can be safely regarded as having greater complexity than a stem plus only one of these suffixes.

Another important contribution of Hungarian rests in the manner in which suffixes combine with noun stems. Many noun stems undergo a phonological change when a suffix is added. Based on the type of change required, nouns are divided into regular forms and irregular forms. Regular forms undergo changes based on predictable morphophonological changes. For example, stems that end in a low vowel are lengthened when a plural suffix is added, as in *kutya-kutyák* 'dog-dogs'. Regular forms constitute an open set; new nouns introduced in the language will conform to one of the stem types that undergo these predictable morphophonological changes when a suffix is added. Regular forms have high type frequency (that is, a large number of stems belong to each regular morphophonological type) in the language.

Irregular forms constitute a closed set, with relatively low type frequency. New nouns introduced in the language will not enter this set. In this sense, the regular–irregular distinction is somewhat like that in English. For example, in English, new nouns that are introduced into the language will conform to regular plural inflection use, with the addition of *–s* to the stem. They will not take on an irregular form modelled after existing plural forms such as *men* or *children*. The latter forms constitute a closed set; their list will not be expanded even with the introduction of new words into the lexicon.

However, in other important respects, the regular–irregular distinction in Hungarian is quite different from that of English. First, the suffixes used in regular and irregular forms are essentially the same; the differences rest in the type of change that occurs within the stem when a suffix is added. The changes that occur in regular stems are generalizable to a large number of nouns and any new nouns that enter the language. In contrast, the changes that occur in irregular stems are limited to those (smaller number of) words that already exist in the language. Furthermore, the stems of irregular forms, although changed from their non-suffixed form, retain most of their segments. For example, the (nominative) singular for 'bread' is *kenyér* whereas the (nominative) plural form is *kenyerek*. Here, the second vowel of the stem (*é*) is shortened (to *e*) when the plural suffix is added.

These similarities and differences between regular and irregular forms in Hungarian suggest that factors that ordinarily influence accuracy of use of irregulars in a language such as English may not operate in the same way in Hungarian. For example, in English, we expect to see that children's success with irregular forms such as *feet* and *mice* will be related to the frequency of occurrence of these words, as they cannot be readily broken down into stems and plural suffixes. However, in the case of Hungarian irregular plural forms, where the stems undergo only small phonological changes and the plural suffixes resemble those used in regular forms, stem frequency of occurrence may play as large a role as word frequency in predicting children's success with these forms. That is, the suffix shared by both regular and irregular forms could allow children to readily supply the suffix from the general paradigm to less well known words requiring irregular forms as well as to less well known words requiring regular forms. Their success with the stems of irregular forms, however, may well depend the frequency of occurrence of the stem of the irregular form, not the frequency of occurrence of the entire (stem plus suffix) inflected word.

Predictions

These characteristics of Hungarian lead to the following predictions of the morphological richness account for the use of noun morphology by Hungarian-speaking children with LI. First, given the rich morphology of Hungarian, children with LI in this language should devote many of their processing resources to this area of language and therefore differ from typically developing peers only in those instances in which their limited processing is taxed, as in combinations of a stem plus two suffixes, as seen when a stem is followed by both a plural suffix and an accusative case suffix. Second, when errors occur, they will be best characterized as near misses. In particular, errors on productions requiring a stem plus plural plus accusative case will be more likely to include the stem and either the plural suffix only or the accusative case suffix only. Errors of this type will be more frequent than productions of the stem only. When an error occurs on an item requiring a stem plus plural suffix or a stem plus accusative case suffix, the error will be more likely to involve production of the stem only. Substitutions of a plural for an accusative case suffix or vice versa are not expected, because their grammatical functions are too distinct. Furthermore, children with LI will not rely on a default form, because, although their representations of inflected words are relatively weak and vulnerable to error, they are associated with the proper grammatical function. Finally, because suffixes are essentially the same when used with both regular and irregular stems, they will be successfully extracted and included in general paradigms. This will allow the children to apply these suffixes to words whose stems may not be well established. When paradigms are built for irregular words, inflected words that are high in frequency of occurrence will show the greatest accuracy; those with lower frequency of occurrence will be inflected but may show overgeneralization in the form of a more common stem type replacing a less common stem type. This effect of word frequency is expected at the stage of learning when paradigms are being built; this stage will be longer: (1) for irregulars than for regulars, because of smaller type frequencies for irregulars; and (2) for children with LI, because of limited processing capacities. Because suffixes are transparent even in the case of irregular forms, it is possible that stem frequency rather than frequency of the entire inflected word will be a better predictor of success when paradigms are being built.

Before describing the study itself, we provide a brief sketch of Hungarian noun morphology and review the available evidence from typically developing Hungarian-speaking children.

The system of noun morphology in Hungarian

Hungarian is a non-configurational language with a very rich system of suffixes. Word order is relatively free; for this reason, noun suffixes convey important information about the grammatical function of the noun in the sentence. As noted earlier, suffixes can combine, and many combinations are possible. For example, a noun can appear in 756 different forms, based on the number and types of suffixes applied to the stem. However, the order of suffixes within a word is fixed. Although Hungarian is often categorized as an agglutinating language given the sequences of suffixes that follow the stem, there are details in the morphology that constitute fusions of grammatical dimensions. For example, agreement inflections attached to verb stems simultaneously mark person and number. In the case of noun morphology, the two suffixes of interest in the present study — (nominative) plural and accusative case — function as agglutinating morphemes, in the order stem plus plural plus accusative. Most suffixes have several allomorphs whose forms are dictated by whether the preceding segment has a relative anterior place of articulation or (in some cases) whether the preceding segment is rounded. These are instances of vowel harmony. There are other complex morphophonological patterns governing suffixation; those relevant to the present investigation will be detailed below.

Noun stem allomorphs in Hungarian can be classified into one of three regular classes and one of four irregular classes. Only three of the four irregular classes are examined in this study. The great majority of stems belong to productive regular classes; irregulars form closed classes with small type frequencies. Unlike in English and some other languages, the noun suffix in Hungarian is readily identifiable with all stem types, regular and irregular. For example, the plural form of every noun, whether regular or irregular, ends in *–k*; it is either the stem that alternates, or the quality of the linking vowel or allomorphy in general that changes in irregular forms (for previous research on Hungarian, see Lukács and Pléh 1999). We first describe the three regular classes.[1]

Stems ending in a low vowel

Low vowels (*a*, *e*) lengthen (to *á* and *é*, respectively) before suffixes. For example, the (nominative) singular *kamra* 'chamber' becomes *kamrák* 'chambers' in the (nominative) plural); the (nominative) singular *csésze* 'cup' becomes *csészék* 'cups' in the (nominative) plural. This is a regular morphophonological change that works with all inflectional suffixes.

Stems ending in a consonant

The vowel of suffixes with a linking vowel is manifest in combinations with stems ending in a consonant. For example, for the (nominative) singular *mester* 'master', the (nominative) plural form is *mesterek* 'masters'. The behaviour of the accusative *–t* is more complicated; when the stem ends with a coronal nasal, liquid, or sibilant fricative, it forms a consonantal cluster as in the accusative singular *mestert* 'master'. When the final consonant does not have any of these particular characteristics, a linking vowel is used with the accusative suffix. For example, whereas the (nominative) singular for the word 'basis' is *alap*, the accusative singular form is *alapok*.

Stems ending in a non-low vowel

With these stems, a linking vowel is not used, and the stem-final vowel does not change. For example, the (nominative) singular *hajó* 'ship' has the (nominative) plural form *hajók* 'ships', and the (nominative) singular *cipő* 'shoe' has the accusative singular form *cipőt*.

Each of the three irregular classes is a closed class. Along with the description of each irregular class, we provide the number of stems in the language that enter into the class.

Epenthetic stems

There are 104 stems in Hungarian that fall into this class. These stems end in a –VCVC sequence, as in *bokor* 'bush' and *terem* 'hall'. When plural and accusative suffixes are added to these stems, the final vowel of the stem is eliminated, resulting in a stem-final VCC sequence. Hence, the (nominative) plural and accusative singular forms for *bokor* become *bokrok* and *bokrot*, respectively.

Shortening stems

This is the largest closed class of stems, with 222 of this type in the language. This type of irregular form is found in cases in which the vowel of the final syllable of the stem is long. When plural or accusative suffixes are added, the vowel becomes short. Examples include the (nominative) singular form *madár* 'bird' becoming *madarak* in the plural and the (nominative) singular *kenyér* 'bread' becoming *kenyeret* in the accusative singular. Shortening stems require a low vowel as a linking vowel, hence *–ak* and *–et* rather than *–ok* and *–ot*, respectively, in these examples.

v-inserting stems

This small class contains seven one-syllable stems ending in a long vowel. When plural or accusative suffixes are attached to these stems, the voiced labiodental fricative /v/ is inserted between the stem final and the linking vowels, and at the same time, the vowel of the stem shortens. Thus, *ló* 'horse' becomes *lovak* 'horses' in the (nominative) plural, and *kő* 'stone' becomes *követ* in the accusative singular.

Data from earlier studies of typical acquisition

The plural and accusative case suffixes are among the first to emerge in the speech of typically developing children, although when they first appear at approximately 18 month of age, they form an unanalysed unit with the stem. Despite the early appearance of these suffixes, the full system of noun allomorphy is a relatively late grammatical attainment. In fact, the irregular classes are not fully mastered until the early school years. According to MacWhinney's (1975, 1978) results from elicited production, children show at least 90% accuracy on regular classes by three years of age, with fewer errors when the stem ends in a vowel (regardless of whether vowel lengthening was required) than when the stem ends in a consonant. Among the irregular classes, /v/-inserting is most accurate (approximately 80% correct), and epenthetic stems are least accurate (approximately 67%) at seven years of age.

A later study of a larger group of children between four and eight years of age by Pléh *et al.* (2002) largely confirmed MacWhinney's (1975, 1978) findings. Suffixes and stems from the regular classes were used with accuracy at the earliest age studied. In addition, over-regularizations of the irregular classes of /v/-inserting and epenthetic stems were seen until eight years of age. However, there were slight deviations from MacWhinney's results in the specific order of acquisition of irregulars. Shortening stems proved to be easier than epenthetic stems, and /v/-inserting stems were most difficult. These results matched the number of stems in the language that belong to these irregular classes (shortening=222 stems, epenthetic stems=104 stems, /v/-inserting=eight stems).

In their study, Pléh et al. (2002) included an examination of the children's use of plural, accusative, and plural plus accusative forms, although they did not systematically test all three forms with the same nouns. They found that the combination of two suffixes was more problematic for the children relative to single suffixes of either type. However, this finding applied primarily to stems from the regular classes. Because the same nouns were not tested across all suffixes, more research is needed to clarify the interaction of multiple suffixation and stem types (for a summary of previous research on acquisition, see Pléh 2000; on adult processing of regular–irregular inflections in Hungarian, see Lukács and Pléh 1999, Lukács 2001, and Pléh and Lukács 2002).

One remaining study of typically developing children's use of multiple suffixes comes from Pléh *et al.* (1997). These investigators examined the spontaneous speech of children between age 1;7 and 2;4, using the Hungarian CHILDES corpus. Their major focus was on locative case suffixes. Suffixes of this type follow plural suffixes. Of the 615 instances of locative case suffixes, 65 were preceded by another suffix in a multiple suffix sequence. Only two of the preceding suffixes were plural suffixes. The remaining instances were sequences of possessive plus locative case suffixes.

To our knowledge, studies of noun suffix use by Hungarian-speaking children with LI have not yet been reported. We turn to a description of the present study on this topic, with an eye toward whether these children's use of noun suffixes conform to predictions based on the morphological richness account.

Method

Participants

A total of 60 children participated in the investigation. Thirty of the children exhibited a language impairment (LI) and 30 were developing language in a typical manner. Fifteen of the children with LI were recruited for a younger group of children with LI; these children ranged in age from 4;10 to 7;2 (mean=6;0). Another 15 children with LI were recruited for an older group of children with LI. These older children ranged in age from 7;10 to 9;10 (mean=9;0). The younger children with LI were selected from two special kindergarten classes for children with language impairment. The older children with LI were selected from two special schools for children with language impairment.[2] All 30 of these children met the criteria for LI. Each child scored above 85 on the Raven Coloured Progressive Matrices (Raven *et al.* 1987), a measure of non-verbal intelligence, passed a hearing screening, and had no history of neurological impairment. Each of these children scored at least 1.5 standard deviations (SDs) below the mean for his or her age on two or more of four language tests administered. Two of the four tests assessed receptive skills, and two evaluated expressive skills. The receptive tests were the

Hungarian standardizations of the Peabody Picture Vocabulary Test (PPVT) and the Test for Reception of Grammar (TROG). The expressive tests were the Hungarian Sentence Repetition Test, and a non-word repetition test.

The Hungarian adaptation and standardization of the PPVT is modelled closely after its English equivalent (Csányi 1974, Dunn and Dunn 1981). The Hungarian adaptation of the original TROG (Bishop 1983) is in the process of being standardized for the age range 4–12 years.[3] Items assess the children's understanding of progressively more difficult grammatical forms. The test consists of 20 blocks, each with four items of the same grammatical construction (for example, sentences with comparatives, post-modified subjects and embedded clauses). The child must point to a picture (from an array of four pictures) that matches the sentence spoken by the experimenter. A block is scored as complete if the child responds correctly to all four pictures in the block. The total score is the number of blocks correctly completed.

The Hungarian Sentence Repetition Test (*Magyar Mondatutánmondási Teszt* (Kas and Lukács 2008) contains 40 sentences, distributed evenly across five types of grammatical constructions. These are simple subject–verb–object (SVO) sentences, simple OVS sentences, and complex sentences containing SS relative clauses, SO relative clauses, and OS relative clauses. The sentences vary in length from eight to 15 syllables within each type. The child is asked to immediately and accurately repeat each sentences presented by the experimenter. Accuracy is measured in terms of the number of correctly repeated sentences.

The non-word repetition test (Racsmány *et al.* 2005) consists of 36 non-words ranging from one to nine syllables in length. Four non-words are used for each length. All non-words conform to the phonotactic patterns of Hungarian. The particular phonological sequences used in the non-words do not reflect the frequency distribution of Hungarian phoneme sequences. However, sequences that are articulatorily difficult for speakers are not employed. The child is asked to repeat each non-word. The child's score is defined as the length at which the child successfully repeated at least two of the four non-words.

The 30 typically developing children were selected to serve as vocabulary-score matches based on their raw scores on the PPVT. Hereafter, these children are referred as the vocabulary control (VC) children. Fifteen of the children (younger VC group) ranged in age from 3;3 to 6;10 (mean=5;3) and were matched according to PPVT scores with the younger LI group. The remaining 15 children (older VC group) ranged in age from 4;5 to 7;10 (mean=6;4) and were matched according to PPVT scores with the older LI group. (It can be seen that these two groups of VC children overlapped in age; for this reason the designation 'younger' and 'older' should be read as 'matched to the younger' and 'matched to the older', respectively.) A typically developing child was considered a match if his or her PPVT score was within 4 points of the PPVT score of a child in the LI group.

Because the children with LI scored below age level on the PPVT, the typically developing children matched on this measure were younger. This age difference between the children with LI and VC children was much larger for the older LI and VC groups than for the younger LI and VC groups. We are not certain of the reasons for the larger age difference in the older groups. One possibility is that the LI and VC groups differ in the growth rate of their vocabulary development. For instance, during the primary school years, children with LI might show a linear growth rate, whereas typically developing children might show inverse exponential growth. We also cannot rule out the unintentional selection of children with more severe language impairments in the older LI group, although we have no basis for assuming this occurred. However, even the younger LI and VC groups differed in age by an average of 9 months. The use of younger typically developing children matched on a non-grammatical measure (such as the PPVT) was designed to detect whether the difficulties of the children with LI on regular and irregular noun morphology exceeded their limitations in vocabulary size. If so, group differences favouring the VC group should be seen.

All of the children in the VC groups scored above –1 SD on each of the four language tests that were administered to the children with LI. Means and ranges for age, PPVT raw score, and scores on the TROG and non-word repetition are provided in table 1.

Procedure

The items used by Pléh et al. (2002) were employed here, supplemented by additional items to ensure an equal number of each of the regular and irregular suffixes of interest. There were 24 items that assessed the children's use of the (nominative) plural suffix, 24 items that evaluated their use of the accusative singular suffix, and 24 items that examined the children's use of the plural plus accusative suffix. The same 24 noun stems were used for all three item types. For each of these three suffix types, four items were used for each of the three regular stem classes and four items were employed for each of the three irregular stem classes. An elicited production task with pictures was used. The items were arranged in 24 sets, with each set designed to assess a (nominative) plural suffix, an accusative case suffix and a plural plus accusative case suffix with the same stem. For each set, the experimenter showed the children a picture of an object and provided the name of the object, using the nominative singular form that constitutes the appropriate form for a labelling context. The children were then shown a second picture and were asked questions that prompted one of the suffix types under investigation. For items designed to assess the (nominative) plural suffix, the question (with corresponding picture) was *Mik ezek?* 'What are these?' For items that assess the accusative singular, the question took the form *Mit néz a fiú?* 'What is the boy watching?' The question *Miket néz a fiú?* 'What (plural) is the boy watching?' was used to test the children's use of the plural plus accusative. In Hungarian, the interrogative pronoun is always case-marked, and, when referring to a plural referent requiring accusative case, is also in the plural. Table 2 provides examples of items for each suffix type and regular and irregular stem class.

Table 1. Mean ages, Peabody Picture Vocabulary Test (PPVT), Test for Reception of Grammar (TROG), and non-word repetition scores (with ranges in parentheses) of the older and younger language impairment (LI) and vocabulary control (VC) groups

	Age	PPVT	TROG	Non-word repetition
LI older	9;0 (7;10–9;10)	88.73 (67–114)	68.67 (59–75)	3 (0–5)
LI younger	6;0 (4.;10–7;2)	80.5 (45–110)	63.93 (48–75)	2.87 (1–5)
VC older	6;4 (4;5–7;10)	89.67 (68–111)	70.33 (56–78)	5.73 (3–7)
VC younger	5;3 (3;3–6;10)	79.8 (49–114)	68.4 (52–80)	5.27 (4–7)

Table 2. Examples of stimuli for each stem class

	Free stem	Accusative	Plural	Plural plus accusative	
1. Epenthetic (*n*=104)	*majom*	*majm**ot***	*majm**ok***	*majm**okat***	monkey
2. Shortening (*n*=222)	*kenyér*	*kenyer**et***	*kenyer**ek***	*kenyer**eket***	bread
3. v-Insertion (*n*=8)	*kő*	*köv**et***	*köv**ek***	*köv**eket***	stone
4. 'Low V'-final	*kutya*	*kutyá**t***	*kutyá**k***	*kutyá**kat***	dog
5. C-final	*asztal*	*asztal**t***	*asztal**ok***	*asztal**okat***	table
6. 'Non-low V'-final	*cipő*	*cipő**t***	*cipő**k***	*cipő**ket***	shoe

Note: The first three stem classes are closed sets and thus described as 'irregular'; the number of stems in the language belonging to each of these sets is also provided. The remaining stem classes are open sets and are described as 'regular'.

Scoring

We categorized errors into several different error types: (1) overgeneralization — that is, the use of a suffix with the base form of the stem (nominative singular) instead of the modified bound form; (2) unmarked form — that is, use of the stem only, the form used for nominative singular nouns; (3) use of a non-target noun constituting a different stem type; (4) simplification of a plural plus accusative case item to a plural suffix only; and (5) simplification of a plural plus accusative case item to an accusative case suffix only. If any of these patterns occurred, the response was considered to be an error. We also performed analyses using an alternative scoring method. In the alternative method, we scored overgeneralizations as if they were correct. We reasoned that this alternative view would provide us with an indication of the children's ability to use the appropriate suffixes even when their command of the alterations needed for the stem was somewhat limited. This alternative method of scoring can be considered a test of the children's ability with the morphosyntactic requirements of the task, apart from morphophonological demands.

Data analysis

The data were examined separately for the older and the younger LI-VC groups, and separately for the regular and irregular stem classes given that these differed in whether they were open or closed classes. We examined the children's percentages of correct responses using a general linear model analysis of variance (ANOVA) with Group as a between-subjects factor and Suffix (plural, accusative, plural plus accusative) as a within-subjects factor. We also tested whether the children's scores were related to frequency of occurrence. For each inflected form, we calculated both the inflected word frequency (the frequency of the exact inflected noun form), the stem frequency, and the stem allomorph frequency. The stem frequency was defined as the free stem form, that is, the form as it appears in nominative singular plus the frequency of the bound stem form(s) when the suffix is added. For example, for 'monkey', the stem frequency is the sum of the frequency of the free stem form *majom* plus the frequency of all related bound stem forms of the same noun, such as *majm-* (table 2). The stem allomorph frequency was the frequency of the bound stem form itself that was required when a suffix was added (for example, the frequency of *majm-*).[4] Given that the forms examined in this study always involved a bound stem, it is plausible that the frequency of this bound stem (allomorph) had a greater effect on the children's pattern of use than the other frequency measures. The source of frequency data was the Hungarian Webcorpus (Halácsy *et al.* 2004, Kornai *et al.* 2006). Calculations employed the logarithm of frequency. We also performed an analysis of the children's errors. In a final series of analyses, we compared the groups of children on their accuracy of suffix use after rescoring overgeneralizations of stems and treating them as correct responses.

Younger LI and VC groups

Regular stem classes

A summary of the findings for accuracy is illustrated in figure 1. A significant difference according to Group was seen, $F(1, 28)=4.76$, $\eta^2=0.145$, $p<0.05$, indicating that the younger VC group produced the items with greater accuracy than the children with LI. A significant difference was also seen for Suffix, $F(2, 56)=7.843$, $\eta^2=0.219$, $p<0.001$. Bonferroni-corrected pairwise comparisons revealed that the children were significantly less accurate on plural plus accusative items than on (nominative) plural items, but performance on accusative singular items did not differ from either. However, these results should be interpreted within the context of the significant Group × Suffix interaction that was also observed, $F(2, 56)=3.279$, $\eta^2=0.105$, $p<0.05$. This interaction was seen because, whereas the younger LI group used plural plus accusative forms with less accuracy than accusative forms, the younger VC group showed no differences in this regard. In addition, the difference between the younger LI group and younger VC group (favouring the latter) was much larger for plural plus accusative forms than for the other suffixes.

The errors made by the younger LI and VC groups are summarized in table 3. For illustration purposes, the table provides the total number of occurrences for each error type. However, statistical analyses used comparisons of means by *t*-tests.[5] For plural suffix items, the most frequent error by the younger children with LI was the production of the free form of the stem (the form used in nominative singular) as in *asztal* 'table' for *asztalok*. These children also produced three overgeneralization

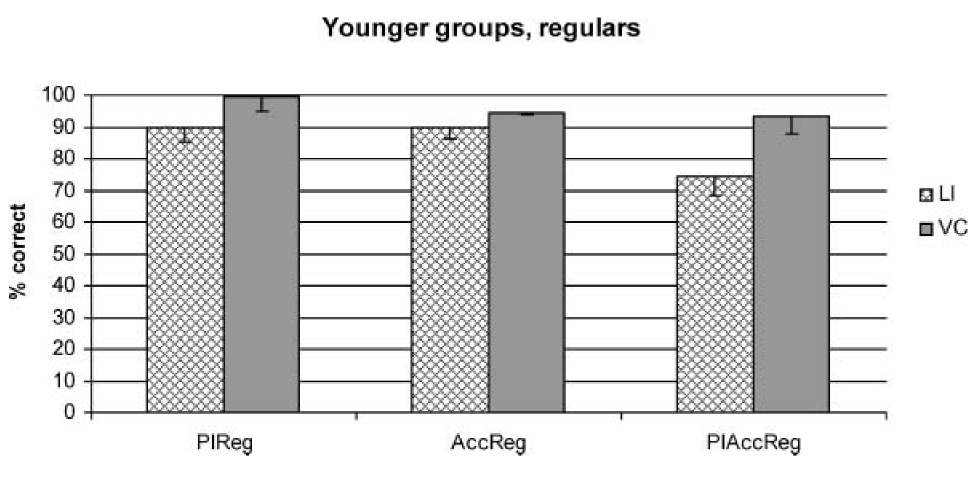


Figure 1. Percentage correct on plural suffixes, accusative case suffixes, and plural plus accusative case suffixes with verbs from the regular stem classes by the younger language impairment (LI) and vocabulary control (VC) groups.

Table 3. Total frequency of each error type by the younger language impairment (LI) and vocabulary control (VC) groups on the items involving regular stem classes

	Overgeneralization		Unmarked		Simplification to plural		Simplification to accusative		Noun Sub	
Suffix	LI	VC	LI	VC	LI	VC	LI	VC	LI	VC
Plural	3	0	10	0	–	–	–	–	0	0
Accusative	6	7	9	2	–	–	–	–	0	1
Plural plus accusative	8	1	3	0	7	1	10	2	1	0

errors. Recall that even though these are referred to as regular stems, vowel length is altered for certain stem classes. An example of one such overgeneralization is *kutyak* instead of *kutyák* where the nominative singular form is *kutya*. The younger VC children made no errors on plural suffix items.

Errors on accusative singular items did not distinguish the younger LI and VC groups as sharply, though the children with LI produced a larger number of stems in these accusative case contexts. For items requiring plural plus accusative suffixes, the LI group not only produced more overgeneralizations and stems than the VC group but also showed a greater tendency to simplify the production to either a plural suffix only or an accusative suffix only. Importantly the LI group was also more likely to use a single suffix in place of a plural plus accusative case suffix than to produce a bare (unmarked) stem, consistent with the expected tendency for near-miss errors, $t(14)=2.17$, $p<0.05$, one-tailed).

Irregular stem classes

Figure 2 provides an illustration of the findings for the younger LI and VC groups' accuracy with the suffixes applied to irregular stem classes. In contrast to the findings for regular stem classes the younger LI and VC groups did not differ in

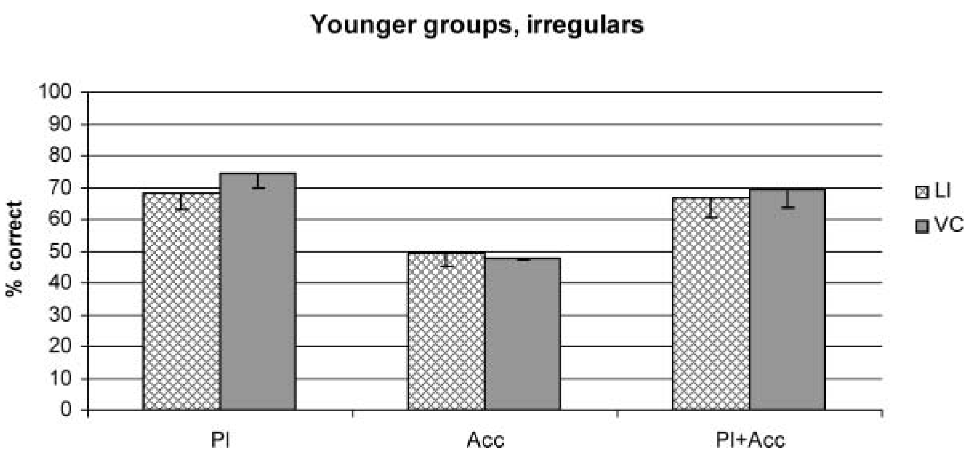


Figure 2. Percentage correct on plural suffixes, accusative case suffixes, and plural plus accusative case suffixes with verbs from the irregular stem classes by the younger language impairment (LI) and vocabulary control (VC) groups.

their accuracy on the items involving irregular stem classes, $F(1, 28)=0.106$, $\eta^2=0.004$, n.s. However, a significant difference was seen for Suffix, $F(2, 56)=31.941$, $\eta^2=0.533$, $p<0.001$. Bonferroni-corrected pairwise comparisons indicated that the accusative case items were used with significantly less accuracy than either the plural items or the plural plus accusative case items ($p<0.001$ in both cases), while the latter two did not differ. The Group × Suffix interaction was not significant, $F(2, 56)=0.803$, $\eta^2=0.028$, n.s.

The two younger groups' errors on irregular stem class items are summarized in table 4. On plural suffix items, the two groups are most distinguishable in the LI group's greater use of stems. This tendency for greater use of stems by the LI group was also seen for accusative case items and for plural plus accusative case items. Interestingly, when children simplified a plural plus accusative case item, they produced an accusative suffix only, never a plural suffix. This type of simplification was more frequent in the data of the children with LI. This pattern is especially noteworthy given that accusative case items, when serving as the target, were produced with lower accuracy than the other suffix types. As was seen for the data for regular stem class items, the LI group was more likely to respond to plural plus accusative case items with the production of a single suffix than to produce a bare stem (although this difference only approached significance, $t(14)=1.35$, $p=0.09$, one-tailed). Finally, it should be noted that both groups of children produced a large number of overgeneralizations. For both groups, overgeneralizations were more common than productions of unmarked stems, $t(14)=6.80$, $p<0.001$ for the VC group, $t(14)=4.90$ for the LI group, $p<0.001$. The two groups did not differ statistically in this respect. Examples are productions for 'monkey/monkeys' such as *majomok*, *majomot*, and *majomokat* for *majmok*, *majmot*, and *majmokat*, respectively. The nominative singular (free stem) form is *majom*.

Older LI and VC groups

Regular stem classes

Figure 3 provides a summary of the older groups' accuracy with suffixes used with regular stem classes. The main effect for Group approached significance, $F(1, 28)=3.390$, $\eta^2=0.108$, $p=0.076$, with higher accuracy achieved by the older VC group. The main effect for Suffix was not significant, $F(2, 56)=1.364$, $\eta^2=0.046$, n.s. The Group × Suffix interaction approached significance, $F(2, 56)=2.563$, $\eta^2=0.084$, $p=0.086$, attributable primarily to the larger difference between the

Table 4. Total frequency of each error type by the younger language impairment (LI) and vocabulary control (VC) groups on the items involving irregular (closed set) stem classes

	Overgeneralization		Unmarked		Simplification to plural		Simplification to accusative		Noun Sub	
Suffix	LI	VC	LI	VC	LI	VC	LI	VC	LI	VC
Plural	39	37	14	5	–	–	–	–	1	1
Accusative	78	86	6	2	–	–	–	–	2	4
Plural plus accusative	34	44	6	2	0	0	16	4	2	3

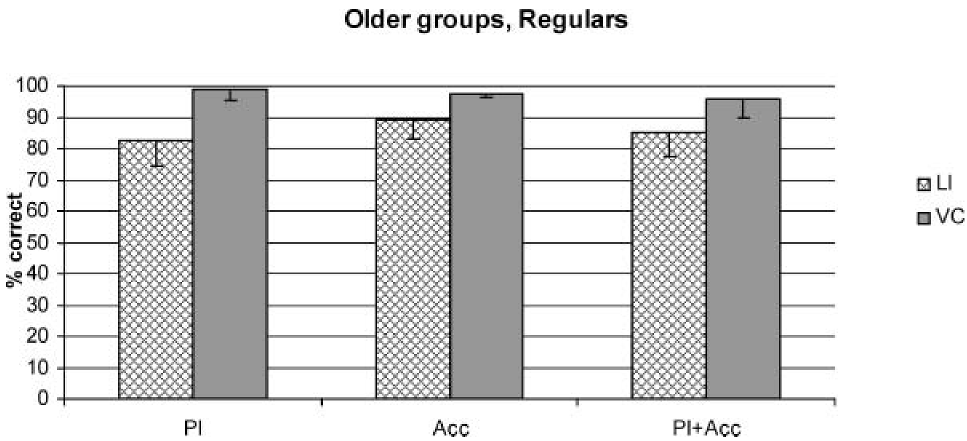


Figure 3. Percentage correct on plural suffixes, accusative case suffixes, and plural plus accusative case suffixes with verbs from the regular stem classes by the older language impairment (LI) and vocabulary control (VC) groups.

older LI and VC groups on their accuracy with plural suffixes than with accusative suffixes.

A summary of the older LI and VC groups' errors can be found in table 5. The clearest differences between the two groups is the greater tendency on the part of the children with LI to produce stems rather than a form with a suffix. They were also somewhat more likely to simplify an item requiring a plural plus accusative form with a form containing a plural suffix only. In addition, the LI group was more likely to produce a different noun than the one expected for the item.

Irregular stem classes

Figure 4 provides an illustration of the older LI and VC groups' accuracy on items requiring suffixes with irregular stem classes. The main effect for Group was not significant, $F(1, 28)=0.023$, $\eta^2=0.001$, n.s. However, the main effect for Suffix was highly significant, $F(2, 56)=11.991$, $\eta^2=0.300$, $p<.001$. Bonferroni-corrected pairwise comparisons indicated that the children were significantly less accurate with accusative case suffixes than with either plural suffixes or plural plus accusative case suffixes ($p<0.01$ in both cases), which, in turn, did not differ. However, a significant

Table 5. Total frequency of each error type by the older language impairment (LI) and vocabulary control (VC) groups on the items involving regular stem classes

	Overgeneralization		Unmarked		Simplification to plural		Simplification to accusative		Noun Sub	
Suffix	LI	VC	LI	VC	LI	VC	LI	VC	LI	VC
Plural	1	0	14	1	–	–	–	–	4	0
Accusative	3	5	10	0	–	–	–	–	4	0
Plural plus accusative	4	2	3	0	5	1	2	1	1	0

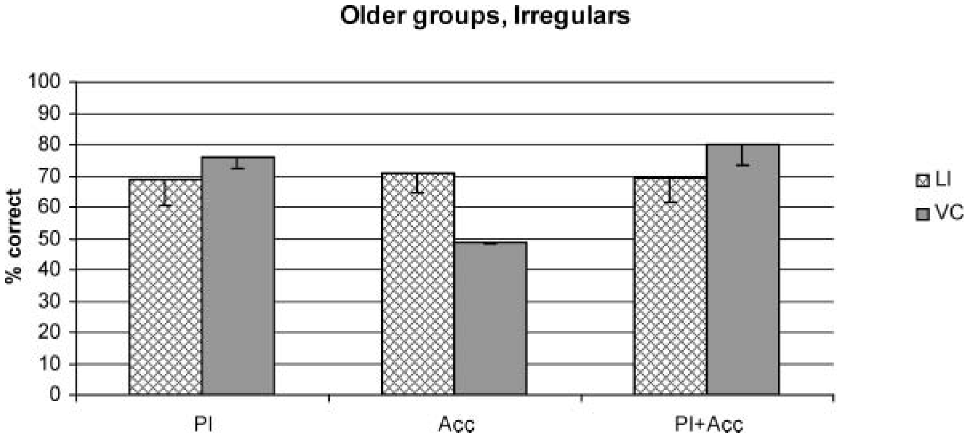


Figure 4. Percentage correct on plural suffixes, accusative case suffixes, and plural plus accusative case suffixes with verbs from the irregular stem classes by the older language impairment (LI) and vocabulary control (VC) groups.

Group × Suffix interaction, $F(2, 56)=14.362$, $\eta^2=0.339$, $p<.001$, indicates that the main effect for Suffix was driven principally by the fact that the older VC children had extraordinary difficulty with accusative case suffix items. In fact, the older VC children were significantly less accurate than the older LI group on this suffix type. In contrast to the older VC children, the accuracy of older children with LI was more uniform across suffix types, showing approximately 70% accuracy, on average, for each type.

Table 6 provides a summary of the older children's errors on items involving irregular stem classes. As was found for the younger children, both the older LI and older VC groups produced a large number of overgeneralizations, for all suffix types. However, an inspection of table 6 provides an answer to why the older VC children were less accurate than the older children with LI on accusative forms; the older VC produced many more overgeneralizations on accusative case items (87) than the LI group (37), $t(28)=3.21$, $p<0.01$, and they produced many more overgeneralizations on these items than on items that assessed the other suffix types. Again, the LI group was more likely to produce a single suffix in contexts requiring a plural plus accusative case suffix than to produce a bare stem, as predicted, $t(14)=1.95$, $p<0.05$, one-tailed.

Table 6. Total frequency of each error type by the older language impairment (LI) and vocabulary control (VC) groups on the items involving irregular (closed set) stem classes

	Overgeneralization		Unmarked		Simplification to plural		Simplification to accusative		Noun Sub	
Suffix	LI	VC	LI	VC	LI	VC	LI	VC	LI	VC
Plural	37	32	11	5	–	–	–	–	4	3
Accusative	37	87	11	1	–	–	–	–	3	3
Plural plus accusative	40	29	5	1	0	0	13	4	4	1

Frequency of occurrence as a predictor of accuracy

Although we assume that the grammatical function of the suffixes and their additive (agglutinating) nature played an important role in the children's production accuracy, it was also important to determine whether frequency of occurrence also played a role. Three types of frequency of occurrence measures were considered. The first was frequency of occurrence of the entire word, that is, the stem plus suffix(es) (for example, *bokrot* 'bush' accusative singular). The potential importance of frequency in the case of irregular forms is obvious, as they constitute a closed set. To use an analogy with English, children are likely to learn that *children* is the plural for *child* much sooner than they learn that *oxen* is the plural for *ox*. However, word frequency could also be a contributing factor in children's accuracy with regular forms.

The second frequency measure considered was the stem frequency of occurrence, defined as the frequency of the sum of all forms containing the free stem form (for example, *bokor-*) plus all bound stem forms (for example, *bokr-*). Because regular forms are usually viewed as being made up of a stem plus an affix, it is reasonable to assume that the frequency of the stem itself might be predictive of children's ability to apply suffixes to it. However, stem frequency is also relevant to irregular forms, given their composition in Hungarian. Recall that irregular forms make use of the same suffixes employed in regular forms. They are 'irregular' in large part because the manner in which their stems are phonetically altered when suffixes are added is limited to a closed set of nouns. It is not clear that the learning mechanisms involved in learning irregular forms, then, is qualitatively different from those involved in learning regular forms. Both regular and irregular forms may be treated as a stem plus suffix(es). The fact that phonetic modifications must be made to the stem relative to the nominative singular form may not be sufficient for the children to treat irregular forms differently from regular forms given that, certain regular forms, too, have stems that must be phonetically modified when a suffix is added (for example, *kutya* 'dog' becomes *kutyák*, *kutyát*, and *kutyákat*), although, unlike in irregulars, these changes are fully predictable from the form of the word. If this assumption is correct, stem frequency of occurrence could constitute a predictor of children's accuracy with irregular as well as regular forms.

The third frequency measure was stem allomorph frequency (for example, the frequency of the form *bokr-*). This measure was employed in case frequency effects were limited to the precise stem allomorph to which a suffix had to be attached.

To determine if these factors could predict performance on the experimental task, we included them in stepwise regression analyses. Specifically, we tested the effects of log word frequency, log stem frequency, and log stem allomorph frequency on the number of correct responses on regular forms and, separately, irregular forms for each of the four groups of children. Only variables that showed a significant correlation ($p<0.05$) with the target variable were entered into the analysis. Results are shown in tables 7 and 8.

The best predictor of using suffixes accurately with regular stem classes was word frequency. However, this proved true only for the younger LI group. No factor proved significant for the remaining three groups. In addition, word frequency accounted for only 17% of the variance for the younger LI group. These findings suggest that, except for the youngest group of children with LI, the children had reached a point in their use of plural, accusative, and plural plus accusative suffixes that the frequency (word, stem, or stem allomorph) with which an item occurred in

Table 7. Best models for predicting the accuracy of use of suffixes with regular stem classes by the younger and older children with language impairment (LI) and vocabulary control (VC) children

		Beta	*p*	R^2
LI older			n.s.	
LI younger	Word frequency	0.41	< 0.05	0.168
VC older			n.s.	
VC younger			n.s.	

Note: n.s., Not significant.

Table 8. Best models for predicting the accuracy of use of suffixes with irregular stem classes by the younger and older children with language impairment (LI) and vocabulary control (VC) children

		Beta	*p*	R^2
LI older	Word frequency	0.66	< 0.001	0.43
LI younger	Word frequency	0.61	< 0.001	0.35
VC older		n.s.		
VC younger	Stem frequency	0.48	< 0.01	0.23

Note: n.s., Not significant.

the language was no longer playing a major role in their success with regular stem classes.

However, from table 8 it is clear that frequency played a more important role with regard to nouns from the irregular stem classes. Word frequency was a predictor for both the younger and older LI groups, explaining 43% and 35% of the variance, respectively. Stem frequency proved to be a stronger predictor for the younger VC group, and no frequency measure was a predictor for the older VC group. Frequency measures accounted for much less variance in the VC groups than in the LI groups. In the younger VC group, if the stem was sufficiently frequent, the children apparently could identify it as part of a closed set, modify the stem accordingly and then add the appropriate suffix(es).

Overgeneralizations of stems as correct responses

To gain a clearer view of the children's accuracy in the use of suffixes whether or not the stem was appropriately formed, we conducted ANOVAs after treating all overgeneralizations of stems as correct responses (provided, of course, that the suffixes were correct). Recall that even some of the regular stem classes show some modification when a suffix is added (for example, *kutya* 'dog' in accusative singular is *kutyát* not *kutyat*). Thus, although the closed set, termed "irregulars", always undergo stem modification when a suffix is added, it was necessary to conduct ANOVAs for regular as well as irregular stem classes using this alternative scoring method.

The ANOVA for the younger LI and VC groups' suffix use with regular stem classes revealed a significant main effect for Group, $F(1, 28)=4.300$, $\eta^2=0.134$, $p<0.05$, indicating that the younger VC group produced the items with greater accuracy than the children with LI. A significant difference was also seen for Suffix, $F(2, 56)=5.560$, $\eta^2=0.166$, $p<0.01$. Bonferroni-corrected pairwise comparisons revealed a difference approaching significance between plural plus accusative items and (nominative) plural items ($p=0.053$) and accusative singular items ($p=0.074$), with weaker performance on multiple suffixation. The Group × Suffix interaction was not significant, $F(2, 56)=1.47$, $\eta^2=0.050$, n.s.

No differences were observed for the younger LI and VC groups' suffix use with irregular stem classes. The non-significant finding for Group, $F(1, 28)=1.630$, $\eta^2=0.055$, n.s., mirrored the previously reported findings when overgeneralizations were scored as errors. However, in this alternative analysis, Suffix also proved non-significant, $F(2, 56)=1.570$, $\eta^2=0.053$, n.s., whereas this factor was significant in the earlier analysis. The difference in the two analyses for Suffix can be easily explained. In the earlier analysis, accusative case items were used with significantly less accuracy than either plural items or plural plus accusative case items. However, accusative case items were also produced with many more overgeneralizations than the other item types,[6] a fact that led to the finding of non-significance in the analysis that treated overgeneralizations as correct. Finally, there was no significant Group × Suffix interaction, $F(2, 56)=0.615$, $\eta^2=0.021$, n.s.

In the alternative analysis of the older groups' accuracy with suffixes used with regular stem classes, the main effect for Group approached significance, $F(1, 28)=3.861$, $\eta^2=0.121$, $p=0.058$, with higher accuracy achieved by the older VC group. The main effect for Suffix was significant, $F(2, 56)=3.89$, $\eta^2=0.122$, $p<.05$. Bonferroni-corrected pairwise comparisons indicated that the scores for plural suffixes were lower than those for accusative case suffixes ($p<0.01$). The Group × Suffix interaction was not significant, $F(2, 56)=1.910$, $\eta^2=0.064$, n.s.

Finally, the alternative analysis of the older children's accuracy with suffixes on items with irregular stem classes showed no significant main effects or interactions. Specifically, Group was non-significant, $F(1, 28)=0.740$, $\eta^2=0.026$, n.s., as was Suffix, $F(2, 56)=1.11$, $\eta^2=0.038$, n.s. and the Group × Suffix interaction, $F(2, 56)=0.137$, $\eta^2=0.003$, n.s. The lack of an interaction contrasts with the earlier analysis, but an inspection of table 6 reveals the source of this difference in findings. The older VC children produced many more overgeneralizations in accusative case items than did the children with LI. In this alternative analysis, overgeneralizations were counted as correct, thus reducing the group differences for this suffix type.

Discussion

Predictions of the morphological richness account

The data obtained in this study allowed us to test several predictions based on the morphological richness account. According to this account, given the rich morphology of Hungarian, children with LI should devote many of their processing resources to this area of language and therefore differ from typically developing peers only in those instances in which their limited processing is taxed. Specifically, they are more likely to score below the level of peers when two suffixes must be applied to the stem. When errors occur on items of this type, they should be near misses, as when the children produce a stem plus plural suffix only or a stem plus accusative case suffix only in place of a stem plus plural plus accusative case suffix. Substitutions involving quite different grammatical functions should not occur. Similarly, the children's partial knowledge of the forms conforming to the different grammatical functions will rule out their reliance on a default form. Suffixes are expected to be applied to stems even when the stem is not well represented in the child's paradigm; in such instances, overgeneralizations of stems are expected. The inflected words that are used correctly are more likely to have either higher word frequency or higher stem frequency.

The comparison between the younger LI and VC groups in the use of regular forms were consistent with the first prediction. The younger children with LI showed lower accuracy with items requiring stem plus plural plus accusative case suffixes than with the other items, and these children showed lower accuracy on these suffix combinations than did the younger VC children. However, for the older children, the difference between the LI and VC groups in their accuracy with regular forms only approached significance (main effect, $p=0.076$; interaction, $p=0.086$). It is not clear if the latter finding reflects a narrowing of group differences across development in the use of these particular suffixes, or in insufficient power in this study to detect very real but smaller group differences at the older ages examined here. The children's use of irregular forms revealed no group differences, in part because all children were generally less accurate with these forms. It is quite possible that the lower type frequency of irregular forms relative to regular forms in the language were partly responsible for these lower accuracy levels.

The finding of group differences being confined to the regular forms should not be interpreted to mean that the children with LI were functioning at expected age level on irregular forms. Recall that the comparison groups were typically developing children who were matched to the children with LI according to receptive vocabulary scores. Given the relatively low receptive vocabulary scores of the LI groups, then, the typically developing children with whom they were compared were somewhat younger. For this reason, it is more accurate to assume that, even in the absence of group differences, the children with LI were probably less proficient than same-age typically developing peers.

As expected by the morphological richness account, errors on items requiring the stem plus plural plus accusative case suffix were usually productions of the stem plus only one of the suffixes. Recall that stem-only productions occurred in the data, but were more frequent if the item required a stem plus a single suffix. Thus, in each of these instances, an error failed to include one suffix, and thus differed minimally from the correct form. Also as predicted, substitutions of a plural suffix for an accusative case suffix or vice versa did not occur. In addition, there was no evidence of the use of a default form. For example, the children could have consistently used bare stems, or relied on a stem and one particular suffix (regardless of correctness), but this pattern was not seen. Thus, it appeared that the children had some degree of knowledge of the appropriate forms, even when they did not succeed in producing the correct response.

The high degree of overgeneralizations was also consistent with predictions. There were numerous instances in which the children used the appropriate suffix but failed to select the appropriate stem form of the noun. Productions of this type could be expected because whereas the suffix appears with both regular and irregular forms, the phonological details of the stem are less transparent and may have had limited strength in the children's word-specific paradigms. A limited degree of overgeneralization was seen for items requiring regular forms. This finding was not surprising given that even certain classes of regular forms require phonological modification when a suffix is added. However, overgeneralizations were much more frequent — as expected — for items requiring irregular forms. These types of productions were abundant in the data for both younger and older children and for the VC as well as the LI groups.

According to the morphological richness account, frequency of occurrence will be a predictor of success because with a greater number of encounters with an inflected word and thus a larger number of opportunities to process it, the representation of the inflected word in the paradigm will be stronger. However, frequency operated in somewhat different ways in the LI and VC groups. For the children with LI, word frequency was a significant predictor. Specifically, for the younger children with LI, word frequency was a predictor of accuracy on both regular and irregular forms. For the older LI group, it was a predictor only for irregular forms. In contrast, the older VC group's accuracy levels with both regular and irregular forms were unrelated to frequency, and, for the younger VC children, stem frequency, rather than word frequency was the better predictor for accuracy on irregular forms. The latter finding suggests that the younger VC children were less dependent on the frequency of occurrence of the entire inflected word. Instead, it appears that accuracy could be achieved when the stems themselves were of higher frequency of occurrence, for the suffixes could be attached by importing them from the general paradigm. The LI groups were instead more sensitive to the frequency of the entire word, including its suffix. Although these children's overgeneralizations clearly showed that they were not entirely dependent on the rote learning of inflected words, the finding of whole-word, rather than stem frequency effects in the LI groups suggests that these children were slower to discard rote learning as a means of learning the suffixed forms of nouns.

To sum up, the overall findings were in keeping with expectations based on the morphological richness account. However, explanations of certain details in the data are not readily apparent. For example, for irregular stem class items, the LI and VC groups did not differ using either scoring method. We expected that the children's accuracy levels would be lower for irregular stem class items than for regular stem class items when the more stringent scoring method was applied. It seemed likely that dealing with the morphophonological challenges involved in irregular stem classes may have exacted a price in the form of adversely affecting the children's choice of suffixes. Less expected was that this adverse effect was as strong or stronger in the case of the VC children. These children held an advantage over the children with LI with regular noun stem classes, yet lost the advantage when the noun stem classes were irregular, even when allowances were made for overgeneralizations of the stem.

However, whereas the challenges posed by irregular stem class items might have been expected to minimize any difference between the younger LI and VC groups, it was surprising that no group difference was seen between the older children with LI and older VC children. The reason for the lack of a difference appears attributable to the rather dramatic increase in overgeneralization by the older VC children, especially for items requiring an accusative case suffix. It appears as if the challenge of determining the form of the stem interfered with the selection of the appropriate suffix.

Other contributions of the data

Along with its contribution to the evaluation of the morphological richness account, the present study also addresses issues that have rarely been applied to the study of children with LI. For example, one of the suffixes examined in this study — the accusative case suffix — deals with a grammatical case that is not easy to examine in an unobstructed manner. In English, an assessment of the use of case marking by children with LI is necessarily limited to pronouns. Accusative case pronouns are not problematic for English-speaking children with LI. However, accusative case pronouns constitute the default form in English. Therefore, the absence of errors such as *Mommy kissed he* could be due to adequate control of accusative case or the use of pronouns that are marked for person, number, and gender but unspecified for case.

In German, nominative case forms serve as the default, and German-speaking children with LI do make errors such as the selection of determiners with, for example, nominative case in contexts that require accusative case. However, because these morphemes also express gender and number information, the children's errors cannot be attributed to case alone. It may be that accusative case constitutes the problem in these instances, or, alternatively, problems dealing with combinations of functions (gender and number in addition to case) may be the source of the problem. (It is also true that in many instances — such as definite and indefinite feminine singular, definite and indefinite neuter singular, and definite plural — there is no distinction between nominative and accusative case in the determiner system of German.).

In contrast to these other languages, Hungarian provides a clear view of children's use of accusative case through the accusative case suffix examined in the present study. This suffix marks only accusative case. Even when it is not the only suffix attached to the stem, its identity remains clear. We found that the children with LI had no special difficulty with accusative case when this suffix was the only one required with a noun stem. As noted earlier, these children experienced greater difficulty when both a plural and an accusative case suffix was required. However, when these forms were reduced to the stem and only a single suffix, the suffix marking accusative case was the most likely to be retained.

Another advantage of Hungarian is that it allowed us to examine how children with LI make use of agglutinating morphology. Given the assumptions of the morphological richness account, we expected less accuracy with noun plus plural plus accusative case suffixes than with noun plus plural suffixes or noun plus accusative case suffixes. However, the nature of the errors made by the children on these agglutinating forms was quite instructive. We found no instances in which the children produced the two suffixes in the incorrect order; plural suffixes always preceded accusative case suffixes, as required in the language. This finding was not attributable to the lack of opportunity for such errors to occur. As can be seen from figures 1–4, even with the more stringent scoring method, the children with LI averaged from over 65% correct to approximately 85% correct in their use of noun plus plural plus accusative case forms. Thus, there were many opportunities for the two suffixes to be produced in the wrong order.

Conclusions

In summary, the findings of this study provided support for several assumptions of the morphological richness account, but also revealed certain details in greater need of explication. As expected in this account, differences between children with LI and younger typically developing children in this inflectionally rich language were not great and occurred principally when multiple suffixation was required. Errors could be viewed as near misses, as errors on stem plus plural plus accusative case suffix items were more likely to be productions of the stem and one of these suffixes rather than stems alone, whereas productions of stems alone were more likely to occur when the item required a stem and one suffix. The children did not resort to default forms, suggesting that the children had some command of the grammatical functions involved. Overgeneralizations were frequent, as expected. Less expected was the finding that the LI groups relied to a great extent on the frequency of occurrence of the entire inflected word, whereas the only frequency metric that was a predictor for the typically developing children was stem frequency, and this held for the younger VC group only. Future research should explore the basis of these differences pertaining to frequency of occurrence effects. It will be important to determine whether these differences merely reflect alternative routes to eventual mastery, or instead reflect a distinction between a successful route (early partial dependence on stem frequency) and a route (partial but prolonged dependence on word frequency) that reflects or even contributes to an impairment in grammatical morphology.

Acknowledgements

This research was supported by Research Grant Number R01 DC00458 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (USA) to Laurence B. Leonard and by Research Grant Number OTKA TS 049840 from the Hungarian National Science Foundation to Csaba Pléh. Ágnes Lukács was a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science. The authors are grateful to the children in Simon Antal Primary School in Vác, in the Dr Nagy László Institute of Special Education in Kószeg, in the ELTE Special Preschool and Early Intervention Centre and in the Zölderdő Preschool for speech therapy and nature preservation for their participation. They also thank the speech therapists in both institutions for their help with screening and organization. **Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Notes

1. For ease of exposition, we use standard Hungarian orthography and do not give phonetic transcriptions. Hungarian orthography is fairly transparent, geminates are marked by double consonants (also by doubling the first letter in a consonant digraph), and accents above vowels mark length. However, not every accented vowel is phonetically equivalent to their short counterpart, so we present the phonetic symbols for Hungarian vowels and non-transparent consonantal letters here. Vowels: a [ɔ], á [a:], o [o], ó [o:], u [u], ú [u:], e [ɜ], é [e:], i [i], í [i:], ö [ø], ő [ø:], ü [y], ű [y:]; and consonants: c [ts], cs [tʃ], dzs [ʤ], g [g], gy [ɟ], j [j], ly [j], ny [ɲ], r [r], s [ʃ], sz [s], ty [ç], zs [ʒ]. Our description is based on Kiefer (1998), Nádasdy and Siptár (1994), and Törkenczy (1994). Linguistic accounts of the alternations are not presented; we provide only the information relevant to understanding the behaviour of the inflected forms examined in this study.
2. The 15 older children with LI and their VC pairs were also part of a larger group of children who participated in a separate study on verb agreement in Hungarian (Lukács et al. forthcoming).
3. The authors are grateful to Professor Dorothy Bishop for providing us with the TROG for this purpose. To date, 600 typically developing children have been tested as part of the standardization process. The scores of the children with LI in this study were compared against the values obtained for the typically developing children participating in the standardization.

4. The authors are grateful to Péter Halácsy for the frequency calculations for bound allomorphs.
5. Results for statistical tests in error analysis are only given where differences are significant.
6. This is due to the phonotactics of the accusative: overgeneralized accusative forms result in phonotactically more well-formed sequences than overgeneralized plural forms in most cases.

References

Bates, E. and MacWhinney, B., 1989, Functionalism and the competition model. In B. MacWhinney and E. Bates (eds), *The Cross-linguistic Study of Sentence Processing* (Cambridge: Cambridge University Press), pp. 3–73.

Bedore, L. and Leonard, L., 1998, Specific language impairment and grammatical morphology: a discriminant function analysis. *Journal of Speech, Language, and Hearing Research*, **41**, 1185–1192.

Bedore, L. and Leonard, L., 2001, Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **44**, 905–924.

Bishop, D. V. M., 1983, *Test for the Reception of Grammar* (published by the author).

Bortolini, U., Caselli, M. C. and Leonard, L., 1997, Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **40**, 809–820.

Csányi, F. I., 1974, *Peabody Szókincs-Teszt* (Budapest: Bárczi Gusztáv Gyógypedagógiai Főiskola).

Dromi, E., Leonard, L., Adam, G. and Zadunaisky-Ehrlich, S., 1999, Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **42**, 1414–1431.

Dunn, L. and Dunn, L., 1981, *Peabody Picture Vocabulary Test — Revised* (Circle Pines, MN: American Guidance Service).

Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I. and Trón, V., 2004, Creating open language resources for Hungarian. In Proceedings of the Language Resources and Evaluation Conference, pp. 203–210.

Kas, B. and Lukács, Á., 2008, Magyar Mondatutánmondási Teszt. Unpublished manuscript, Budapest University of Technology and Economics.

Kiefer, F., 1998, Alaktan. In: È. K. Kiss, F. Kiefer and P. Siptár (eds), *Új magyar nyelvtan* (Budapest: Osiris), pp. 187–293.

Kornai, A., Halácsy, P., Nagy, V., Trón, V. and Varga, D., 2006, Web-based frequency dictionaries for medium density languages. In A. Kilgarriff (ed.), *Proceedings of the 2nd International Workshop on Web as Corpus*, Marco Baroni. ACL-06 (Paris: European Language Resources Association), pp. 1–9.

Leonard, L., 1989, Language learnability and specific language impairment in children. *Applied Psycholinguistics*, **10**, 179–202.

Leonard, L., 1998, *Children with Specific Language Impairment* (Cambridge, MA: MIT Press).

Leonard, L., Hansson, K., Nettelbladt, U. and Deevy, P., 2004, Children with specific language impairment in English and Swedish: a crosslinguistic study. *Language Acquisition*, **12**, 219–246.

Leonard, L., Sabbadini, L., Leonard, J. and Volterra, V., 1987, Specific language impairment in children: a crosslinguistic study. *Brain and Language*, **32**, 233–252.

Leonard, L., Salameh, E. K. and Hansson, K., 2001, Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics*, **22**, 619–639.

Lukács, Á., 2001, Szabályok és kivételek: a kettős modell érvényessége a magyarban. In C. Pléh and A. Lukács (eds), *A magyar morfológia pszicholingvisztikája* (Budapest: Osiris), pp. 119–152.

Lukács, Á., Leonard, L. B., Kas, B. and Pléh, C., 2009, The use of tense and agreement by Hungarian-speaking children with language impairment. *Journal of Speech, Language and Hearing Research*, **52**, 1–20.

Lukács, Á. and Pléh, Cs., 1999, Hungarian cross-modal priming and treatment of nonsense words supports the dual process hypothesis. Commentary on Harald Clahsen's Rules of Language. *Behavioral and Brain Sciences*, **22**, 1030–1031.

MacWhinney, B., 1975, Rules, rote and analogy in morphological formations by Hungarian children. *Journal of Child Language*, **2**, 65–77.

MacWhinney, B., 1978, The acquisition of morphophonology. *Monograph of the Society for Research in Child Development*. 174.

MacWhinney, B., 1987, The Competition Model. In B. MacWhinney (ed.), *Mechanisms of Language Acquisition* (Hillsdale, NJ: Lawrence Erlbaum), pp. 249–308.

Nádasdy, Á. and Siptár, P., 1994, A magánhangzók. In F. Kiefer (ed.), *Strukturális magyar nyelvtan. 2. kötet: Fonológia* (Budapest: Akadémiai), pp. 42–181.

Pléh, C., 2000, A magyar morfológia pszicholingvisztikai aspektusai. In F. Kiefer (ed.), *Strukturális magyar nyelvtan 3: Morfológia* (Budapest: Akadémiai), pp. 951–1021.

Pléh, Cs. and Lukács, Á., 2002, A szabályok és a kettős disszociációs elv a nyelv agyi reprezentációjában. In E. Sz. Vizi, F. Altrichter, K. Nyíri and Cs. Pléh (eds), *Agy és tudat* (Budapest: BIP), pp. 153–168.

Pléh, Cs., Palotás, G. and Lőrik, J., 2002, *Nyelvfejlődési szűrővizsgálat* (Budapest: Akadémiai).

Pléh, Cs., Vinkler, Zs. and Kálmán, L., 1997, Early morphology of spatial expressions in Hungarian children: a CHILDES study. *Acta Linguistica Hungarica*, **44**, 249–260.

Racsmány, M., Lukács, Á., Németh, D. and Pléh, Cs., 2005, A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle*, **4**, 479–505.

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

Rice, M., 2003, A unified model of specific and general language delay: grammatical tense as a clinical marker of unexpected variation. In Y. Levy and J. Schaeffer (eds), *Language Competence Across Populations: Toward a Definition of SLI* (Mahwah, NJ: Lawrence Erlbaum Associates), pp. 63–94.

Rice, M. and Wexler, K., 1996, Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research*, **39**, 1239–1257.

Rice, M. and Wexler, K., 2001, *Rice/Wexler Test of Early Grammatical Impairment* (San Antonio, TX: Psychological Corporation).

Törkenczy, M., 1994, A szótag. In F. Kiefer (ed.), *Strukturális magyar nyelvtan. 2. kötet: Fonológia* (Budapest: Akadémiai), pp. 273–393.

Wexler, K., 1998, Very early parameter setting and the unique checking constraint. *Lingua*, **106**, 23–79.

Wexler, K., 2003, Lenneberg's dream: learning, normal language development, and specific language impairment. In Y. Levy and J. Schaeffer (eds), *Language Competence Across Populations: Toward a Definition of Specific Language Impairment* (Mahwah, NJ: Lawrence Erlbaum), pp. 11–61.

Study 3.

Kemény F., Lukács Á. (2010) Impaired procedural learning in language impairment: results from probabilistic categorization *Journal of Clinical and Experimental Neuropsychology.* 32:3, 249–258. doi: 10.1080/13803390902971131

Thesis 6. Cognitive impairments in SLI do not selectively target language; deficits also occur in skill learning outside the language domain, most prominently for sequentially organized stimuli (3, 8)

Thesis 7. In concert with the Procedural Deficit Hypothesis, procedural learning is vulnerable in SLI, while processes of declarative learning and retention are relatively intact (3, 8, 12)

JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
2009, i*First*, 1–12

Psychology Press
Taylor & Francis Group

Impaired procedural learning in language impairment: Results from probabilistic categorization

Ferenc Kemény[1] and Ágnes Lukács[1,2]

[1]Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary
[2]Research Institute of Linguistics, Hungarian Academy of Sciences, Budapest, Hungary

The Weather Prediction (WP) Task is a classical task of probabilistic category learning generally used for examining the dissociation of procedural and declarative memory. The current study focuses on performance of children with language impairment (LI) and compares their performance to that of typically developing (TD) children and adults with the aim of testing the procedural deficit hypothesis of LI (PDH; Ullman & Pierpont, 2005), which states that language impairment is not a specific linguistic phenomenon, but results from the dysfunction of a more general cognitive system: the procedural system. To test the generality of the procedural impairment, we needed a task that is dissimilar from language in that it does not build on sequential information. Children with language impairment show deficient learning on the Weather Prediction Task, which already appears at the early stages of the task. These results, in line with the PDH, point to the deficit of the procedural system in language impairment going beyond the language system. Whether this deficit is selective to the procedural system or is complemented by deficits in the declarative system is the subject of future studies.

Keywords: Implicit learning; Procedural system; Probabilistic categorization; Language impairment; Procedural deficit hypothesis.

Specific language impairment (SLI) is a developmental disorder usually characterized by focal disorder of the linguistic domain. Children with language impairment show a significant delay in language abilities in spite of not having any hearing deficits, neurological disorders, environmental deprivation, or mental retardation that could account for their language problems (e.g., Bishop, 1992; Leonard, 1997; Ullman & Pierpont, 2005). The core deficit concerns grammar, manifests itself as a difficulty in using suffixes and/or specific syntactic structures, and often persists into school years.

Children with a central deficit in the domain of language are often claimed to have *specific* language impairment, a term implying that language or grammar can be selectively impaired in an otherwise intact cognitive system. This view is problematic for several reasons. The disorder is very heterogeneous, and although many attempts have been made to identify specific subgroups (e.g., Aram, Morris, & Hall, 1993; Bishop & Adams, 1992; Vargha-Khadem, Watkins, Alcock, Fletcher, & Passingham, 1995; Whitehurst et al., 1991), a proper system of subcategorization with diagnostic validity is still missing. Another key problem is that

We are grateful to all adult participants and to the children and their teachers in the Dr. Nagy László Institute of Special Education in Kőszeg and in the primary school in Kerék Street in Budapest for their cooperation. Special thanks go to Attila Buknicz, István Kemény, and Bence Fadgyas for their help, to Dezső Németh and Szabolcs Kéri for their comments and suggestions, and to Karolina Janacsek and Gergely Éliás for their help in statistics. The research was funded by National Institutes of Health (NIH) R01DC00458 to Larry Leonard and the Országos Tudományos Kutatási Alapprogramok (Hungarian Scientific Research Fund, OTKA) TS 049840 to Csaba Pléh. During the research Ágnes Lukács was a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science.

Address correspondence to Ferenc Kemény, Department of Cognitive Sciences, Budapest University of Technology and Economics (BME), Stoczek u. 2, Budapest 1111, Hungary (E-mail: fkemeny@cogsci.bme.hu).

http://www.psypress.com/jcen DOI: 10.1080/13803390902971131

specific language impairment in many (probably in most) cases turns out to be not as specific as claimed, and impairments in several nonlinguistic abilities tend to accompany the grammatical deficit, including motor control of oral and fine movements, hypothesis testing and categorization (on both linguistic and nonlinguistic material), mental rotation, sequencing, word retrieval, phonological discrimination, simultaneous execution, and, perhaps most apparently, executive functions (Leonard, 1997; Ullman & Pierpont, 2005).

Besides theories that treat language or grammar as a functionally and anatomically isolated module with the potential for selective impairment, there have been several proposals for more basic mechanisms behind nonlinguistic impairments leading to language impairment. Most well known are claims about deficits in processing rapidly changing auditory input (Tallal & Piercy, 1973) and a decreased capacity of phonological short-term memory (Gathercole & Baddeley, 1990). Most relevant to this paper is a theory that tries to integrate all the linguistic and nonlinguistic deficits observed in SLI into a neurobiological model. The procedural/declarative model (Ullman & Pierpont, 2005) of language claims that there is a clear dissociation within language between the grammar and the lexicon, since they are functions of different memory systems. Grammar is a procedural function, while the lexicon is based on declarative memory. These two memory systems are dissociated not just functionally, but also anatomically, as is explained later in detail. This model has a specific prediction for the nature of language impairment called the procedural deficit hypothesis of specific language impairment (PDH; Ullman & Pierpont, 2005). On this view, language impairment is a result of abnormal development of brain structures underlying the procedural memory system responsible for learning cognitive and motor skills (for sequence and rule learning) and, among them, grammar. Developmental disorders of such a system should result in deficits of skills that rely on procedural learning within both the linguistic and nonlinguistic domains.

The PDH thus suggests that the developmental disorder termed specific language impairment is not specific to language, but is rather a deficit of the more general system of procedural memory. As it has already been stated the procedural system is responsible for the acquisition not only of motor skills, but also of cognitive skills (Knowlton, Mangels, & Squire, 1996a) like probabilistic category learning or rule learning, an example of which is learning and using grammatical rules (Ullman et al., 1997). Based on the model, if children with SLI indeed have a more general procedural deficit, they are expected to show lower performance on different sorts of tasks, linguistic and nonlinguistic, requiring the soundness of the procedural system. Accordingly, earlier research found deficient performance in implicit sequence learning in language impairment (Tomblin, Mainela-Arnold, & Zhang, 2007). The study compared the performance of adolescents with and without language impairment. Both groups showed learning on the serial reaction time (SRT) task, but the learning rate of language-impaired adolescents was significantly slower than that of the control group.

There are also results showing that language-impaired subjects have problems with the processing of rapid sequential information in the auditory domain (Tallal & Piercy, 1973) and also have deficits with other motor functions, like orofacial motor movements, that are best exemplified in speech movement (Vargha-Khadem et al., 1998; Watkins, Dronkers, & Vargha-Khadem, 2002a; Watkins et al., 2002b). At the same time manual praxis, generally, seems to be impaired too, but there is one study suggesting that manual praxis can be intact in certain cases (Vargha-Khadem et al., 1995). The comorbidity of language impairment and poor motor skills (both sequential and nonsequential) seems to be quite high (Hill, 2001).

What seems clear from the literature is that language impairment often involves a wide variety of motor deficits. There are several results indicating the comorbidity of sequence learning and language impairment (LI), and the implicit acquisition of nonsequential motor information also seems to be impaired. This suggests that language impairment is not properly characterized by linguistic features only and cannot be considered a pure linguistic phenomenon.

The implicit–explicit and the declarative–procedural distinction

The nature of implicit learning has been well studied throughout the last few decades, although not many conceptual works focused on the different aspects of multiple memory systems (see Poldrack & Foerde, 2008, for details). Implicit learning is the incidental acquisition of complex information with difficulty in explicitly recollecting the information acquired (Meulemans, van der Linden, & Perruchet, 1998). The concept of implicit learning has mostly been discussed in a multiple memory systems model (Knowlton, Squire, & Gluck, 1994), in which human memory is not homogenous: It has implicit and explicit functions. The former does not involve explicit awareness, while the latter does, so the distinction is based on the mode of recollection of knowledge and its availability to conscious information processing.

Although the definition cited above does not imply it, acquiring information in the traditional implicit learning tasks (like the serial reaction time task, artificial grammar learning, or probabilistic categorization) requires a lot more time and learning trials than acquiring information through explicit learning. This difference can be accounted for by the distinction of Squire, Knowlton, and Musen (1993), on the basis of the representations involved. The declarative system manipulates factual representations that have clear boundaries, while the procedural system uses more dynamic representations, which are like procedures and are mostly acquired in an incremental way.

The dissociation of the declarative and procedural systems (as explained earlier) cannot only be diagnosed by behavioral symptoms: The two systems are anchored in different brain regions (e.g., Knowlton & Squire, 1993; Ullman & Pierpont, 2005; though this view is not general, see, e.g., Voermans et al., 2004; Yin & Knowlton, 2006). While the functioning of the declarative system mainly relies on the activity of the mediotemporal lobe (Knowlton & Squire, 1993), the procedural system is based on more diverse regions, of which the most important are the basal ganglia, the fronto-striatal pathways, the cerebellum, Broca's area, and other areas handling movement, action planning, and motor execution (Ullman & Pierpont, 2005).

It has been known that the procedural system is responsible for the acquisition of motor skills (Knowlton & Squire, 1993), but the same system is responsible for the acquisition of nonmotor cognitive skills including the acquisition of a categories (Knowlton et al., 1996a) or the abstraction of a prototype (Nosofsky, Stanton, & Zaki, 2005). A wide variety of phenomena touching upon procedural learning are traditionally covered by the implicit learning literature: motor sequence learning (Meulemans et al., 1998), perceptual sequence learning (Remillard, 2003), artificial grammar learning (Aslin, Saffran, & Newport, 1999; Reber, 1967), probabilistic category learning (Knowlton et al., 1994), and so on.

There have been very few studies that directly addressed the problem of implicit/procedural learning in language impairment, but these scarce and controversial findings together with results from patients with neurodegenerative disorders of the procedural system (like Parkinson's and Huntington's syndrome) urge research in this field. There are results showing that sequence learning performance is decreased in dyslexia (Stoodley, Harrison, & Stein, 2006) and language impairment (Tomblin et al., 2007), and the same is true of Parkinson's (Siegert, Taylor, Weatherall, & Abernethy, 2006) and Huntington's syndrome (Knopman & Nissen, 1991). Artificial grammar learning (AGL) at the same time does not seem to be impaired in dyslexia (Rüsseler, Gerth, & Munte, 2006); there are studies finding both intact (Witt, Nuhsman, & Deuschl, 2002) and impaired (Smith & McDowall, 2006) AGL performance in Parkinson's syndrome. Research suggests that Huntington's syndrome does not involve impairment in AGL (Knowlton et al., 1996b). Results of one study exploring the link between language/learning disability and performance on the AGL task (Plante, Gómez, & Gerken, 2002) suggest that adults with language/learning disability show low performance on the AGL task, a task that is expected to model sensitivity to word order cues, which is displayed by even one-year-olds (Gómez & Gerken, 1999). This is another suggestion that language impairment is related to implicit learning abilities, but the nature of the deficit needs further investigation in different domains of implicit learning.

The Weather Prediction Task

One of the most frequently used probabilistic category learning tasks is the Weather Prediction (WP) Task (Knowlton et al., 1994). This is a dichotic decision-making task. Participants are presented with an image of a combination of one, two, or three of four cues (which can be either different tarot cards or geometrical shapes). They have to decide whether the pattern they see predicts SUNSHINE or RAIN and have to respond accordingly. As soon as they have made their choice a feedback appears to show whether they were right or wrong. This makes the WP task different from AGL and SRT: The WP task includes feedback-based incremental learning (Shohamy, Myers, Onlaor, & Gluck, 2004). This type of category learning is a cognitive skill involving activity in the procedural system (Knowlton et al., 1996a). There have been hardly any studies examining the relationship between the three traditional implicit learning tasks—that is, the SRT, AGL, and WP tasks. We are only aware of one (Aczel & Gonci, 2005) that found no correlation between performance on the three tasks.

The WP task is also similar to some other tasks measuring implicit learning. The Iowa Gambling Task (IGT) is an implicit learning task to measure risk taking (Bechara, Damasio, Damasio, &

Anderson, 1994). The key feature of the task is to compute and clash the short- and long-term benefits and punishments within the task. A central neural structure behind performance is the ventromedial prefrontal cortex, which is proposed to merge the information from the emotional somatic markers supporting decision making. Though there have been studies focusing on the role of the prefrontal cortex in the WP task (Kincses, Antal, Nitsche, Bartfai, & Paulus, 2004), the primary areas active during probabilistic category learning are the basal ganglia (Hopkins, Myers, Shohamy, Grossman, & Gluck, 2004; Knowlton et al., 1996a; Knowlton et al., 1994; Poldrack et al., 2001; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999; Shohamy et al., 2004). This difference might be due to the fact that a mistake in the WP task does not involve strong emotions, while emotions play a central role in the IG task; the IGT has been widely used as a research tool for studying the somatic marker hypothesis (Bechara et al., 1994; Dunn, Dalgleish, & Lawrence, 2006).

In the original version of the task (Knowlton et al., 1994) there are altogether four cues, and each cue has its own predictive value. Cue 1 predicts SUNSHINE in 77% of all cases, Cue 2 in 58%, Cue 3 in 42% and Cue 4 in 23%. The task consists of several blocks of 50 trials. Learning performance is measured by the difference between accuracy on the first and later blocks. People are expected to achieve a performance above 70% correct on this task (Gluck, Shohamy, & Myers, 2002; Knowlton et al., 1994; Knowlton et al., 1996b). According to the Rescorla–Wagner law (Knowlton et al., 1994; Rescorla & Wagner, 1972) participants are expected to learn the response to those cues first that have the best predictive value—that is, the cues that determine the outcome most. In the case of the original task these are the cues with 77% and 23% probability values. After the individual cues are learnt people are expected to integrate them into a whole pattern and (implicitly) start calculating the predictive values and give the answer according to the result.

The Weather Prediction Task proved to be a useful way of testing implicit learning in different adult neuropsychological disorders. In their study Knowlton and her colleagues (1994) compared the performance of amnesic patients with an age-matched control group. The etiologies of their subject were heterogeneous: The group included patients with Korsakoff syndrome, bilateral brain infarct, bilateral traumatic brain injury, and anoxia. The only feature that was shared by all patients was that they had amnesia associated with a mediotemporal lobe (MTL) impairment. During the first 50 trials there was no difference between the amnesic and control groups, both showing learning. Significant differences emerged in the later blocks only: The control group's performance was significantly better than that of the clinical group. Under this assumption, one would expect to see MTL activity during the second half of the task, a prediction that was borne out by Poldrack and his colleagues' positron emission tomography (PET) studies (Poldrack et al., 2001; Poldrack et al., 1999), which showed that during the early stages of the feedback-based Weather Prediction Task the striate, the caudate nucleus (NC), and the cortico-NC pathways are active while the activity of the MTL only increases in the second half of the task, probably as a result of verbalization.

In a subsequent study Knowlton and her colleagues (1996a) found a double dissociation between the two systems. Their results showed that people with Parkinson's syndrome showed significantly lower performance than amnesic and control participants in the first 50 trials. Their performance did not differ from chance at the beginning of the task, but later they started to improve. According to Knowlton and colleagues this is due to the fact that learning in the early stages relies on activation of the procedural system. At later stages of the task procedural activation decreases, while declarative processes emerge. With the emergence of declarative strategies, the rate of the procedural load decreases, and patients with impaired procedural functioning start to show better performance on the task—that is, Parkinson's patients start to improve. Due to the same reason the performance of amnesic patients declines after performing well on the first blocks. The theory of the early activation of the procedural system and the later declarative functioning has been tested in a PET study (Poldrack et al., 2001; Poldrack et al., 1999), which confirmed Knowlton and colleagues' (1996; 1994) results.

Possible strategies for solving the Weather Prediction Task

Unconscious strategies for solving the Weather Prediction Task are essential to understand the diversity of performance within the typical population and also to explain differences between clinical and control groups. There have been several attempts to identify the behavioral difference associated with the switch from procedural to declarative functioning. Since it is one of the key points of the present research, it is necessary to go into detail about these strategies. Gluck and colleagues (2002) were searching for an answer as to whether there are differences in robust strategies that are responsible for this variance. They identified three different strategies:

- Multicue strategy
- Singleton strategy
- One-cue strategy.

The multicue strategy is the most optimal way of solving the Weather Prediction Task. It is basically following all four cues and averaging them before decision. The use of this strategy throughout the task leads to more than 80% correct answers. This is the strategy we would probably use if the task was an explicit mathematical task—that is, if we are facing the combination of two cues, one with 23%, the other with 58% predictive strength for sunshine, we calculate the average (40.5%), decide whether it is above 50% (SUNSHINE) or below that (RAIN) and give an answer accordingly.

In the singleton strategy participants answer consistently if cues appear on their own, and they answer randomly when cues appear in combination. As a result, cue-based answers are given only if any of the cues appears alone—that is, seeing Cue 1 or Cue 2 alone, participants predict SUNSHINE, but if only Cue 3 or Cue 4 appears on the screen, they predict RAIN. Any other patterns (all combinations of cues) lead to random responses, resulting on these trials in an average performance of 50%. The optimal use of this strategy leads to about 70% correct performance.

The one-cue strategy is the consistent use of one cue as a predictor of an outcome. It means that, for example, the participant consistently answers SUNSHINE when Cue 1 is present (either alone or in a combination), but if it is not then his answers do not differ from chance. This strategy leads to a performance just above 60%.

Preliminary research with 30 participants showed that 27 used the singleton strategy, and 3 used the one-cue strategy on the first 50 trials. Later on more and more of them started to use the multicue strategy (Gluck et al., 2002).[1] Gluck and colleagues argued that the improvement of the participants' performance is not only quantitative, a proposal that is supported by PET studies (Poldrack et al., 2001; Poldrack et al., 1999) showing that different brain regions are active in the different stages of the task. The shift in activation patterns from the procedural to the declarative system can also be observed and distinguished behaviorally—that is, based on strategy use.

It has to be pointed out that from the perspective of the procedural–declarative dissociation the one-cue and the singleton strategies do not differ. According to Gluck and colleagues (2002) this behavioral difference is almost the same as the difference between the one-cue strategies—that is, they are different manifestations of the same learning process. The important difference lies in the distinction of the single (one-cue and singleton) and multicue strategies. Single strategies always rely on the appearance of a single cue (which is set by either the specifics of the cue, i.e., one can use a strategy relying on Cue 1 OR Cue 2 OR Cue 3 OR Cue 4 alone, or the number of the cue, i.e., any of the cues ALONE). Multicue strategies on the other hand involve the use of more cues and the comparison of cue combinations along their predictive value: Single strategies (i.e., one-cue and singleton strategies) rely on the procedural system, while the multicue strategy relies on the declarative system.

To test the strategy-use hypothesis, Hopkins and colleagues (2004) repeated Knowlton et al.'s (1994) study in a later experiment with a more homogenous clinical group with only patients with specific bilateral MTL damage (Hopkins et al., 2004). The overall performance of the amnesic group was significantly lower than that of the control group. As expected, the difference between the clinical and control groups was not seen during the first 50 trials and only appeared in later phases of the task. In later phases of learning, control participants switched to the multicue strategy, but amnesic patients were unable to do so. While the multicue strategy seemed to be difficult for the hypoxic patients, all their performance patterns were fit properly to the single strategies. This confirms the hypothesis that the use of the multicue strategy is probably hippocampus dependent, while the single strategies can rely on the procedural system (Knowlton et al., 1994).

Shohamy and colleagues (2004) repeated another experiment of Knowlton et al. (1996a). They compared patients with mild Parkinson's syndrome with a control group. Both groups showed learning during the first 50 trials, and there were no group differences. Control participants—consistent with the earlier results—gradually switched from single-cue strategies to multicue strategy, while this change did not appear in Parkinson's patients. Shohamy and colleagues (2004) suggested that the discrepancy between their and Knowlton and colleagues' (1996a) result is due to the differences in the severity of the state of patients with Parkinson's disease participating in the two studies. These results are not in concert with the procedural–declarative

[1]Gluck et al. (2002) and studies that followed used the predictive values of 80%, 60%, 40%, and 20% in the WP task.

model of probabilistic category learning, a discrepancy that we address in more detail in the Discussion section.

A critical review of literature on the Weather Prediction Task suggests that the introductory parts of the task rely on single strategies with procedural activity, while the later phases build on multicue strategies associated with activity in the declarative system. Deficits of the procedural system (like Parkinson's syndrome) are expected to be associated with a decreased overall performance, especially in the early phases of the task, which is manifested by an inability to use single strategies. As the declarative system starts to be active—between approximately the second and third blocks of 50 trials—improvement should be observed. The reason for this could be a compensatory reliance on declarative strategies for those who were unable to use single strategies due to procedural malfunction. If the procedural deficit hypothesis of language impairment (Ullman & Pierpont, 2005) is correct, we expect children with language impairment to show a performance pattern similar to that of Parkinson's patients and to show (a) no learning in the first 50 trials, and/or (b) learning in the third block of 50 trials, and/or (c) an inability to switch from single to multicue strategies. To test this hypothesis, we compared WP performance of a group of children with language impairment to typically developing children matched on chronological age. This study is the first to examine probabilistic category learning in children with LI (and also typically developing, TD, children). As we used a modified version of the original WP task to suit children, we first tested adults to set the baseline of learning and to make comparisons to other studies possible.

METHOD

Participants

Adults

A total of 16 adults (5 female, 11 male) participated in the study. All of them were recruited at the Budapest University of Technology and Economics and participated voluntarily in the study for credits. Their mean age was 20;5 with a standard deviation of 1;7. They were informed about the purpose of the research after the study. All participants provided written informed consent, in accordance with the principles set out in the Declaration of Helsinki and the stipulations of the local Institutional Review Board.

Children with language impairment

A total of 16 children (5 girls, 11 boys) were selected for the language-impaired group. Their mean age was 11;3 with a standard deviation of 1;3. All language-impaired children were students of an institute of special education in Kőszeg. All of these children met the criteria for LI. Each child scored above 85 on the Raven Colored Progressive Matrices (Raven, Court, & Raven, 1987), a measure of nonverbal intelligence. All children passed a hearing screening, and no child had a history of neurological impairment. Each child scored at least 1.5 standard deviations below age norms on at least two of four language tests administered. These four tests included two receptive tests and two expressive tests. The receptive tests were the Hungarian versions of the Peabody Picture Vocabulary Test (PPVT; Csányi, 1974; Dunn et al., 2006) and the Test for Reception of Grammar (TROG; Bishop, 1983). The expressive tests were the Hungarian Sentence Repetition Test (Magyar Mondatutánmondási Teszt, MAMUT; Kas & Lukács, 2007) and a nonword repetition test (Racsmány, Lukács, Németh, & Pléh, 2005).

Performance of the LI group was compared to that of a control group of 16 children matched individually on chronological age to children in the LI group (4 girls, 12 boys; all from a primary school in Budapest). Their mean age was 11;3 (*SD* 1;2). All children in the LI and in the TD groups 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.

Procedure

Participants were presented with a special version of the Weather Prediction Task. In this version the procedure was the same as that explained earlier: Participants were shown a pattern of one, two, or three cues (in this case out of four simple geometric forms), and for each combination they had to tell whether it predicted sunshine or rain. Immediate feedback was given showing whether they were right or wrong. Each of the four cues had its own probability of determining SUNSHINE: 90%, 70%, 30%, and 10%. These probabilities were adjusted to make the task easier for children, as preliminary studies with the original weights did not show any learning in this age group. For the same purpose, we deviated from the routine of earlier studies in presenting participants with only three blocks of 50 trials, since we tried to lower the requirements for children. After every 25 trials there was a recess-slide, which was on as long as one of the buttons was pressed. Number and percentage of correct answers were measured for each block, and learning was signaled by improvement across the three blocks.

The outcome could be predicted in 84% based on the four cues that were present either alone or in combination. In 84% of the trials the expected answer and the given feedback are the same. In the remaining 16% the expected answer is not the same as the feedback given. In cases like 0110 in which Cue 2 and Cue 3 appear the average predictive value is 50%—that is, participants will probably respond by chance on these trials. There were trials where the feedback was not consistent with the expected answer. If a cue has 90% predictive value for sunshine that means that 1 out of 10 trials will have a feedback contradicting the usual outcome (i.e., RAIN instead of SUNSHINE). Since we did not want to investigate pure one-to-one associative learning, even those trials that presented Cue 1 alone, without any other cues, had to have some cases in which the feedback was RAIN instead of SUNSHINE. Altogether 84% of the feedbacks were predictable from the cues presented.

Strategy fitting

Three different strategies (multicue, singleton, and one-cue) were fit with regression to the performance of each participant. A strategy was assigned to a block of a participant if the average deviation from the expected performance based on that specific strategy did not exceed 0.1, an arbitrary criterion set by earlier literature (Gluck et al., 2002). For trials that had a predictive value (probabilities differing from 50%: all trials except those with Cue 2 and Cue 3 or Cue 1 and Cue 4—two opposing cues—present at the same time without any other cues in the multicue strategy) the square of the difference was calculated (1 or 0). For trials without predictive probabilities the sum of answers was drawn from the half of the trials (since we expect 50% accuracy on these trials), and this value was divided by the number of trials (to get the measure of difference/trial) and raised to the power of two (square of difference by trial). After this, the differences for predictive and nonpredictive trials were summed up and were divided by the complete number of trials (by 50 in each block). This way we were able to get an average deviance per trial. This value was calculated for all strategies—that is, the multicue strategy, the singleton strategy, and the 4 one-cue strategies. A participant was credited with using the multicue strategy if the value of average deviance did not exceed 0.1. If it did, the single strategies were fit, and the best fit strategy was assigned unless average deviance exceeded 0.1 from each the single strategy, in which case there was no strategy assigned to the block.

Note that earlier studies (Hopkins et al., 2004; Shohamy et al., 2004) have shown that from a functional point of view the difference between the singleton and the one-cue strategy is irrelevant (just as the differences between each one-cue strategy are), as these two strategies both build on the same underlying anatomical functioning—that is, both singleton and one-cue strategies rely on the declarative system, while the multicue strategy uses the procedural system.

RESULTS

Accuracy

A 3 (group) × 3 (blocks) repeated measures analysis of variance (ANOVA) was employed with group as between-subjects variable and blocks as within-subject variable to see whether the overall number of correct answers based on the predictive values of all four cues differs between blocks and groups. There was a significant main effect of group, $F(2, 46) = 15.584$, $\eta^2 = .409$, $p < .001$, showing that there is a significant difference between the groups with adults giving the most correct answers, followed by TD children, and children with LI giving the least. Post hoc least significant difference (LSD) tests revealed that the LI group differed from both the TD and adult groups significantly ($p < .001$ in both comparisons, see also Figure 1, Table 1), while performance in the two typically developing groups did not differ ($p = .109$). There was a significant main effect of block, $F(2, 46) = 7.361$, $\eta^2 = .141$, $p < .001$, showing that participants' performance improved with time. The Group × Block interaction did not appear to be significant, $F(4, 46) = 0.882$, $\eta^2 = .038$, $p = .478$ (Figure 2).

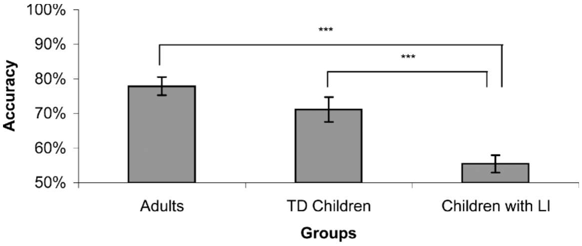


Figure 1. Overall performance on the PCL task by groups (% correct).

TABLE 1
Performance by groups: Descriptive statistics

	Block 1		Block 2		Block 3		Total	
Groups	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*	*Mean*	*SE*
Adults	71.09	3.84	79.56	3.07	83.07	3.46	77.91	2.64
TD children	66.02	4.58	73.31	3.94	74.09	3.68	71.14	3.53
Children with LI	53.91	2.11	54.43	3.70	57.94	3.94	55.43	2.52

Note. In percentages. TD = typically developing. LI = language impairment.

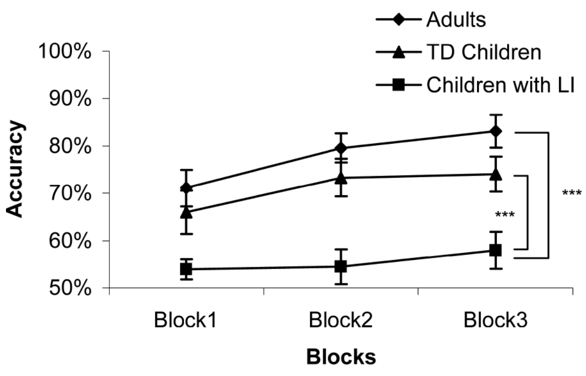


Figure 2. Performance of the three groups on the PLC task by blocks (% correct).

A one-way multivariate ANOVA was employed for three dependent variables (performance on Block 1, performance on Block 2, performance on Block 3) and one between-subjects variable (group). The ANOVA revealed a significant main effect of group on Block 1, $F(2) = 5.824$, $\eta^2 = .206$, $p < .01$. Post hoc LSD tests revealed a significant difference between LI and TD children ($p < .05$) and between LI children and adults ($p < .01$), but the difference between the two control groups was not significant ($p = .332$).

The ANOVA revealed a significant main effect of group on Block 2, $F(2) = 13.303$, $\eta^2 = .372$, $p < .001$. Post hoc LSD tests revealed a significant difference between LI and TD children ($p < .01$) and between LI children and adults ($p < .001$), but the difference between the two control groups appeared to be not significant ($p = .224$).

The ANOVA revealed a significant main effect of group on Block 3, $F(2) = 11.849$, $\eta^2 = .345$, $p < .001$. Post hoc LSD tests revealed a significant difference between LI and TD children ($p < .01$) and between LI children and adults ($p < .001$), and the difference between the two control groups was approaching significance ($p = .093$).

A striking difference between the LI and control groups is that the LI group does not show any evidence of improvement in the first two blocks; they seem to learn only in the third one, when the performance rises up to 57.94% (a performance level that is significantly better than chance; $t = 4.456$, $p < .001$). At the same time both control groups show greater improvement between the first and second blocks. Pairwise comparisons (LSD tests) at the .05 level confirm this difference for the control groups, but the difference is not significant between these blocks for the LI group (Figure 2).

Best fit strategies

Of the pool of 48 participants, altogether 15 switched to multicue strategy. Of the 16 adults, 10 arrived at the multicue strategy, and 5 of the 16 control children also managed to do so, but no children with LI applied this strategy at any point in solving the Weather Prediction Task. Among those who did not switch to multicue strategy, there were 6 adults, 9 control children, and 5 children with LI who used one of the single strategies (i.e., either singleton or one-cue strategy). A total of 2 control children and 11 children with LI failed to show any sign of strategy use (the average deviation exceeded 0.1) A two-way chi-square comparison of strategies by groups revealed a significant difference in strategy use between the groups, $\chi^2(2, N = 48) = 27.146$, $p < .001$ (see Table 2).

TABLE 2
Strategy use by groups

	Adults	*TD children*	*Children with LI*
Switch to multicue	10	5	—
Single strategies	6	9	5
No strategies	—	2	11

Note. TD = typically developing. LI = language impairment. The distribution of strategy use by the three groups differed significantly. The distribution in the TD and LI groups also differed significantly, while there was no difference between TD children and adults.

DISCUSSION

Results show that children with LI perform significantly worse on the Weather Prediction Task than either adults or typically developing children matched on chronological age. It is also clear that children with LI are less able to rely on strategies described by Gluck and colleagues (2002), and even those who seem to develop a one-cue strategy seem to be unable to switch to the more effective multicue strategy. The result that only 5 of the 16 children of the LI group showed any sign of strategy use suggests that children with language impairment have a more fundamental problem with making use of even the simplest single-cue strategies. Their results fall behind age-level expectancies, since typically developing children matched on chronological age learn and perform significantly better. TD children learn at a level closer to, but still significantly below, adult performance; they are less likely to switch to multicue strategy, and they use one of the single strategies instead.

The two control groups perform at a similar level and show the same rate of development across blocks on the Weather Prediction Task. The difference between the overall performance of adults and children seems to differ by only 5%. As can be seen in Figure 2, the difference between the two typically developing groups already appears in the first two blocks (although this difference might grow, since in the third block the performance of the two control groups was approaching significance). As we have explained earlier, early learning is linked to the procedural system, and the difference between children and adults is probably explained by late maturation of the fronto-striatal pathways (Casey, 2005; Thomas et al., 2004).

Our results on the performance of adults are not consistent with that of the Gluck et al. (2002) study. In the present study, 62.5% of adult participants switched to multicue strategy (10 out of 16), whereas in Gluck et al.'s study around 40% of all participants managed to switch to multicue by the 4th block (Figure 7, p. 416). This discrepancy can probably be due to the fact that in our result the predictive values of each cue were higher than that of the Gluck et al. (2002) study.

There is a great difference in accuracy, learning rate, and strategy use between control children and children with LI. Overall performance of LI children lags behind that of control children, and the difference from chance in their performance only reached significance by the end of the third block (and even then, it did not reach 60%). The performance of control children in the first 50 trials is already approximately the level of performance of LI children in the last 50 trials, which shows a clear deficit in the task performance already in the early phases, indicative of impairments in procedural functioning. The difference in accuracy is not the only marker of impaired learning in LI. Strategy analysis shows that children with LI are unable make use of any of the three possible strategies, and even best fit strategies seem to be less efficient and lead to lower performance for them. Besides failing to switch to the multicue strategy, children with LI are also less able to use the single strategies that would be the basis of switching.

As earlier behavioral (Hopkins et al., 2004; Knowlton et al., 1996a; Knowlton et al., 1994) and imaging (Poldrack et al., 2001; Poldrack et al., 1999) results show, the early stages of learning on the Weather Prediction Task, which include the singleton and one-cue strategies, seem to rely on the procedural system, while the later, representational stages of probabilistic categorization require hippocampal activity (Knowlton et al., 1994; Poldrack et al., 1999). This is shown by phenomena that the deficit of the declarative system in amnesia leads to an early learning using singleton and one-cue strategies and later deficits in the multicue strategy (Hopkins et al., 2004; Knowlton et al., 1994). Deficits of the procedural system—that is, Parkinson's syndrome, Huntington's syndrome—show impaired learning at the early stages, and, lacking single-cue strategies, patients do not get the opportunity of switching to multicue strategy.

Results of children with language impairment are similar to results of Parkinson's patients by Knowlton and colleagues (1996a) and Shohamy et al. (2004) for the early phases of learning. Language-impaired children—just like Parkinson's patients—show impaired performance on the Weather Prediction Task, already in its early stages of the first 50 trials. Unlike Parkinson's patients, though, they show very little learning and little evidence of strategy use through three blocks of 50 items. A major difference is that in Shohamy et al.'s study, Parkinson's patients were reported to be able to use single strategies, which were concluded to require declarative memory, and the authors interpreted results for patients with Parkinson's syndrome as an inability to switch to multicue strategy. In the light of other results from the literature, this conclusion is problematic (Hopkins et al., 2004; Knowlton et al., 1996a; Knowlton et al., 1994).

A reason for this inconsistency might be Shohamy and colleagues' (2004) focus on strategy analysis. Strategy analysis requires an arbitrary criterion to be set. Participants who reach the criterion are assigned the strategy. A total of 2 participants who were not assigned any strategies were excluded from the statistical analysis. Since there were only 12 patients with Parkinson's syndrome in the study, the exclusion of 2 patients is enough to lead to the incorrect assumption that patients with Parkinson's syndrome are impaired on later phases of the task, as almost 15% of the clinical group was unable to even start to solve the task. This is especially important considering the fact that the study tested patients with mild Parkinson's syndrome who were relatively well functioning.

Including those patients who did not fit any strategies would probably show a different picture, suggesting that Parkinson's patients are impaired on the early stages of the WP task, which would be consistent with earlier results. As the review of behavioral (Hopkins et al., 2004; Knowlton et al., 1996a; Knowlton et al., 1994) and PET (Poldrack et al., 2001; Poldrack et al., 1999) studies suggested, early phases of the WP task rely on the procedural system, while during the later phases activation shifts towards the declarative system. If Parkinson's patients are less able to fit single strategies, then they are impaired on the earlier phases, which is the procedural part of the task. Amnesic patients use single strategies properly (Hopkins et al., 2004), since these strategies do not require proper functioning of the declarative system. Taken together, these results imply that prototype learning relies on the procedural system, while specifying the individual characteristics and combining the predictive values is a declarative, hippocampus-dependent function. In the light of earlier research, our results indicate that children with LI show a procedural deficit: They do not learn on the initial stages of the WP task, and they also show a more severe inability to use strategies.

If, following Shohamy et al.'s (2004) procedure, children with LI who did not fit any strategies were excluded from our study, we could argue that the focal problem of children with LI on the WP task is that they are not able to switch to the use of the multicue strategy. In our case, it would imply the exclusion of 69% of the LI group. The exclusion of no-fits would lead us to the false conclusion that the central problem of children with LI in solving the Weather Prediction Task is a declarative deficit, manifest in the failure of switching to the multicue strategy.

The finding that children with LI show very little learning and strategy use throughout the task supports Ullman and Pierpont's (2005) hypothesis that children with language impairment have a more general cognitive problem in procedural learning going beyond language and argues against the specificity of language problems in SLI. Whether this deficit is selective to the procedural system, or is complemented by deficits in the declarative system (as suggested by severe vocabulary problems in LI), is the subject of future studies.

The procedural deficit hypothesis (Ullman & Pierpont, 2005) puts language impairment into a broader context by predicting impairments in nonlinguistic abilities, locating these deficits in the domain of procedural learning. Results of the present study are in concert with the PDH and, as such, draw attention to the importance of focusing on impairments outside the language domain as well, both in diagnosis and in training of LI.

Our results also raise a number of other questions suggesting possible lines of further research not only for language impairment, but also for probabilistic category learning. At what age does the difference in performance level and strategy use between children and adults disappear? Would this qualitative difference also disappear in LI children with age at a slower rate than in typical development, or is the qualitative difference maintained? What makes the Weather Prediction Task so much different from artificial grammar learning (Aslin et al., 1999) where children at very early ages can properly differentiate between legitimate and illegitimate sequences solely based on statistical properties of the stimuli? Would we get the same difference between LI and control children on the AGL or SRT tasks? These are yet open research questions with important consequences for both LI and implicit learning.

Original manuscript received 3 February 2009
Revised manuscript accepted 14 April 2009
First published online day month year

REFERENCES

Aczél, B., & Gönci, D. (2005). *Unity or fractionality of implicit learning: A methodological aspect*. Paper presented at CogSci2005—27th Annual Meeting of the Cognitive Science Society, Stresa, Italy, 21–23 July, 2005.

Aram, D., Morris, R., & Hall, N. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research*, *36*, 580–591.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1999). Statistical learning in linguistic and nonlinguistic domains. In B. MacWhinney (Ed.), *Emergence of language*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15.

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

Bishop, D. V. M. (1992). The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *33*, 3–66.

Bishop, D. V. M., & Adams, C. (1992). Comprehension problems in children with specific language impairment—literal and inferential meaning. *Journal of Speech and Hearing Research*, *35*, 119–129.

Casey, B. J. (2005). Frontostriatal and frontocerebellar circuitry underlying cognitive control. In M. I. Posner, U. Mayr, E. Awh, & S. W. Keele (Eds.), *Developing individuality in the human brain: A tribute to Michael I. Posner*. Washington, DC: American Psychological Association.

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

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews*, *30*, 239–271.

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

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory*, *9*, 408–418.

Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109–135.

Hill, E. L. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, *36*, 149–171.

Hopkins, R. O., Myers, C. E., Shohamy, D., Grossman, S., & Gluck, M. (2004). Impaired probabilistic category learning in hypoxic subjects with hippocampal damage. *Neuropsychologia*, *42*, 524–535.

Kas, B., & Lukács, Á. (2007). *Magyar Mondatutánmondási Teszt* [Hungarian Sentence-Repetition Test]. Manuscript in preparation.

Kincses, T. Z., Antal, A., Nitsche, M. A., Bartfai, O., & Paulus, W. (2004). Facilitation of probabilistic classification learning by transcranial direct current stimulation of the prefrontal cortex in the human. *Neuropsychologia*, *42*, 113–117.

Knopman, D., & Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia*, *29*, 245–254.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996a). A neostriatal habit learning system in humans. *Science*, *273*, 1399–1402.

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science*, *262*, 1747–1749.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic category learning in amnesia. *Learning & Memory*, *1*, 106–120.

Knowlton, B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., Swenson, M., & Butters, N. (1996b). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology*, *10*, 538–548.

Leonard, L. B. (1997). *Children with specific language impairment*. Cambridge, MA: The MIT Press.

Meulemans, T., van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, *69*, 199–221.

Nosofsky, R. M., Stanton, R. D., & Zaki, S. R. (2005). Procedural interference in perceptual classification: Implicit learning or cognitive complexity? *Memory & Cognition*, *33*, 1256–1271.

Plante, E., Gómez, R., & Gerken, L. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*, *35*, 453–462.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Moyano, J. C., Myers, C., et al. (2001). Interactive memory systems in the human brain. *Nature*, *414*, 546–550.

Poldrack, R. A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience and Biobehavioral Reviews*, *32*, 197–205.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, *13*, 564–574.

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

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

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 581–597.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. P. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton Century Croft.

Rüsseler, J., Gerth, I., & Munte, T. F. (2006). Implicit learning is intact in adult developmental dyslexic readers: Evidence from the serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology*, *28*, 808–827.

Shohamy, D., Myers, C. E., Onlaor, S., & Gluck, M. A. (2004). Role of the basal ganglia in category learning: How do patients with Parkinson's disease learn? *Behavioral Neuroscience*, *118*, 676–686.

Siegert, R. J., Taylor, K. D., Weatherall, M., & Abernethy, D. A. (2006). Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology*, *20*, 490–495.

Smith, J. G., & McDowall, J. (2006). When artificial grammar acquisition in Parkinson's disease is impaired: The case of learning via trial-by-trial feedback. *Brain Research*, *1067*, 216–228.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology*, *44*, 453–495.

Stoodley, C. J., Harrison, E. P. D., & Stein, J. F. (2006). Implicit motor learning deficits in dyslexic adults. *Neuropsychologia*, *44*, 795–798.

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

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y. H., et al. (2004). Evidence of developmental differences in implicit sequence learning:

An fMRI study of children and adults. *Journal of Cognitive Neuroscience*, *16*, 1339–1351.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development*, *3*, 269–293.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., et al. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, *9*, 266–276.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*, 399–433.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 930–933.

Vargha-Khadem, F., Watkins, K. E., Price, C. J., Ashburner, J., Alcock, K. J., Connelly, A., et al. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 12695–12700.

Voermans, N. C., Petersson, K. M., Daudey, L., Weber, B., van Spaendonck, K. P., Kremer, H. P. H., et al. (2004). Interaction between the human hippocampus and the caudate nucleus during route recognition. *Neuron*, *43*, 427–435.

Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002a). Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain*, *125*, 452–464.

Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., et al. (2002b). MRI analysis of an inherited speech and language disorder: Structural brain abnormalities. *Brain*, *125*, 465–478.

Whitehurst, G. J., Arnold, D. S., Smith, M., Fischel, J. E., Lonigan, C. J., & Valdezmenchaca, M. C. (1991). Family history in developmental expressive language delay. *Journal of Speech and Hearing Research*, *34*, 1150–1157.

Witt, K., Nuhsman, A., & Deuschl, G. (2002). Intact artificial grammar learning in patients with cerebellar degeneration and advanced Parkinson's disease. *Neuropsychologia*, *40*, 1534–1540.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*, 464–476.

Study 4.

Lukács, Á., Leonard, L. B., Kas, B. (2011) "The dog chase the cat": grammaticality judgments by Hungarian-speaking children with language impairment. *Acta Linguistica Hungarica*, 58:1–2, 24–38. doi: 10.1556/ALing.58.2011.1–2.2

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6)

Thesis 2. Agreement deficits in SLI are better explained by processing difficulties than by a selective grammatical impairment targeting agreement. (1, 4)

Acta Linguistica Hungarica, Vol. 58 (1–2), pp. 24–38 (2011)
DOI: 10.1556/ALing.58.2011.1–2.2

"THE DOG CHASE THE CAT": GRAMMATICALITY JUDGMENTS BY HUNGARIAN-SPEAKING CHILDREN WITH LANGUAGE IMPAIRMENT*

ÁGNES LUKÁCS[a, b] – LAURENCE B. LEONARD[c] – BENCE KAS[b, d]

[a]Department of Cognitive Science,
Budapest University of Technology and Economics, Hungary

[b]Research Institute of Linguistics, Hungarian Academy of Sciences, Hungary

[c]Purdue University, USA

[d]Bárczi Gusztáv Faculty of Special Pedagogy, Eötvös Loránd University, Hungary

Corresponding author: Ágnes Lukács, address: Stoczek u. 2.,
H–1111 Budapest, Hungary, e-mail: alukacs@cogsci.bme.hu

Abstract: In a previous study of language production, a group of Hungarian-speaking children with language impairment (LI) committed a larger number of errors than typically developing peers on verb inflections that mark person, number, tense, and definiteness (Lukács et al. 2009b). However, the error forms produced often differed from the correct form by only a single dimension (e.g., person, number, tense, or definiteness) with no single dimension proving consistently problematic. In the present study, we sought to determine whether a similar pattern applied to the children's understanding of verb inflections, as reflected in a grammaticality judgment task. We compared the performance of 17 Hungarian-speaking children with language impairment (LI) between ages 8;0 and 11;9 with typically developing children between 6;10 and 11;1 years individually matched on receptive vocabulary raw scores (VC) and also to a control group of children matched on chronological age (AC; between 8;1-12;1). We obtained grammaticality judgments for 68 sentences, including 56 ill-formed sentences that contained a single error of person, number, tense, definiteness, or morphophonology. As the AC group performed at ceiling, the analysis focused on comparisons between the LI and VC groups. Besides comparing accuracy scores in the two groups, we tested how well performance could be predicted by a test of grammatical comprehension (TROG) and a measure of

* This research was supported by research grant R01 DC00458 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (USA) to Laurence B. Leonard and by OTKA TS 049840 from the Hungarian National Science Foundation to Csaba Pléh. Ágnes Lukács was a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science. The authors are grateful to the children in the Dr Nagy László Institute of Special Education in Kőszeg for their participation. We also thank the speech therapists in both institutions for their help with screening and organization and Miklós Győri for his helpful comments on the manuscript.

dc_1365_16

nonword repetition ability obtained prior to the administration of the grammaticality judgment task. There were no significant group differences in the accuracy of grammaticality judgments. Both groups recognized well-formed sentences, and agreement errors of number, person or definiteness, significantly more accurately than tense or morphophonological errors. Although there was no difference between performance levels of the LI and the VC groups, we found differences between the two groups in the types of measures that were most closely tied to performance on the grammaticality judgment task. Performance in the LI group was strongly associated with nonword repetition span, while in the VC group, TROG performance was associated with grammaticality judgment performance. These results suggest that the same level and pattern of performance can be supported by different background mechanisms in typical and atypical language development.

Keywords: language impairment, Hungarian, agreement, grammaticality judgments

1. Introduction

Many children with language impairment (LI) exhibit significant limitations in their use of verb inflections. These limitations are especially striking in Germanic languages, but seem somewhat less dramatic in languages with a rich verb inflection system such as Italian or Spanish. The types of verb inflection errors produced by children with LI also seem to vary according to the language being acquired. In English, German, Dutch, Swedish, and French, correct use of finite verb inflections seems to alternate with the inappropriate use of infinitive forms in finite verb contexts (Rice–Wexler 1996; Leonard et al. 1997; Oetting–Horohov 1997; Marchman et al. 1999; Norbury et al. 2001; Redmond 2003; Rice et al. 1997; Jong 1999; Leonard et al. 2004; Paradis–Crago 2001). In English, this is seen when children with LI alternate between, for example, *Mommy likes ice cream* and *Mommy like ice cream.* These errant productions do not appear to be simple omissions. In English, infinitives are bare stems. In languages that employ overt inflections for infinitives, the substitutions take the form of an overt infinitive inflection replacing an overt finite inflection (e.g., *spring-a* 'to run' in place of *spring-er* 'runs' in Swedish).

In contrast, in null-subject languages with a rich verb inflection system (like Hungarian), most errors are substitutions of one finite inflection in place of another. Furthermore, the majority of these substitutions can be characterized as 'near misses'. For example, in Italian, children with LI are more likely to use a first person singular inflection in place of a first person plural inflection than to use a third person singular inflection. Yet, the latter will be the most frequent substitute in contexts that require use of a third person plural inflection. When tense as well as person and number are marked on the verb, these near misses are even more apparent (Bortolini et al. 1997). For example, in Spanish, a child with LI might produce a third person singular past form or a third person plural present form in contexts requiring a third person plural past form (Bedore–Leonard 2001).

We recently reported a similar finding for Hungarian-speaking children with LI (Lukács et al. 2009b). Hungarian was an especially valuable language to study because verbs are inflected not only for person, number, and tense, but also for definiteness. Whereas the verb must agree with the subject in person and number, it must agree with the direct object in definiteness. Thus, four dimensions must be considered in selecting the appropriate verb inflection. We found that the children with LI were less accurate than typically developing peers. However, all 24 inflections examined were used by the children both in correct contexts and as (incorrect) substitutes for other inflections. Errors that differed from the correct form by a single dimension constituted the most frequent error, even though by randomly selecting a suffix, differences in two or three dimensions actually had higher probabilities, given the verb inflection system of Hungarian. Strikingly, no substitute differed from the correct form by four dimensions even though all 24 inflections were occasionally used as a substitute at some point.

Along with errors of person, number, tense, or definiteness, the children with LI also produced errors in morphophonology, expressing the correct set of features but using an incorrect allomorph in doing so. In a subsequent study, we noted that Hungarian-speaking children with LI also have difficulty with the morphophonology of noun use, as when the children sometimes correctly expressed the plural and accusative case of a noun but failed to alter the phonological form of the stem to accommodate these inflections as is required for many nouns (Lukács et al. 2009a).

The pattern of errors seen in the productions of Hungarian-speaking children with LI clearly reflect a significant degree of grammatical knowledge on the children's part. Without such knowledge, substitutions would be haphazard, or the children would resort to the overuse of a default form. Instead, most errors approximated the correct form. We attributed the errors of the children with LI to processing limitations. Given a relatively large number of dimensions (person, number, tense, and definiteness) that had to be considered for the retrieval of the correct form, these children sometimes retrieved a form that was similar to the correct form in its composition, yet differed in one detail, with no particular dimension standing out as highly problematic. The purpose of the present study was to further evaluate this assumption of processing limitations as a basis for the verb inflection errors that are seen in Hungarian-speaking children with LI.

If processing ability is severely taxed when children with LI must consider a relatively large number of dimensions, the resulting difficulties should not be confined to production in the moment. These difficulties should adversely affect the degree to which the inflections are learned and incorporated into the children's grammars. When children hear inflections in the input, they must hypothesize the dimensions that these inflections reflect. Inflections that require multiple dimensions to be considered require more processing resources and, for this reason, will be learned more slowly if these resources are limited. As a result, the strength of the representations of these inflections in the grammars of the children (i.e., the degree to which they are learnt and incorporated into the child's grammar) will be lower than the representations of the same inflections in the grammars of typically developing children. Weaker representations are likely to be more difficult to retain in comprehension tasks as well as more difficult to retrieve in production tasks. In the present study, we employ a grammaticality judgment task to test this assumption. Specifically, we present children with both well-formed sentences and sentences that are ungrammatical in a single dimension. If the processing limitation view is correct, children with LI should occasionally miss the errors in the ungrammatical sentences but show no extraordinary difficulty with any particular dimension. Their performance profile across error types should approximate that of a group of younger typically developing children.

A second goal of the present study was to see whether measures of grammatical comprehension and verbal short-term memory predict accuracy in judging the grammaticality of sentences in children with LI and in typically developing children to the same degree. We assume that measures of grammatical comprehension such as those found in picture-pointing tasks would serve as a significant predictor for typically developing children's success in distinguishing grammatical from ungrammatical sentences. The amount of variance would probably be due to differences in the task employed (such as picture-pointing versus grammaticality judgment) and the particular composition of items in the two tasks. However, the general type of skill likely to be the most relevant in both cases is the child's understanding of grammar.

For children with LI, other types of predictors may also prove important. If, as we have assumed, processing difficulties are involved in children's language impairment, measures of such skills might account for unique variance in the children's grammaticality judgment performance. This does not imply that processing abilities do not play a role in typical development (TD), but we expect that the task we employed would not tax TD children's processing abilities to a great extent at this age.

In the present study, we use a test of nonword repetition along with a grammatical comprehension measure as predictors of the children's success in judging the grammaticality of sentences that differ in accuracy (grammatical, ungrammatical) and type of error (error of person, number, tense, definiteness, morphophonology). Nonword repetition places demands on verbal short term memory, an area of processing that is often found to be vulnerable in children with LI (see e.g., Bishop et al. 2006; also see e.g., Archibald–Gathercole 2007 on a more complex approach and the need for further specifications). For a language such as Hungarian, such a measure may prove especially revealing because inflected Hungarian nouns and verbs often involve extended phonological sequences that must conform to rather complex morphophonological rules. Retention of these sequences would seem to be a prerequisite to learning and comprehension as well as to retrieval for production.

2. Materials and methods

2.1. Participants

The experimental group consisted of 17 Hungarian-speaking, monolingual children diagnosed with language impairment between ages 8;0 and11;9. Inclusive criteria for the language impaired group were significant deviation from age norms (−1.5 SD) on two out of the following four language tests: a test of grammatical comprehension (Test for the Reception of Grammar, TROG, Bishop 1983), a test of receptive vocabulary (Peabody Picture Vocabulary Test, PPVT, Csányi 1974), a sentence repetition test (Hungarian Sentence Repetition Test/Magyar Mondatutánmondási Teszt, MAMUT, Kas–Lukács in preparation) and a test of nonword-repetition (Racsmány et al. 2005). The four screening tests were selected partly for practical reasons (these are the only language tests in Hungarian that have either age norms or data from a large sample of children available), but they also have theoretical motivation: they focus on specific functions that are systematically found to be impaired in SLI (grammatical comprehension (TROG), vocabulary (PPVT), verbal short term memory (nonword repetition) and grammatical production (sentence repetition; also taxing verbal STM). Children with an IQ below 85 (Raven et al. 1987), or a history of hearing impairment or any neurological conditions were excluded from the study.

Performance of the LI group was compared to two control groups, one matched individually on receptive vocabulary (PPVT) scores, the other matched individually on chronological age. Data for the three groups are summarized in Table 1.

Table 1

Data for the language impaired (LI) and the control groups matched on vocabulary scores (VC) and on chronological age (AC). Significant differences are shown in shading (both significant at $p < 0.001$; nonword repetition: $F(1, 33) = 27.64$, $\eta^2 = 0.515$; sentence repetition: $F(1, 33) = 16.78$, $\eta^2 = 0.351$)

	LI	VC	AC
N:	17	17	17
Mean age (range)	10;2 (8;0–11;9)	7;11 (6;10–11;1)	10;2 (8;1–12;1)
Mean PPVT score (range)	105.00 (77–150)	103.47 (76–150)	
Mean TROG scores	71.35	72.31	
Mean nonword repetition scores	3.76	6.18	
Mean sentence repetition	27.06	36.94	
Mean LAPP scores (productive vocabulary)	34.82	33.69	

Statistical comparisons confirmed that the LI and VC groups were not only highly similar in receptive vocabulary but also in their performance on the TROG. Note that it would be misleading to conclude that the children with LI had no deficit in vocabulary and grammatical comprehension; their scores were matched to those of TD children who were more than two years younger on average. Scores on a vocabulary production measure are also provided in Table 1. The LI group and younger TD children did not differ on this measure. On the other hand, differences favoring the younger TD children were seen on both the sentence repetition measure and the nonword repetition measure.

2.2. Method

2.2.1. Grammaticality judgments

We tested sensitivity to morphological errors by reading sentences to children, and asking them to judge the grammatical well-formedness of the sentence according to the following instruction: "Small children often make errors when they speak. Did you know that? I know a two year-old, and now I am going to read you some of what he said. After each sentence, tell me whether he said it right or not. If he made an error, also correct the sentence he said."

Participants had to tell whether the sentence was ok or not ok, and if their answer was no, we asked them to correct them. Sentences were depicting imaginary scenarios—describing actions carried out by animals. The main focus of the study was sensitivity to agreement errors. Only errors on a single dimension were included, and these were complemented by a set of sentences with morphophonological errors to test performance on non-agreement errors as well.

The test battery contained 68 sentences, in the following categories:

(1) well-formed sentences ($N = 12$)

Az oroszlán kergeti a lovat.
'The lion chase-pres.3sg.def the horse'

(2) sentences with agreement errors ($N = 48$)

(a) definiteness errors ($N = 16$)
*A majmok mostak a hintát.
'The monkeys wash-past.3pl.indef the swing'

(b) person errors ($N = 16$)
*A nyulak építetek egy várat.
'The rabbits build-pres.2pl.indef a castle'

(c) number errors ($N = 8$)
*A tehenek épít egy alagutat.
'The cows build-pres.3sg.indef a tunnel'

(d) tense errors ($N = 8$)
*Tegnap a kutyák tolnak egy ágyat.
'Yesterday the dogs push-pres.3pl.indef a bed'

(3) sentences with morphophonological errors ($N = 8$)

*Az oroszlán a toronyt építi.
'The lion builds the tower-acc' (grammatical: *tornyot*)

2.2.2. Comprehension of grammatical structures

The Hungarian adaptation of the original TROG (Bishop 1983) is being standardized (by Ágnes Lukács, Miklós Győri and Sándor Rózsa) on children from 4 to 12 years of age.[1] Items assess the children's comprehension of increasingly more difficult grammatical structures. The test consists of 20 blocks, each with 4 sentences of the same construction (such as sentences with comparatives, postmodified subjects and embedded clauses). The test has a booklet containing 80 pages, each with 4 pictures, and on each page the child must point to the picture that matches the sentence spoken by the experimenter. A block is considered completed if the child responds correctly to all 4 pictures in the block. Performance is measured in terms of number of blocks correctly completed.

2.2.3. Nonword repetition

The nonword repetition test (Racsmány et al. 2005) requires the repetition of meaningless but phonotactically licit strings of Hungarian phonemes. The test contains 36 nonwords between 1 and 9 syllables in length. Each length is represented by 4 nonwords. The phonological structure of the nonwords does not reflect frequency distributions of Hungarian phoneme sequences, but the test avoids sequences that would be articulatorily difficult for speakers. The span of the participant is the highest syllable number for which s/he could correctly repeat at least 2 out of the 4 nonwords.

3. Results

3.1. Grammaticality judgments

A judgment was considered correct if it involved accepting a correct sentence or rejecting an ungrammatical one. Rejecting an ungrammatical sentence was scored as correct even if the child could not correct the sentence. The children's performance is summarized in Figure 1, based on percentage correct for each item type. As can be seen from the figure,

[1] We thank Professor Dorothy Bishop for providing us with the TROG for this purpose. Thus far, 600 typically developing children have been seen as part of the norming process; the scores for the children with LI were compared against the values obtained for the typically developing children.

the AC group performed at very high levels of accuracy. An analysis of variance (ANOVA) revealed a significant main effect for group, $(2, 48) = 14.26$, $\eta^2 = 0.343$, $p < 0.001$, with the AC group showing significantly greater accuracy than the two remaining groups.

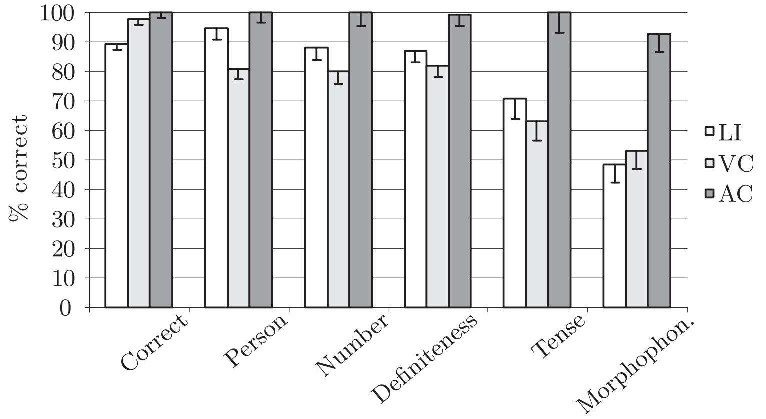


Fig. 1

Performance of the LI, VC and AC groups on different types of errors in the grammaticality judgment task

For the main analysis for accuracy we calculated A' for agrammatical sentence types to adjust for a possible bias of children accepting sentences rather than rejecting them (cf. Linebarger et al. 1983).[2] The A' data were then analyzed using a general linear model ANOVA with Group as a between-subjects factor and Error type (fully grammatical, tense error, definiteness error, person error, number error, morphophonological error) as a within-subjects factor. A summary of the A' findings appears in Figure 2. Again, there was a main effect for Group, $F(2, 46) = 18.04$, $\eta^2 = 0.440$, $p < 0.001$, owing to the greater accuracy on the part of the AC group. Because the ceiling level performance of these children distorted the data, the ANOVA was re-calculated after excluding this group. The

[2] Following Rice et al. (1999), we used the formula described in Linebarger et al. (1983) to calculate scores: $A' = 0.5 + (y-x)(1+y-x)/4y(1-x)$ where y represents the correct judgements of grammatical sentences ("hits") and x the incorrect judgements of ungrammatical sentences ("false alarms"). A strong tendency to reject sentences will result in an A' value approximately around 0, a tendency to accept sentences result in an A' value of around 0.5. An A' value close to 1.0 shows good discrimination between grammatical and ungrammatical.

dc_1365_16

LI group maintained 17 participants, but only 15 children were included in the VC group, because of divisions by zero. The ANOVA did not show a significant main effect for Group, $F(1, 30) = 2.742$, $\eta^2 = 0.084$, n.s. Error type had a significant main effect, $F(4, 120) = 19.79$, $\eta^2 = 0.397$, $p < 0.001$, but the interaction of Group and Error type was not significant $F(4, 120) = 1.17$, $\eta^2 = 0.038$, n.s. Bonferroni-adjusted pairwise comparisons showed that the performance on the following error types differed significantly. Morphophonological errors were the most difficult to detect for both groups, with A' scores for this item type lagging significantly behind definiteness errors ($p < 0.001$), number errors ($p < 0.001$), and person errors ($p < 0.001$), but they did not differ from tense errors. Tense errors were significantly more difficult than number ($p < 0.05$) or person ($p < 0.01$) errors. All other pairwise differences were nonsignificant.

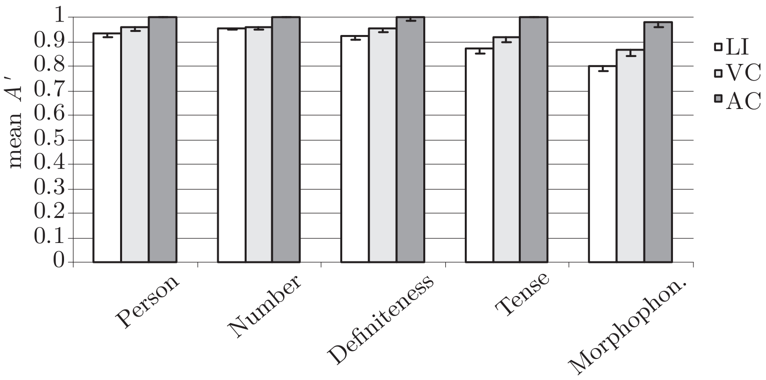


Fig. 2

A' values of the LI, VC and AC groups on different types of errors in the grammaticality judgment task

3.2. Potential predictors of grammaticality judgment performance

To determine whether grammatical comprehension as measured by the TROG or nonword repetition ability were associated with the children's accuracy in making grammaticality judgements of verb inflections, we included them in stepwise regression analyses. Only variables that showed a significant correlation ($p < 0.05$) with the target variable were entered into the analysis. Table 2 shows results for the two groups for the VC and LI groups. The ceiling level performance of the AC group obviated use of regression analyses for these children. For the VC group, TROG scores modeled performance best, explaining 38.5% of variance in the data, while nonword repetition did not show a significant correlation with GJ performance. For the LI group, nonword repetition span proved to be the best model, explaining 34.3% of variance in the data; for this group, TROG scores did not show a signfiicant correlation with GJ performance.

Table 2

Models of performance by individual differences in the LI and VC groups

		Beta	Sig	R^2
VC	TROG	0.653	<0.01	0.385
LI	Nonword repetition	0.619	<0.01	0.343

4. Discussion

The LI group showed significantly lower accuracy in their grammaticality judgments than their same-age TD peers. Relative to the VC group, the children with LI were similar not only in their overall accuracy, but also in their profile of performance across item types. However, the LI and VC groups differed in the factors that were predictive of performance on the grammaticality judgment task. For the VC group, a grammatical comprehension measure explained a significant amount of variance. For the LI group, in contrast, nonword repetition ability proved to be the primary predictor. Thus, despite the similar accuracy levels of the LI and VC groups, the factors proving predictive differed, with no overlap.

These results seem most compatible with a processing limitation view of LI, for several reasons. First, the children with LI showed performance levels that were very similar to those of TD children (viz., the VC group) who were approximately two years younger than the children with LI. The similar performance profile across item types for these two groups suggests that the children with LI were not disproportionately weak in select areas of grammar. Instead, they were only relatively weak in the same areas that proved weakest in the VC group as well. Importantly, errors were distributed across all item types; the children with LI occasionally accepted sentences containing a person error, a number error, a definiteness error, a tense error, or a morphophonological error. Such errors resemble the "near-miss" errors reported for Hungarian-speaking children with LI in production (Lukács et al. 2009b). These errors were productions of a form that differed from the target form by only one dimension, but the particular dimension in error varied from item to item.

For the children with LI, as with the children in the VC group, items containing tense errors were more likely to be incorrect than most other item types. This relative difficulty might also have an explanation in terms of processing demands. Note that in the items with errors of person, number, or definiteness, the violation was proximal. That is, either the subject and immediately following verb differed in person or number, or the verb and the immediately following direct object differed in definiteness. In contrast, for items containing a tense error, the entire subject–verb–direct object sequence showed correct agreement. The sentence was ungrammatical because the temporal adverb preceding the subject indicated that past, rather than present tense should be used with the verb. Such a violation was distal and therefore required retention of the temporal information appearing in sentence-initial position to determine that the verb (appearing after the subject) was not in the proper form.

The factor serving as a significant predictor of the LI group's grammaticality judgment accuracy—nonword repetition ability—can also be interpreted within a processing limitation framework. Nonword repetition requires the retention of sound sequences. Hungarian is a language that involves the detection and retention of sound sequences to a greater extent than most languages. Specifically, attached to the verb stem are tense and agreement inflections that reflect a variety of morphophonological patterns. Adult-like ability requires not only an attention to grammatical accuracy, but also to which allomorphs must be used and how they must be modified according to the phonological context. Given that the children with LI made more errors on item types that contained morphophonological errors than on item types containing local agreement errors, it is clear that they were far from mastery levels in their grasp of the morphophonology of their language. For these children, it is possible that difficulty with the retention of sound sequences was part of their problem. However, nonword repetition did not serve as a good predictor of the typically developing group's grammaticality judgements. A plausible reason for this is the significantly higher average level and smaller within-group variability of nonword repetition abilitiy. It seems that the VC children have reached the level of phonological processing ability required for this specific sentence processing task while the LI children as a group have not. Yet VC children did not outperform the LI children in the experimental task, as would be predicted on the basis of the difference in nonword repetition ability. We assume that there is another factor, namely, school routine, that might contribute to the results and represent a counter-balance in favour of the LI children. Judging the grammaticality of sentences requires not only the processing of the sentence, but also a conscious reflection on the structure, that is, a kind of metalinguistic consideration. Children must have a notion of linguistic error and a routine in recognizing and correcting them. Since this kind of task typically emerges in school settings, and the VC group consisted of children who were on average two and a half year younger than the LI children, this difference in school experiences might have counterbalanced the effect of differences in phonological ability.

We are somewhat surprised that the grammatical comprehension measure (TROG) was not a significant predictor of the grammaticality judgment performance of the children with LI, as it was for the VC group. Even if difficulty with sound sequences was an important factor for the LI group, we had assumed that the grammatical nature of the TROG and our judgment task would result in a stronger relationship between the two measures. To be sure, the difference in tasks (picture-pointing versus judgment) would reduce the amount of variance that could have been explained by the grammatical comprehension measure.

It is also true that the item emphasis of the two measures differed to a considerable degree. Our grammaticality judgment task emphasized tense, agreement, and morphophonology, whereas the emphasis of the TROG is on syntactic structure (e.g., postmodified subjects, center-embedded relative clauses, comparative structures etc.). Given the relationship between the two measures seen for the TD children, it appears that one common route in learning is one in which language structure and morphosyntax/morphophonology are learned as complementary components of grammar or in which progress in one component is used to benefit the other.

Although our data do not allow conclusions on developmental pathways, we speculate that for the children with LI, a different learning route seems possible. Specifically, it appears that these children's learning of morphosyntax/morphophonology was more tied to sequential phonological information than to (syntactic) structural information. This may be an alternative route to grammatical learning that is not unique to the LI population. However, given that this route has thus far only been associated with a group of children diagnosed with language learning problems, additional research is needed to determine whether a relationship between memory for sound sequences and grammatical inflection ability might serve as a clinical marker of language impairment, or whether it simply reflects one of several learning routes that any child—typical or language impaired—might adopt.

References

Archibald, Lisa M. D. – Susan E. Gathercole 2007. Nonword repetition and serial recall: Equivalent measures of verbal short-term memory? In: Applied Psycholinguistics 28 : 587–606.

Bedore, Lisa – Laurence B. Leonard 2001. Grammatical morphology deficits in Spanish-speaking children with specific language impairment. In: Journal of Speech, Language, and Hearing Research 44 : 905–24.

Bishop, Dorothy V. M. 1983. Test for the reception of grammar. The author's publication.

Bishop, Dorothy V. M. – Caroline Adams – Courtenay F. Norbury 2006. Distinct genetic influences on grammar and phonological short-term memory deficits: Evidence from 6-year-old twins. In: Genes, Brain and Behavior 5 : 158–69.

Bortolini, Umberta – Cristina M. Caselli – Laurence Leonard 1997. Grammatical deficits in Italian-speaking children with specific language impairment. In: Journal of Speech, Language and Hearing Research 40 : 809–20.

Csányi, F. Yvonne 1974. Peabody szókincsteszt [Peabody Vocabulary Test]. Bárczi Gusztáv Gyógypedagógiai Főiskola, Budapest.

Jong, Jan de 1999. Specific language impairment in Dutch: Inflectional morphology and argument structure (Groningen Dissertations in Linguistics 28). Groningen University, Groningen.

Kas, Bence – Ágnes Lukács in preparation. Magyar Mondatutánmondási Teszt [Hungarian Sentence Repetition Test]. Ms.

Leonard, Laurence B. – Julia A. Eyer – Lisa M. Bedore – Bernard G. Grela 1997. Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. In: Journal of Speech, Language and Hearing Research 40 : 741–53.

Leonard, Laurence B. – Kristina Hansson – Ulrika Nettelbladt – Patricia Deevy 2004. Specific language impairment in children: A comparison of English and Swedish. In: Language Acquisition 12 : 219–46.

Linebarger, Marcia – Myrna Schwartz – Eleanor Saffran 1983. Sensitivity to grammatical structure in so-called agrammatic aphasics. In: Cognition 13 : 361–92.

Lukács, Ágnes – Laurence B. Leonard – Bence Kas 2009a. The use of noun morphology by children with language impairment: The case of Hungarian. In: International Journal of Language & Communication Disorders 6 : 1–26.

Lukács, Ágnes – Laurence B. Leonard – Bence Kas – Csaba Pléh 2009b. The use of tense and agreement by Hungarian-speaking children with language impairment. In: Journal of Speech, Language, and Hearing Research 52 : 98–117.

Marchman, Virginia – Beverly Wulfeck – Susan Ellis Weismer 1999. Morphological productivity in children with normal language and SLI: A study of the English past tense. In: Journal of Speech, Language, and Hearing Research 42 : 206–19.

Norbury, Courtenay F. – Dorothy V. M. Bishop – Josie Briscoe 2001. Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. In: Journal of Speech, Language, and Hearing Research 44 : 165–78.

Oetting, Janna B. – Janice Horohov 1997. Past-tense marking by children with and without specific language impairment. In: Journal of Speech, Language, and Hearing Research 40 : 62–74.

Paradis, Johanne – Martha Crago 2001. The morphosyntax of specific language impairment in french: An extended optional default account. In: Language Acquisition 9 : 269–300.

Racsmány, Mihály – Ágnes Lukács – Dezső Németh – Csaba Pléh 2005. A verbális munkamemória magyar nyelvű vizsgálóeljárásai [Hungarian tools for examining verbal working memory]. In: Magyar Pszichológiai Szemle 4 : 479–505.

Raven, John – John Court – John C. Raven 1987. Raven's progressive matrices and Raven's colored matrices. H. K. Lewis, London.

Redmond, Sean 2003. Children's productions of the affix *-ed* in past tense and past participle contexts. In: Journal of Speech, Language, and Hearing Research 46 : 1095–109.

Rice, Mabel L. – Karen Ruff Noll – Hannelore Grimm 1997. An extended optional infinitive stage in German-speaking children with specific language impairment. In: Language Acquisition 6 : 255–95.

Rice, Mabel L. – Kenneth Wexler 1996. Toward tense as a clinical marker of specific language impairment in English-speaking children. In: Journal of Speech, Language, and Hearing Research 39 : 1239–57.

Rice, Mabel L. – Kenneth Wexler – Sean M. Redmond 1999. Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. In: Journal of Speech, Language, and Hearing Research 42 : 943–61.

Study 5.

Leonard, L. B, Lukács, Á., Kas, B. (*2012*) Tense and Aspect in Childhood Language Impairment: Contributions from Hungarian. *Applied Psycholinguistics, 33,* 305–328. doi: 10.1017/S0142716411000361

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6)

Thesis 4. Difficulties in aspect marking in production but not in comprehension in past tense forms suggest a processing problem instead of a selective impairment of aspect marking (5)

Applied Psycholinguistics **33** (2012), 305–328
doi:10.1017/S0142716411000361

Tense and aspect in childhood language impairment: Contributions from Hungarian

LAURENCE B. LEONARD
Purdue University

ÁGNES LUKÁCS
Budapest University of Technology and Economics and Hungarian Academy of Sciences

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

Received: August 20, 2009 Accepted for publication: September 6, 2010

ADDRESS FOR CORRESPONDENCE
Laurence B. Leonard, Department of Speech, Language, and Hearing Sciences, 500 Oval Drive, Heavilon Hall, Purdue University, West Lafayette, IN 47907. E-mail: xdxl@purdue.edu

ABSTRACT
Previous studies of children with language impairment (LI) reveal an insensitivity to aspect that may constitute part of the children's deficit. In this study, we examine aspect as well as tense in Hungarian-speaking children with LI. Twenty-one children with LI, 21 TD children matched for age, and 21 TD children matched for receptive vocabulary scores were tested on their comprehension and production of both imperfective and perfective verb forms in past tense contexts. Although the groups did not differ in their comprehension performance, the children with LI were less accurate than both comparison groups in producing both imperfective and perfective forms. Based on these results, it appears that children with LI have difficulties selecting the appropriate aspectual marking in past tense contexts.

A problem in the expression of tense is among the most ubiquitous findings in the literature on children with language impairment (LI). This is decidedly the case for English-speaking children with LI; these children use present and past tense forms less consistently than younger typically developing (TD) children as well as same-age TD peers (Leonard, Eyer, Bedore, & Grela, 1997; Marchman, Wulfeck, & Ellis Weismer, 1999; Norbury, Bishop, & Briscoe, 2001; Oetting & Horohov, 1997; Redmond, 2003; Rice & Wexler, 1996). In other Germanic languages that mark present and past tense through inflection, such as Swedish, tense is also problematic for children with LI (Hansson, Nettelbladt, & Leonard, 2000). Studies of languages that express past tense through auxiliary + participle constructions, as in German and Dutch, show that children with LI often omit the auxiliary, thus rendering tense unspecified (de Jong, 1999; Rice, Noll, & Grimm, 1997). Even studies of Romance languages such as Italian and French that employ the auxiliary + participle construction for past tense reveal problems with the expression of tense on the part of children with LI, due to frequent omission of the auxiliary (Leonard, Sabbadini, Leonard, & Volterra, 1987; Paradis & Crago, 2001).

The present study is concerned with children with LI who are acquiring Hungarian. As will be seen, this language permits an assessment of tense as well as another temporally related detail of grammar, aspect. Following a review of tense and aspect in the current LI literature, we discuss the contributions of studying Hungarian-speaking children with LI.

One prominent account of problems with tense is that children with LI are especially slow in developing the principle that tense is obligatory in main clauses (Rice, 2003; Rice & Wexler, 1996, 2001; Rice, Wexler, & Hershberger, 1998). According to this view, tense is represented in the grammars of children with LI so that when tense is expressed it is used appropriately. The problem rests in the children's failure to grasp that tense must always be used in a sentence. For Germanic languages, this account holds for both auxiliaries and verb inflections that mark tense; for most Romance languages and null-subject languages more generally, this account holds for auxiliaries in particular (Wexler, 1998, 2003).

Recently, studies have appeared that also raise questions about these children's command of the temporal details pertaining to tense. In languages such as English, tense and aspect interact in important ways, and young TD children do not appear to keep these notions distinct in their early interpretation of sentences. Tense refers to the temporal relationship between the time of the event being described and some other time, usually the time of speaking. Thus, a sentence such as *She read the newspaper article* refers to an event that occurred before the time of speaking. Aspect differs from tense in referring to the temporal distribution of an event independent of the event's location in time (Comrie, 1976). For example, *She was reading the newspaper article* and *She is reading the newspaper article* differ in tense but both refer to an event (reading) that extends across some period of time. In contrast, two sentences can employ the same lexical verb and tense but can differ in aspect. The sentence *She read the newspaper article (and enjoyed it)* indicates not only an event in the past but also that the event was completed. However, in *She was reading the newspaper article (when the phone rang)*, the event occurred in the past but there is no implication that the event (reading the newspaper article) had been completed.

There are two forms of aspect. One of these is lexical aspect, pertaining to the meaning of the verb in combination with the contents of the larger predicate. An important lexical aspect distinction is the distinction between telic events and atelic events. Telic events specify an endpoint or completion; atelic events do not specify an endpoint. Thus, *We cleaned the garage in an hour* is telic, whereas *We cleaned the garage for an hour* is atelic. Note that the same lexical verb (*cleaned* in this case) and tense (past tense) can be used for atelic as well as telic events because it is the larger predicate, which is the verb in combination with the particular lexical items or phrases that follow (as in *in an hour* versus *for an hour*), that determines telicity.

Grammatical aspect is the other form of aspect. Here, the distinction is made between the perfective and the imperfective (Bybee, 1985). The perfective involves viewing the described event "from the outside, as a completed whole" (Wagner, 2001, p. 663). In contrast, the imperfective views the event from within; there is no assumption of completion. In English, the imperfective is reflected in the use of the progressive. In the example, *He was writing the love letter*, there is no basis for knowing if the event was completed. English is much less transparent with regard to the perfective. Constructions such as the present perfect (as in *We have eaten*) do not express perfective aspect but rather express present relevance of past events (Comrie, 1976). However, past tense in English can have a default perfective interpretation. For example, the sentence *He wrote the love letter* has a perfective interpretation. Note, however, that a slight change in the predicate can alter the interpretation; *He wrote love letters* can refer to an activity pursued in the past with no assumption of completion.

TENSE AND ASPECT INTERACTIONS IN CHILDREN WITH LI

In a language such as English, young TD children's initial use of verb inflections seems to reflect influences of both lexical and grammatical aspect (Bloom, Lifter, & Hafitz, 1980; Shirai & Andersen, 1995). Children's early use of *-ed* (and irregular past forms) appears to be closely associated with completed actions as expressed in telic predicates (e.g., *He pushed me*). The initial use of *-ing* tends to mark imperfective aspect focused principally on atelic predicates in the present (e.g., *Look, horsie running*). Although this interaction between tense and aspect may be viewed as challenging for the young language learner, it seems likely that the strong correlations in the language between continuous actions and present tense on the one hand and completed actions and past tense on the other hand provide children with a means of hooking into the tense system of the language. That is, these relationships between particular forms of aspect and tense may be facilitative.

The possibility that aspect may facilitate the acquisition of tense seems more plausible given the differences seen between TD children and children with LI with regard to their sensitivity to aspectual information in the comprehension and production of tense. For example, Leonard et al. (2007) found that young TD children acquiring English were more likely to use past tense in utterances with telic predicates than in utterances with atelic predicates, whereas somewhat older children with LI, whose overall use of past tense was lower than that of the younger TD children, were no more likely to use past tense with telic predicates than with atelic predicates. Even the use of the progressive *-ing* inflection, which is often assumed to be a relative strength among children with LI (Leonard, 1998), was not used as frequently by children with LI as by same-age TD peers in past progressive contexts. Whereas both younger and older TD children were more likely to use *-ing* when referring to past events involving atelic predicates than when referring to past events involving telic predicates, the children with LI showed no such distinction.

Leonard and Deevy (2010) examined tense–aspect interactions in the sentence comprehension of English-speaking children with LI and younger and older TD children. They found that both groups of TD children were more accurate in understanding telic predicates in the past progressive if the actions had been completed than if they were left incomplete. The children with LI, whose overall accuracy was lower, showed no difference as a function of whether the action had been completed.

Collectively, these findings led Leonard and Deevy (2010) to propose that children with LI may be relatively insensitive to aspectual distinctions that can interact with tense, and therefore these children cannot take advantage of potential starting points for learning past tense. For example, young TD children seem to be sensitive to the completion of actions and, because completed actions are often referred to with past tense, completion can direct the children's attention to past tense forms. If children with LI are relatively insensitive to the linguistic relevance of action completion, their attention to past tense forms may be delayed.

Although these previous studies of children with LI were designed to assess whether children with LI are missing important cues to past tense by their insensitivity to aspectual information, it can also be argued that aspect itself may be relatively weak in this clinical population. As noted earlier, English is not an ideal language in which to test this possibility given that only the imperfective is marked, in the form of the progressive. Even the English progressive is not very transparent given that it is often used to refer to actions in the present even when the continuous nature of the action is not of particular relevance to the speaker or listener. For example, a picture of a girl walking is likely to be described with the progressive, as in *A/the girl is walking*, not to emphasize the continuous nature of the action but simply to name the action. Speakers of other languages might well use a simple present tense form (the equivalent of *A/the girl walks*) to describe the same picture. Indeed, the Leonard et al. (2007) finding that children with LI were less likely than TD peers to use *-ing* in the past progressive may suggest that the children had not clearly disassociated progressive aspect from present tense.

The literature on LI in other languages also reveals signs that these children may have weaknesses in the area of aspect. Fletcher, Leonard, Stokes, and Wong (2005) found that Cantonese-speaking children with LI were less likely to use aspect markers than younger and older TD peers in contexts where these are highly likely. In Cantonese, aspect markers are monosyllabic morphemes that immediately follow lexical verbs to specify that the action has been (or will be) completed, is (or was) ongoing, and so on. Cantonese does not mark tense. Thus, if tense were the sole temporally related feature that was problematic for children with LI, Fletcher et al. (2005) might have found no differences between the LI group and the TD groups. However, a difference was observed that suggested that aspect may be an area of weakness in LI.

Evidence from German also seems consistent with the notion that aspect is a relatively weak area of language in children with LI. Penner, Schulz, and Wymann (2003) and Schulz and Wittek (2003) found that German-speaking children with LI were less likely than TD peers to require evidence of an endpoint when judging telic verbs as appropriate in contexts such as "Has the girl closed the door?" There is even a hint from the literature that there may be a nonlinguistic source for some of this insensitivity. Kelly and Rice (1994) found that a group of children with LI differed from TD peers in their preferred interpretations of novel verbs. Whereas the TD children were more likely to prefer a change of state interpretation than a motion interpretation, no such preference was shown by the children with LI.

THE ROLE OF ADVERBS

In many languages, adverbs express notions of time that can complement the use of tense or aspect. Adverbs or adverbial phrases such as *Yesterday or A minute ago* frequently accompany past tense forms, whereas adverbs such as *still* and *already* often accompany aspect forms. Children with LI seem to acquire certain adverb forms earlier than tense forms (Moore & Johnston, 1993). Children's familiarity with temporal adverbs might serve to facilitate their development of compatible tense or aspect forms. However, given the overlap between adverbs and tense/aspect, there is also the risk that children may go through a period during which the use of an adverb is viewed as a sufficient expression of past/present time or completed/incomplete action, and hence, when an adverb is used, the tense/aspect form is omitted (Krantz & Leonard, 2007).

TENSE AND ASPECT IN HUNGARIAN

In this paper, we examine the use of tense and aspect in children with LI who are acquiring Hungarian. This language offers an especially useful view of tense and aspect because of the manner in which these two temporally related notions are marked. Past and present tense are distinguished by the presence versus absence of a verb inflection, and perfective and imperfective aspect are distinguished by the presence versus absence of a prefix (Kiefer, 1994). Tense marking and aspect marking can each occur without the other. When past tense occurs without perfective marking (hence, 0 + verb stem + past), the interpretation is that of an action in the past in which completion cannot be assumed, which is an imperfective interpretation. When the perfective prefix occurs without past tense (hence, perfective + verb stem + 0), completion is expressed without an assumption of the event occurring in the past. These details permit an inspection of how (or whether) children make the imperfective–perfective distinction when this distinction does not require tense. The remaining combinations involve the simultaneous expression of perfective aspect and past tense (perfective + verb stem + past) and the simultaneous expression of imperfective aspect and present tense, a combination that requires no overt morphology (0 + verb stem + 0). Examples appear in Table 1. Hungarian also employs agreement inflections. These are separate morphemes that follow the past tense morpheme in sequence, in agglutinating fashion. In the present study, we controlled for agreement by employing only third-person singular contexts. Verb forms in these contexts have the simplest structure.

A common test for the imperfective and perfective that is applicable to Hungarian is whether the event is divisible (Dowty, 1979). In the case of the imperfective, for a sentence such as *Jeanette 9:00-től 9:15-ig pakolt a bőröndbe* "Jeanette was packing the suitcase from 9:00 to 9:15," it is true that Jeanette was packing at many time points between 9:00 and 9:15. In the case of the perfective, the event is not divisible in this way. For example, in *Jeanette 15 perc alatt bepakolt a bőröndbe* "Jeanette had packed the suitcase in 15 minutes," it is not true that at different

Table 1. *Imperfective and perfective aspect forms in present and past tense for the third person singular form of pakol "pack" (e.g., "pack a suitcase")*

	Imperfective	Perfective
Present	*pakol* (0 + verb stem + 0)	*bepakol* (perf + verb stem + 0)
Past	*pakolt* (0 + verb stem + past)	*bepakolt* (perf + verb stem + past)

points in the interval Jeanette had packed the suitcase. (In translation, we use the English past perfect as the closest approximation to the Hungarian perfective in past tense.) In most cases, the imperfective is the default interpretation when a verb overtly marked only for tense and agreement is used. In past tense, such verbs therefore describe an action in the past but do not imply that the action had been completed. Most often, the imperfective is unrestricted, in the sense that the action described is not dependent on the temporal structure of another event. An example in English would be *I was reading most of the day*; in Hungarian, the verb conveying this act would be a simple verb in past tense.

Hungarian makes use of several different prefixes to mark the perfective. For example, whereas *meg-* is employed to express the perfective in "had eaten" (*meg-ette*), the prefix *ki-* is employed to express the perfective in "had painted" (*ki-festette*). Because the use of a prefix to express perfective aspect is a productive process, whereas the particular prefix that must be selected is often verb-specific, young children can make errors in the particular prefix they use with a verb, although such errors can be recognized as an attempt to express the completion of the action.

In the present study, the items selected to test the imperfective–perfective distinction are verbs used with definite direct objects. Verbs in these contexts without a prefix are uniformly interpreted as imperfective; when the prefix is added, a perfective interpretation is required (Kiefer, 1994, p. 450). We also examine the role played by adverbs in facilitating (or inhibiting) the children's use of aspect. As noted earlier, the complementary nature of particular adverbs and particular types of aspect may assist children in selecting the appropriate aspect form to use. However, it is also plausible that children could treat the adverb as a sufficient means of communicating the completed or incomplete nature of the event and thus fail to express aspect when adverbs are involved.

Despite the independent marking of tense and aspect in Hungarian, young TD children acquiring this language are prone to refer to pictures of ongoing actions (e.g., someone drinking juice) using the imperfective with present tense but to refer to pictures showing a completed action (e.g., a picture of someone holding a glass with only a few drops in it) using the perfective with past tense (Pléh, 1992). In this study, we assess the children's ability with the imperfective–perfective distinction while keeping tense constant.

Finally, we include in our task items that serve at once as filler items and as a further control. Prefixes are required to express perfective aspect; therefore, it is possible that any difficulties with perfective items are not due to perfective aspect per se but are due to the necessity that the prefix must be attached to the

Table 2. *Test scores (z scores) for the children with language impairment*

z Score	PPVT	TROG	Nonword Repetition	Sentence Repetition
Mean	−0.63	−0.59	−1.06	−5.39
Range	−2.74 to 1.49	−3.52 to 0.96	−2.61 to 1.27	−8.37 to −1.68
SD	0.94	1.38	0.97	2.11

Note: PPVT, Peabody Picture Vocabulary Test; TROG, Test for the Reception of Grammar.

verb. Fortunately, Hungarian also makes use of prefixes for functions other than marking perfective aspect. Prefixes can also serve functions such as marking the direction, mode, or intensity of an action when used with particular types of verbs. By including items of this type in our experimental task, we provide the children with occasional items that do not focus on aspect and do so in a manner that allows us to assess the children's more general ability with prefixes.

METHOD

Participants

The LI group consisted of 21 monolingual Hungarian-speaking children between the ages of 4 years, 10 months (4;10) and 7;2 who had been diagnosed with LI. All children were enrolled in language intervention programs prior to their recruitment for this study. Inclusive criteria for the LI group were significant deviation from age norms (−1.5 *SD*) on at least two of the following four language tests: a test of grammatical comprehension (Test for the Reception of Grammar [TROG]; Bishop, 1983), a test of receptive vocabulary (Peabody Picture Vocabulary Test [PPVT]; Csányi, 1974), a sentence repetition test (Hungarian Sentence Repetition Test Magyar Mondatutánmondási Teszt; Kas & Lukács, 2011) and a nonword-repetition test (Racsmány, Lukács, Németh, & Pléh, 2005). A summary of the children's performance on these tests, computed in terms of *z* scores, appears in Table 2.

Tests that were adaptations of existing tests (PPVT and TROG) were standardized following modifications of lexical and grammatical details to ensure that the resulting items were relevant to, and developmentally appropriate for Hungarian. The Hungarian version of the TROG consists of 20 blocks, each with four sentences of the same construction. Examples of construction types include comparatives, postmodified subjects, and embedded clauses. Changes from the English version (beyond translation) were necessitated because the Hungarian pronouns do not make a distinction according to gender, suffixes are used to express notions that are expressed by prepositions in English, and, where passive sentences are used in English, Hungarian makes use of active sentences with object–verb–subject word order (with case marking on the object), among others.

It can be seen from Table 2 that the children with LI earned extremely low *z* scores on the sentence repetition test. Indeed, all 21 children in the LI group

Table 3. *Ages and raw scores on the PPVT for the three groups of children*

	LI	VC	AC
Age			
Mean	5;9	4;10	5;8
Range	4;10–7;2	3;3–6;6	4;8–7;3
SD	8 months	9 months	12 months
PPVT			
Mean	69.95	69.25	83.81
Range	27–110	27–106	59–114
SD	19.76	19.48	15.70

Note: Age mean and range are in years;months. PPVT, Peabody Picture Vocabulary Test; LI, language impairment; VC, vocabulary control; AC, age control.

scored lower than −1.5 *SD* on this test. The sentence repetition test is newly developed, and it is possible that the existing normative data for TD children are not representative, resulting in uncharacteristically low *z* scores for the LI group. However, although it is possible that the sentence repetition abilities are not quite as low as these extreme *z* scores suggest, recall that to be included in the LI group, a child had to score −1.5 *SD* or lower on at least two of the language tests. Therefore, inclusion of a child in the LI group was not based on the sentence repetition test score alone. Finally, children with a nonverbal IQ below 85 (Raven Coloured Progressive Matrices; Raven, Court, & Raven, 1987), or a history of hearing impairment or any neurological conditions were excluded from the study.

Performance of the LI group was compared to that of two TD control groups. All TD children were enrolled in regular preschools or elementary schools and were regarded as developing in an age-appropriate manner according to teacher and parent report. The 21 children in one of these comparison groups were matched individually with children in the LI group according to receptive vocabulary (PPVT) raw scores to within four points. At the time of the study, the TROG was in the process of being standardized; we therefore used the PPVT as the basis for matching, as normative data for this test were well established. Because the production task involved completing the examiner's sentences with short (three-word) responses, this task did not place an utterance length burden on any of the children. The PPVT-matched children ranged in age from 3;3 to 6;6, which we refer to here as the vocabulary control (VC) group. Because the children with LI scored below age level on the PPVT, the children in the VC group were considerably younger than the children with LI. The 21 children in the other comparison group were matched to the children in the LI group to within 2 months of age. Their ages ranged from 4;8 to 7;3; we refer to this group here as the age control (AC) group. All children in the comparison groups scored within the normal range on the PPVT. A summary of the three groups' age and PPVT scores is provided in Table 3.

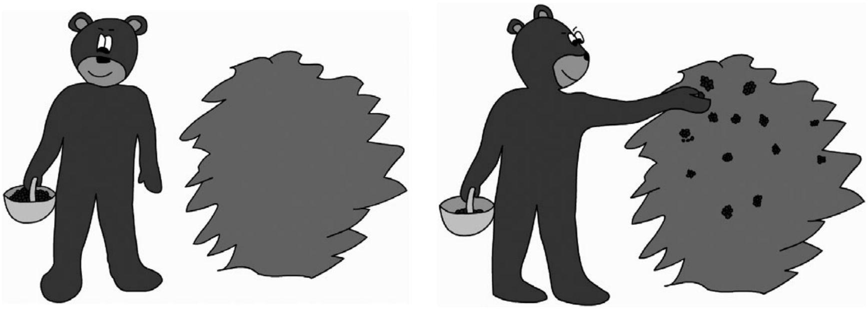


Figure 1. The picture pair for the target sentence *Amikor lefényképezték, a medve leszedte a málnát.* ("When this picture was taken, the bear had picked the raspberries.").

Procedure

Comprehension. Two tasks were devised to assess the children's ability with the imperfective–perfective distinction. The first was a comprehension task. Two drawings were presented on a page, one depicting a completed action, the other depicting the same action in progress. An example is shown in Figure 1. All drawings depicted an animal performing (or having just performed) a humanlike action. All test sentences included a definite article plus subject noun, a verb in past tense in either imperfective or perfective form, and a direct object consisting of a definite article plus object noun. In constructions of this type, the presence or absence of a prefix makes the clearest distinction between the imperfective and perfective (Kiefer, 1994, p. 450). To ensure that the sentences without prefixes were compatible with an imperfective interpretation, the examiner described the following scenario to the child (given here in English): "Yesterday a photographer went to the zoo, and took pictures of what the animals were doing. Help me sort the pictures. Point to the picture that shows what I say." The examiner then preceded each item with "When this picture was taken . . ." This clause, *Amikor lefényképezték,* is literally translated as "When [they] had photographed [him/her] . . ." and is perfective (with the prefix *le-*) and in past tense.

Fifty items were used; 40 of these were divided into the conditions of interest and the remaining 10 items were filler items. Four conditions were employed, with the same 10 verbs used in each condition. Twenty of the items were sentences in the imperfective and the remaining 20 were sentences in the perfective. The imperfective and perfective items were subdivided, resulting in 10 of each type that had an accompanying adverb and 10 of each type that had no accompanying adverb. Thus, the same verb appeared in an imperfective with adverb condition, an imperfective with no adverb condition, a perfective with adverb condition, and a perfective with no adverb condition. The particular adverb used (*még mindig* "still" for the imperfective, már teljesen "already completely" or *már az összes* "already . . . the whole" for the perfective) was consistent only with, and could be viewed as providing semantic support for the particular type of aspect with which it was paired. Examples of four items are shown in (1); the entire list of items appears in Appendix A. Recall that English does not possess a distinct form to mark perfective aspect; in these examples, the English past perfect ("had eaten") is used as the closest equivalent to convey that the action had been completed following a clause such as "When this picture was taken . . . "

(1) a. *A malac még mindig ette a jégkrémet*
"The pig was still eating the ice cream"
b. *A majom ette a tortát*
"The monkey was eating the cake"
c. *A nyúl már megette a répát*
"The rabbit had already eaten the carrot"
d. *A tigris megette a levest*
"The tiger had eaten the soup"

The 10 verbs used in each of the four conditions were *ette/megette* ("was eating/had eaten"), *itta/megitta* ("was drinking/had drunk"), *vágta/kivágta* ("was cutting out/had cut out"), *szedte/leszedte* ("was picking/had picked"), *festette/kifestette* ("was painting/had painted"), *rajzolta/lerajzolta* ("was drawing/had drawn"), *építette/felépítette* ("was building/had built"), *mosta/kimosta* ("was washing/had washed"), *vágta/levágta* (a haját) ("was cutting/had cut (her hair)"), and *szerelte/megszerelte* ("was fixing/had fixed"). The second verb form listed in each pair is the perfective, containing the prefix. As noted earlier and as can be seen from these 10 verbs, there is more than one perfective prefix in Hungarian (e.g., *meg-*, *le-*, *ki-*); the particular prefix employed depends in large part on the specific verb used. As will be seen, this fact was taken into consideration in the scoring.

For each of the 40 test items, the target and foil drawings depicted the same subject, ongoing/completed action, and object. The left versus right location of the target drawing on the page was counterbalanced across the four conditions.

The 10 filler items involved pairs of pictures that could be distinguished on the basis of a difference in prefixes that served a different (nonperfective) function (e.g., a mouse having just run up the stairs versus a mouse having just run down the stairs, a pig having just turned on the tv versus a pig having just turned off the tv). If children performed poorly on the perfective items, these filler items could allow us to determine if the problem rested with prefixes in general or only with those that marked perfective aspect.

Production. The comprehension task provided an indication of whether the children had a basic understanding of the imperfective–perfective distinction when tense was held constant. The production task (always administered after the comprehension task) placed greater demands on the children because the children were given the same type of background about a photographer at the zoo, but this time were asked to complete the sentence produced by the examiner by supplying the verb and direct object. This required the child to choose not only the appropriate aspect but also the appropriate tense. For example, after being told that the photographer took pictures of what the animals were doing, the child was given the prompt *Amikor lefényképezték* (e.g., *a ló . . .*)

"When this picture was taken (e.g., the horse . . .)." Again, 10 items involving the same verbs were used for each of the four conditions. For items in the imperfective with adverb condition and the perfective with adverb condition, the examiner provided the adverb immediately after the name of the subject, as in *Amikor lefényképezték* (e.g., *a majom már* . . .) "When this picture was taken (e.g., the monkey already . . .)." In Hungarian, adverbs of this type often precede the verb, and thus, literally translated, have the order seen in "The lion still was fixing the car" and "The rabbit already had eaten the whole carrot." Therefore, the child's response in all conditions required only the verb (with appropriate tense and aspect) and direct object. The target pictures used in the comprehension task were also used for the production task. The comprehension and production tasks were administered on separate days.

Scoring

The scoring of the comprehension task was simply the number correct out of 10 items in each of the four conditions. (The filler items were scored in the same way, to verify the children's understanding of simple nonperfective prefixes.) The probability of a correct response by chance was 50% as there were two alternative drawings (target and foil) for each item. Results will be discussed in terms of percentages so that the children's level of performance relative to chance is always clear.

Scoring of the production task was performed in two ways. The first scoring method required both past tense and the use of the correct type of aspect (imperfective or perfective). If children used a different but plausible verb and correctly marked it for tense and aspect, it was scored as correct. In addition, for perfective items, if children produced a different prefix than the one ordinarily associated with the verb in the particular context being tested, but the prefix served to mark perfective aspect (and past tense was marked), it was scored as correct. It should be noted that in all of the instances in which an unexpected prefix was used, its use was grammatical even though it was not the prefix that would be selected by adult speakers in that particular context. If the child added an adverb to a response that otherwise met our tense and aspect criteria as just described, the response was treated as correct. Errors were productions of an imperfective form in a perfective context or vice versa, whether with the target verb or with another plausible verb. Productions of the appropriate imperfective or perfective form but with present tense were likewise treated as errors. Finally, if a verb form other than an imperfective or perfective form (regardless of tense) was used, it was scored as an error.

The first scoring method was regarded as demanding because the children had to retain the past tense information contained in the examiner's prompt while selecting the appropriate form for imperfective or perfective aspect. We also adopted a second scoring method that was identical to the first except that we allowed either a present tense or a past tense form of the verb. This second scoring method allowed us to evaluate the children's command of the imperfective–perfective distinction even if the past tense information in the examiner's prompt was not retained in the child's response.

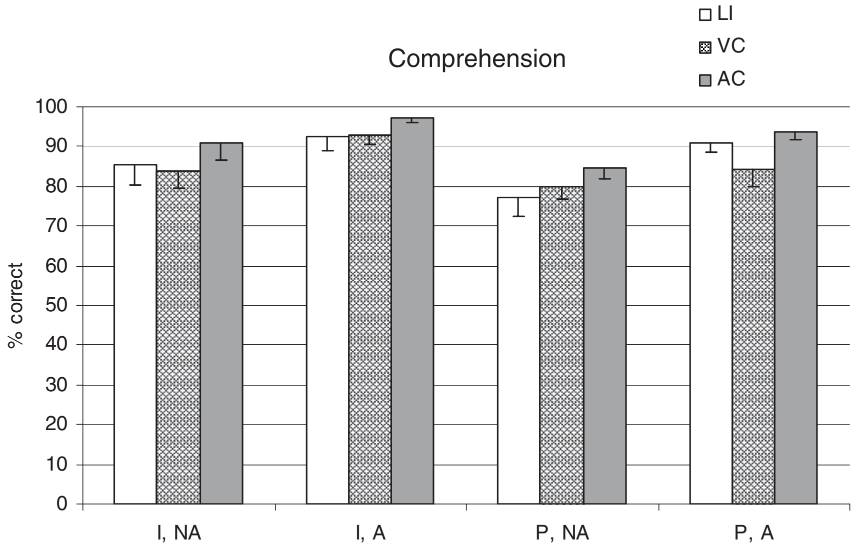


Figure 2. The mean percentage correct on the comprehension task. Error bars are standard errors. I, imperfective; P, perfective; NA, no adverb; A, adverb.

RESULTS

Comprehension

The comprehension task provided an estimate of the children's understanding of the imperfective–perfective distinction when (past) tense was held constant. A mixed model analysis of variance (ANOVA) was used, with participant group (LI, VC, AC) as a between-subjects factor and aspect (imperfective, perfective) and adverb (adverb, no adverb) as within-subject factors. The assumption of homogeneity of variance for the ANOVA was met. Participant group was not significant, F (2, 60) = 2.33, p = .11, and an inspection of Figure 2 reveals that the largest numerical differences were between the VC and AC children, with the children in the LI group falling at intermediate levels. A main effect was found for aspect, F (1, 60) = 5.80, p = .02, partial $\eta^2 = 0.09$, with higher accuracy on imperfective items than on perfective items. Adverb was highly significant, F (1, 60) = 35.80, $p < .001$, partial $\eta^2 = 0.37$. Items containing an adverb in the prompt were more likely to be correct than those with no adverb. No interactions were significant (all Fs < 1.00, all ps > .40).

As can be seen from Figure 2, the children's accuracy was relatively high. For example, the children with LI scored above chance, even on items in the most difficult condition, the perfective with no adverb condition, t (20) = 5.98, $p < .001$. However, given that scores were not at ceiling level, it appeared useful to determine the degree to which the errors that were seen might be due in part to a weakness in comprehending prefixes in general. Recall that the filler items involved contrasts between nonaspectual prefixes (e.g., running up vs. running

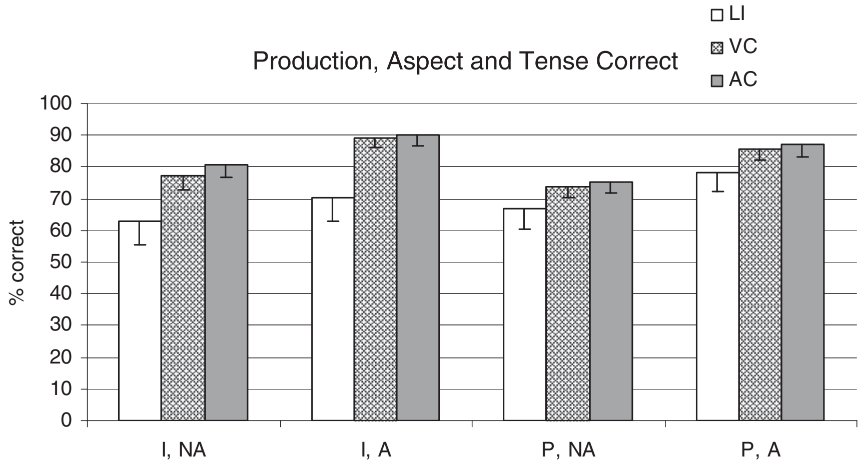


Figure 3. The mean percentage correct on the production task with the scoring method that required both tense and aspect to be correct. Error bars are standard errors. I, imperfective; P, perfective; NA, no adverb; A, adverb.

down the stairs). The children's accuracy on these items was quite high (LI, $M =$ 90%, $SD = 10.95$; VC, $M = 97.1\%$, $SD = 0.46$; AC, $M = 93.8\%$, $SD =$ 0.74). Furthermore, a comparison between the children's performance on the filler items and their performance on the items of the perfective with no adverb condition (which required an understanding of the perfective prefix for a correct response) revealed a significant difference favoring the filler items, F (1, 60) = 36.17, $p < .001$, partial $\eta^2 = 0.38$. Participant group only approached significance, F (2, 60) = 2.82, $p = .07$, partial $\eta^2 = 0.09$. The interaction was not significant, F (2, 60) = 0.04, $p = .96$.

Production

Tense and aspect correct. The first scoring method for the production task required that the children responded with a past tense form as well as the correct aspect (imperfective or perfective). A mixed-model ANOVA was employed with participant group (LI, VC, AC) as a between-subjects variable and aspect (imperfective, perfective) and adverb (adverb, no adverb) as within-subjects variables. Again, the assumption of homogeneity of variance was met. A main effect was found for participant group, F (2, 60) = 4.21, $p = .019$, partial $\eta^2 = 0.121$. Least significant difference testing revealed that the children with LI were less accurate than both the AC group ($p = .010$) and the VC group ($p = .024$). The AC and VC group did not differ ($p = .729$). Aspect was not significant, F (1, 60) = 0.03, $p = .86$. However, adverb proved significant as a main effect, F (1, 60) = 35.91, $p < .001$, partial $\eta^2 = 0.37$, with higher scores for the adverb conditions than for the no-adverb conditions. No interactions were significant (all Fs < 1.00, all ps > .34). A summary of the results can be seen in Figure 3.

Table 4. *Distribution of responses (mean percentages) on the production task*

	Correct	Error			
		Aspect	Tense	Tense + Aspect	Other
LI	69.77	18.33	4.40	2.14	5.36
VC	81.68	11.67	2.26	0.10	4.29
AC	83.23	9.40	1.90	0.71	4.76

Note: LI, language impairment; VC, vocabulary control; AC, age control.

Table 5. *Distribution of aspect errors and use of perfective prefixes (mean percentages) on the production task*

Type of Production	LI	VC	AC
Aspect error			
Imperfective for perfective	49	51	50
Perfective for imperfective	51	49	50
Perfective prefix			
Correct prefix form	66	72.2	80.1
Unexpected prefix form (correct in aspect)	6.4	7.6	1.1

Note: LI, language impairment; VC, vocabulary control; AC, age control.

Table 4 provides the distribution of the children's production responses. As can be seen, errors of aspect only, in which the children produced a past tense imperfective form in a past tense perfective context or a past tense perfective form in a past tense imperfective context, constituted the most frequent error type for all three groups, although such errors were considerably more frequent in the responses of the children with LI. It is important that these errors occurred in both directions. For the children with LI, of the aspect-only errors, 49% were productions of the imperfective on items requiring the perfective, and 51% were productions of the perfective on items obligating the imperfective. Examples appear in (2) and (3), respectively. For the VC group, these errors were 51% and 49%, respectively, and for the AC group, these errors were evenly divided (50% and 50%, respectively). An especially noteworthy detail is that the productions of perfective forms in imperfective contexts involved commission errors: using a prefix where none was required. For the verbs included in our items, the prefixed forms have lower frequencies of occurrence in Hungarian than the nonprefixed forms. A summary of the aspect substitution errors appears in Table 5.

(2) Examiner:	*Amikor lefényképezték, a tehén . . . . (megitta a tejet)*	
	When this picture was taken, the cow. . . (had drunk the milk)	
Child:	*itta a tejet*	
	was drinking the milk	
(3) Examiner:	*Amikor lefényképezték, a szamár. . . (itta a sört)*	
	When this picture was taken, the donkey. . . (was drinking the beer)	
Child:	*megitta a sört*	
	had drunk the beer	

Recall that children's production of a perfective prefix form that differed from the one expected in the context was scored as correct provided that it occurred in a perfective context. As can be seen from Table 5, such productions did not constitute a large percentage of the perfective prefixes produced. It is important that all of the observed productions of this type were grammatical; they were simply not the form that an adult speaker would select in that particular context. For example, some children produced *levágta a fát* "had cut the tree" rather than *kivágta a fát*. When referring to cutting fingernails or beards, only *levágta* can be used, but in the context of cutting trees, adults would select *kivágta*, even though *levágta* is not ungrammatical.

Errors of tense only were considerably less frequent than errors of aspect only in all three groups, although the children with LI produced more of these errors than did their TD compatriots. These errors were instances in which the children produced the present tense form in place of the past tense form, but employed the proper form of aspect. It can also be seen from Table 4 that the children were more likely to err in aspect or tense than to produce an error involving both. Finally, "other" errors were seen in all three groups. These included instances in which children produced the Hungarian equivalent of "was done with" or "finished" in contexts requiring perfective forms. Along with being semantically imprecise, these forms do not involve perfective prefixes. Another error placed in the "other error" category occurred when children produced the Hungarian equivalent of "almost" plus a verb in perfective form in a context suitable for an imperfective form.

Aspect correct, tense requirement relaxed. The second scoring method allowed responses that were in present tense as well as past tense, provided that aspect (imperfective, perfective) was correct. An ANOVA identical to the one used with the first scoring method was applied to the scores based on the second scoring method. The assumption of homogeneity of variance was met. Participant group was again significant, $F(2, 60) = 3.46$, $p = .04$, partial $\eta^2 = 0.10$. Least significant difference testing revealed that the children with LI were less accurate than both the AC group ($p = .013$) and the VC group ($p = .027$). The AC and VC group did not differ ($p = .768$). Aspect was approaching significance, $F(1, 60) = 3.09$, $p = 0.08$, partial $\eta^2 = 0.05$. However, adverb was highly significant, $F(1, 60) = 36.68$, $p < .001$, partial $\eta^2 = 0.38$. Again, no interactions were significant (all Fs < 1.00, all ps $> .35$). The findings are illustrated in Figure 4.

From a comparison of Figures 3 and 4, it can be seen that the children's performance levels did not increase appreciably with the second scoring method,

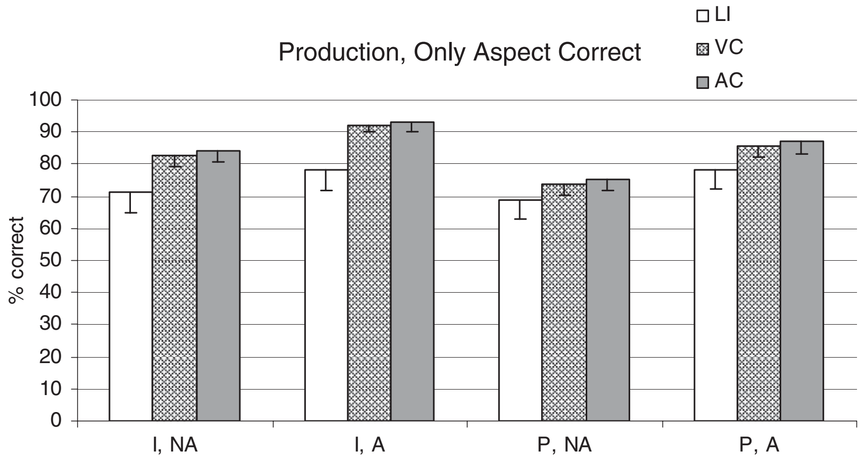


Figure 4. The mean percentage correct on the production task with the scoring method that required only aspect to be correct. Error bars are standard errors. I, imperfective; P, perfective; NA, no adverb; A, adverb.

and group differences were again apparent. These findings suggest that a problem with tense was not a major factor, either globally or specifically within any of the four conditions.

DISCUSSION

This study was based on the premise that aspect may be vulnerable in children with LI. Previous studies in other languages suggested not only that children with LI might be insensitive to aspectual information as it interacts with tense, but that aspect might be an area of weakness independent of tense. Hungarian provided an excellent opportunity to examine the status of aspect in children with LI, as this language permits a view of the imperfective–perfective distinction that can be made without the involvement of tense.

Results from our comprehension task suggested that the children with LI had an understanding of the imperfective–perfective distinction that was comparable to that of their TD peers. They performed above the level of chance on even the most difficult of the item types, those involving the perfective without an accompanying adverb. Although these children, like their TD peers, occasionally made errors, their performance suggested considerable knowledge of this aspectual distinction, at least as it was tested here.

However, the production performance of the children with LI revealed weaknesses even relative to the younger TD children. Our initial scoring required appropriate use of (past) tense as well as correct use of the imperfective or perfective. Analysis based on that scoring showed that the children with LI were less accurate than both the VC and the AC children. Such a finding could be attributable to a problem with tense. However, when we scored the children's productions solely in terms of their use of the imperfective or perfective, the same group differences were found. Scores improved very little, suggesting that choice of the wrong tense was not a central factor in the findings. We interpret these findings to mean that the children with LI had some weakness in the expression of aspect, more specifically, in the expression of imperfective and perfective aspect. Before discussing the implications of this interpretation, we consider alternative interpretations, and provide reasons for why we find them unconvincing.

First, it might be argued that the perfective prefix, which is a monosyllabic syllable in word-initial position, is not perceptually salient and, because children with LI are often reported to have perceptual limitations, these limitations may have made the children's acquisition of the perfective prefix more difficult. This does not seem likely given that, in Hungarian, the first syllable of a word is stressed. Thus, whereas a verb without a prefix will have stress on the first syllable of the stem, the same verb with a prefix will have stress on the prefix itself. Another reason for not preferring a perceptual explanation is that the children with LI were similar to the TD children in their comprehension of perfective (and imperfective) items; the children with LI scored lower than the TD children only in production. A related argument might be that the past tense inflection requires production of /t/, a brief consonant that might have been difficult in production as well as in perception. However, in the context in which past tense occurred in this study—with a definite direct object—the past tense allomorphs were the syllabic forms /ta/ or /te/. Another reason to believe that this factor could not have played a major role is that the children with LI showed lower production accuracy than the VC and AC children even when we excluded tense accuracy from consideration in the scoring.

Second, it might be argued that the static nature of the visual stimuli made the distinction between imperfective and perfective events more difficult to discern. Certainly the ongoing action depicted in the target pictures for imperfective items might have made the expression of past tense more difficult. However, in this case, the most likely error would be one of using the imperfective in present tense. This was not the typical error. As just noted, when tense was ignored in the scoring, the children with LI nevertheless scored lower than the VC and AC children on all item types, including those assessing the imperfective.

It still might be argued that perfective aspect was inadequately represented by our pictures. That is, the result of the action was depicted (e.g., first drawing in Figure 1) but without the action visible in the picture, the children had to infer that the action had actually been performed. Perhaps the ability to infer was the obstacle for the children, not the use of perfective aspect. However, two observations cast some doubt on such an interpretation. First, the children were administered the comprehension task prior to the production task. The comprehension task employed the same pictures and target sentences. Thus, the children had already heard the examiner provide the correct description of the pictures intended to reflect perfective aspect, albeit in a prior testing session. Second, and of importance, most of the children's errors on perfective items were productions of the imperfective.

Such productions contain the name of the action (e.g., *ette* "was eating"). Thus, even when children chose the wrong form of the verb, they evidently could infer the action performed by observing the result.

Current accounts of verb morphology deficits in children with LI were not designed to handle the types of difficulties observed in this study. For example, the agreement deficit account of Clahsen and colleagues (Clahsen, Bartke, & Göllner, 1997; Clahsen & Dalalakis, 1999) does not seem applicable, as all of our contrasts in comprehension and production involved the same third-person singular subject. Person and number information did not have to be considered, yet the children with LI nevertheless performed below the level of the VC and AC children on the production task. Similarly, because Hungarian is a null-subject language, it does not fall within the purview of the "extended optional infinitive" account of Rice, Wexler, and their colleagues (Rice, 2003; Rice et al., 1998; Rice & Wexler, 1996).

In null-subject languages, many verb inflection errors by children with LI can be regarded as "near-miss" errors, which are errors that differ from the correct form on only one of several possible details (e.g., Bedore & Leonard, 2001). In a recent study of tense and agreement by Lukács, Leonard, Kas, and Pléh (2009), Hungarian-speaking children with LI showed the same near-miss pattern, such as producing a present tense first person singular form in place of a present tense first person plural form or producing a past tense third person plural form in place of a past tense first person plural form. No single inflection type was consistently in error, and no evidence of a default form was found. Instead, the children seemed to have difficulty consistently selecting the inflected form that possessed the correct combination of features. Lukács et al. (2009) characterized the difficulty as one of processing, given that the children exhibited sufficient knowledge to generate the correct form in most instances but sometimes erred in ways that suggested retrieval of a form that approximated but did not match the target form.

There were details in the data of the present study that also suggest that processing factors were at work. First, the children with LI scored lower than the VC and AC children in production but did not differ significantly from their TD peers in comprehension. Evidently the act of retrieving the appropriate form for production was relatively more difficult for the children with LI than for the TD children even though their recognition of the correct form in the comprehension task was comparable to that of their peers.

Second, the children with LI did not err primarily on one type of aspectual form. Instead, occasional errors were made on aspect items of each type. It was also clear that the children did not rely on a single type of form as a default form. A likely choice for a default form would have been a verb that lacked both a prefix and a past tense inflection, given that forms of this type have no morphological embellishments and are actually grammatical in a present tense imperfective context. However, instead, the children with LI usually produced perfective items correctly but sometimes used the imperfective form instead, and usually produced the imperfective items correctly but occasionally added a prefix that resulted in an (inappropriate) perfective form. That the children varied their response, even when in error, also serves as evidence that the children were clearly not simply mimicking the form given in the examiner's prompt, which never varied.

Third, errors of aspect only and errors of tense only were each more frequent than errors of both aspect and tense. These differences are consistent with a near-miss pattern of errors.

Fourth, the children with LI were less accurate than the VC and AC children but showed the same pattern of performance across conditions. There were no significant interactions involving group, and all groups showed increased accuracy when the examiner's prompt included an adverb that was compatible with the imperfective or perfective form expected for the item. The similar pattern of performance across conditions in the three groups suggests that the children with LI had greater difficulty across the whole task and were not disproportionately affected by any one aspect type.

However, one important detail in the findings suggests that a processing explanation is not sufficient. The near-miss errors included the use of perfective forms in contexts requiring imperfective forms. These errors were commission errors on the part of the children: adding a prefix to the verb stem where one was not needed. Usually processing limitations result in the omission or simplification of the target in some way. However, in these particular instances, the form produced by the child contained more information (a prefix) than was required.

This finding of commission errors suggests that the children's difficulties were due in part to some confusion about the expression of aspect, at least in past tense contexts. The children seemed to have had some grasp of imperfective and perfective aspect given that they performed above the level of chance in comprehension. However, selecting the appropriate form for production may have required greater command of aspect than the children possessed. The difficulty could have involved uncertainty about the particular type of aspect that applied to the context, or to the particular form (e.g., a prefix or no prefix) that was to be used to express the selected aspect.

That the presence of adverbs facilitated the children's accuracy suggests that the children's limited ability with aspect was not confined to recognizing the correct form (an ability assessed on the comprehension task). The adverbs employed were closely associated with the respective type of aspect ("still" with the imperfective, "already" with the perfective). The presence of the adverb in the examiner's prompt, then, may have served as an important cue as to the type of aspect—imperfective or perfective—the child should express during the production task. It is also plausible that the children's prior experience with the adverbs led to the development of associations between particular adverbs and particular verb forms (e.g., perfective prefixes follow "already," whereas no prefix is used following "still"). Given that the VC and AC groups also benefited from the presence of adverbs, it is clear that their skill with aspect was not yet consolidated. Clearly, though, the children with LI were even more limited in their skill with aspect; even with the benefit of adverbs in the prompt, these children's level of accuracy never reached 80% with either type of aspect, even with the more lenient scoring method.

That the children with LI differed from the VC children on imperfective as well as perfective aspect use has important implications for current views on the status of progressive -*ing* in the grammars of English-speaking children with LI. In the adult grammar of English, progressive -*ing* marks imperfective aspect. Given the early appearance of -*ing* in the utterances of English-speaking children with LI (see Leonard, 1998), it has been assumed that this inflection is a relative strength in these children, and by extension, so, too, is the use of imperfective aspect. Yet, in the present study, the children with LI had considerable difficulty in the use of imperfective aspect. This is especially noteworthy because, in Hungarian, imperfective aspect requires no prefix.

The finding of weaknesses in the use of imperfective aspect by children with LI can also be found in other languages. In particular, Fletcher et al. (2005) found that Cantonese-speaking children with LI made significantly less use of both imperfective and perfective forms than did younger TD compatriots. Together, these findings suggest to us that the use of -*ing* by English-speaking children with LI may have been overinterpreted. Instead of reflecting an age-appropriate command of imperfective aspect, these children's use of -*ing* may reflect in part their awareness that this form is commonly used when describing actions in the present. An intention to convey the continuous nature of the action may not be present.

In summary, the findings of the present study suggest that children with LI may have a problem with temporally related information that extends beyond tense. Aspect also appears to be a vulnerable area. Specifically, it appears that children with LI may have difficulties with the grammatical expression of aspect, at least in past tense contexts. A clearer picture of this vulnerability will no doubt emerge in future studies that employ tasks that differ in their processing demands and languages that vary in how clearly they separate aspect from tense.

APPENDIX A

Target sentences in comprehension and production

1. I, NA	Amikor lefényképezték, a birka fésülte a cicát.	When this picture was taken, the sheep was combing the cat.
2. I, NA	, a szamár itta a sört.	The donkey was drinking the beer.
3. I, NA	, a majom ette a tortát.	The monkey was eating the cake.
4. I, NA	, az elefánt rajzolta az almát.	The elephant was drawing the apple.
5. I, NA	, a szarvas szerelte a biciklit.	The deer was fixing the bike.
6. I, NA	, a ló építette a falat.	The horse was building the wall.
7. I, NA	, a medve vágta a fenyőfát.	The bear was cutting the pine tree.
8. I, NA	, a zebra vágta a füvet.	The zebra was cutting the grass.
9. I, NA	, a birka mosta az inget.	The sheep was washing the shirt.
10. I, NA	, a macska szedte a barackot.	The cat was picking the apricot.
1. I, A	, a medve még mindig fésülte a kutyát.	The bear was still combing the dog.
2. I, A	, a macska még mindig itta a vizet.	The cat was still drinking the water.
3. I, A	, a malac még mindig ette a jégkrémet.	The pig was still eating the ice cream.
4. I, A	, a róka még mindig rajzolta az autót.	The fox was still drawing the car.
5. I, A	, az oroszlán még mindig szerelte az autót.	The lion was still fixing the car.
6. I, A	, a nyúl még mindig építette a várat.	The rabbit was still building the castle.
7. I, A	, a farkas még mindig vágta a férfi haját.	The wolf was still cutting the man’s hair.
8. I, A	, az elefánt még mindig vágta a pálmafát.	The elephant was still cutting the pine tree.
9. I, A	, a tehén még mindig mosta a lepedőket.	The cow was still washing the sheets.
10. I, A	, az egér még mindig szedte a cseresznyét.	The mouse was still picking the cherry.
1. P, NA	, a zebra megfésülte a kislányt.	The zebra had combed the girl(’s hair).
2. P, NA	, a tehén megitta a tejet.	The cow had drunk the milk.
3. P, NA	, a tigris megette a levest.	The tiger had eaten the soup.
4. P, NA	, a ló lerajzolta a csillagot.	The horse had drawn the star.
5. P, NA	, a malac megszerelte a tévét.	The pig had fixed the tv.
6. P, NA	, a medve felépítette a házat.	The bear had built the house.
7. P, NA	, a róka levágta a körmét.	The fox had cut his fingernails.
8. P, NA	, az oroszlán kivágta a fát.	The lion had cut the tree.

APPENDIX A (*cont.*)

9. P, NA	, a nyúl kimosta a zoknikat.	The rabbit had washed the socks.
10. P, NA	, a medve leszedte a málnát.	The bear had picked the raspberry.
1. P, A	, a farkas már teljesen megfésülte a medvét.	The wolf had completely combed the bear.
2. P, A	, a majom már megitta az összes kólát.	The monkey had already drunk all the coke.
3. P, A	, a nyúl már megette az egész répát.	The rabbit had already eaten the entire carrot.
4. P, A	, szarvas már teljesen lerajzolta a házat.	The deer had completely drawn the house.
5. P, A	, a krokodil már teljesen megszerelte a mosógépet.	The crocodile had completely fixed the washing machine.
6. P, A	, az oroszlán már teljesen felépítette a tornyot.	The lion had completely built the tower.
7. P, A	, a kutya már teljesen levágta a Mikulás szakállát.	The dog had completely cut Santa's beard.
8. P, A	, a farkas már teljesen kivágta a kaktuszt.	The wolf had completely cut the cactus.
9. P, A	, a tigris már kimosta az összes nadrágot.	The tiger had washed all the trousers.
10. P, A	, a birka már leszedte az összes dinnyét.	The sheep had picked all the watermelons.
1. F	, a ló éppen kikötötte a cipőfűzőjét	The horse had just untied its shoelace.
2. F	, a maci éppen szétcsavarta a csövet.	The bear had just screwed apart the tube.
3. F	, a majom éppen becsomagolta az ajándékot	The monkey had just wrapped the presents.
4. F	, a krokodil éppen kigombolta a kabátját.	The crocodile had just unbuttoned his coat.
5. F	, a zebra éppen ledobta a labdát.	The zebra had just thrown down the ball.
6. F	, az egér éppen felszaladt a lépcsőn.	The mouse had just run up the stairs.
7. F	, a nyúl éppen lemászott a létrán.	The rabbit had just climbed down the ladder.
8. F	, a cica éppen összepakolta a játékokat.	The cat had just packed.
9. F	, a malac éppen bekapcsolta a tévét.	The pig had just turned on the tv.
10. F	, a szamár éppen levette a sapkáját.	The donkey had just taken off his cap.

Note: I, imperfective; NA, no adverb; A, adverb; P, perfective; F, filler.

ACKNOWLEDGMENTS

This research was supported by Research Grant R01 DC00458 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (to L.B.L.), and by OTKA TS 049840 from the Hungarian National Science Foundation (to C.P.). Ágnes Lukács was a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science. The authors are grateful to the children in the Dr. Nagy László Institute of Special Education in Köszeg, the ELTE Special Preschool and Early Intervention Centre, and the Zölderdő Preschool for Speech Therapy and Nature Preservation for their participation. We also thank the speech therapists in both institutions for their help with screening and organization and to Patricia Deevy for her helpful comments.

REFERENCES

Bedore, L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*, 905–924.

Bishop, D. V. M. (1983). *Test for the Reception of Grammar.* Unpublished manuscript.

Bloom, L., Lifter, K., & Hafitz, J. (1980). Semantics of verbs and the development of inflection in child language. *Language, 56*, 386–412.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics, 10*, 151–171.

Clahsen, H., & Dalalakis, J. (1999). Tense and agreement in Greek SLI: A case study. *Essex Research Reports in Linguistics, 24*, 1–25.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Csányi, F. I. (1974). *Peabody Szókincs-Teszt*. Budapest: Bárczi Gusztáv Gyógypedagógiai Föiskola.

de Jong, J. (1999). *Groningen dissertations in linguistics: Vol. 28. Specific language impairment in Dutch: Inflectional morphology and argument structure*. Groningen: University of Groningen.

Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht: Kluwer Academic.

Fletcher, P., Leonard, L., Stokes, S., & Wong, A. M.-Y. (2005). The expression of aspect in Cantonese-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*, 621–634.

Hansson, K., Nettelbladt, U., & Leonard, L. (2000). Specific language impairment in Swedish: The status of verb morphology and word order. *Journal of Speech, Language, and Hearing Research, 43*, 848–864.

Kas, B., & Lukács, Á. (2011). *Magyar Mondatutánmondási Teszt*. Manuscript in preparation.

Kelly, D., & Rice, M. (1994). Preferences for verb interpretation in children with specific language impairment. *Journal of Speech and Hearing Research, 37*, 182–192.

Kiefer, F. (1994). Aspect and syntactic structure. In F. Kiefer & K. Kiss (Eds.), *Syntax and semantics: Vol. 27. The syntactic structure of Hungarian* (pp. 415–464). San Diego, CA: Academic Press.

Krantz, L., & Leonard, L. (2007). The effect of temporal adverbials on past tense production by children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 50*, 137–148.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L., & Deevy, P. (2010). Tense and aspect in sentence interpretation by children with specific language impairment. *Journal of Child Language, 37*, 397–418.

Leonard, L., Deevy, P., Kurtz, R., Krantz Chorev, L., Owen, A., Polite, E., et al. (2007). Lexical aspect and the use of very morphology by children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 50*, 759–777.

Leonard, L., Eyer, J., Bedore, L., & Grela, B. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 741–752.

Leonard, L., Sabbadini, L., Leonard, J., & Volterra, V. (1987). Specific language impairment in children: A crosslinguistic study. *Brain and Language, 32*, 233–252.

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

Marchman, V., Wulfeck, B., & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research, 42*, 206–219.

Moore, M., & Johnston, J. (1993). Expressions of past tense by normal and language-impaired children. *Applied Psycholinguistics, 14*, 515–534.

Norbury, C. F., Bishop, D. V. M., & Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *Journal of Speech, Language, and Hearing Research, 44*, 165–178.

Oetting, J., & Horohov, J. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 62–74.

Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: An extended optional default account. *Language Acquisition, 9*, 269–300.

Penner, Z., Schulz, P., & Wymann, K. (2003). Learning the meaning of verbs: What distinguishes language-impaired children from normally developing children? *Linguistics, 41*, 289–319.

Pléh, Cs. (1992). Verbal prefixes in Hungarian children. In I. Kenesei & Cs. Pléh (Eds.), *Approaches to Hungarian* (Vol. 4, pp. 269–281). Szeged: JATE.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, Cs. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle, 4*, 479–505.

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

Redmond, S. (2003). Children's productions of the affix *-ed* in past tense and past participle contexts. *Journal of Speech, Language, and Hearing Research, 46*, 1095–1109.

Rice, M. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of SLI* (pp. 63–94). Mahwah, NJ: Erlbaum.

Rice, M., Noll, K. R., & Grimm, H. (1997). An extended optional infinitive stage in German-speaking children with specific language impairment. *Language Acquisition, 6*, 255–295.

Rice, M., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*, 1239–1257.

Rice, M., & Wexler, K. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. San Antonio, TX: Psychological Corporation.

Rice, M., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1412–1431.

Schulz, P., & Wittek, A. (2003). Opening doors and sweeping floors: What children with specific language impairment know about telic and atelic verbs. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (Vol. 2. pp. 727–738). Somerville, MA: Cascadilla Press.

Shirai, Y., & Andersen, R. (1995). The acquisition of tense–aspect morphology: A prototype account. *Language, 71*, 743–762.

Wagner, L. (2001). Aspectual influences on early tense comprehension. *Journal of Child Language, 28*, 661–681.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint. *Lingua, 106*, 23–79.

Wexler, K. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 11–61). Mahwah, NJ: Erlbaum.

Study 6.

Lukács, Á., Leonard, L. B., Kas, B. (2013). Case marking in Hungarian children with specific language impairment. *First Language*, 33/4. 331-353. doi: 10.1177/0142723713490601

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6.)

Thesis 5. Problems with case marking in SLI suggest lexical and processing deficits instead of a selective case marking impairment within grammar (6)

Article

FIRST LANGUAGE

First Language
33(4) 331–353

DOI: 10.1177/0142723713490601
fla.sagepub.com
SAGE

Case marking in Hungarian children with specific language impairment

Ágnes Lukács
Hungarian Academy of Sciences, Hungary

Bence Kas
Hungarian Academy of Sciences, Hungary; Eötvös Loránd University of Sciences, Hungary

Laurence B. Leonard
Purdue University, USA

Abstract
This study examines whether children with specific language impairment (SLI) acquiring a language with a rich case marking system (Hungarian) have difficulty with case, and, if so, whether the difficulty is comparable for spatial and nonspatial meanings. Data were drawn from narrative samples and from a sentence repetition task. Suffixes were tested both in their spatial and nonspatial meanings. Participants with SLI were compared to same-age peers and younger typically developing children matched on receptive vocabulary scores (VC children). Results show that although case marking errors are very rare in spontaneous speech in Hungarian children with SLI, the number of case-marked nouns and of different case markers is significantly lower in children with SLI. In the elicited production task, overall performance of the children with SLI was significantly below that of VC children, but children with SLI and VC children scored higher with spatial than with nonspatial meanings. The results are in line with expectations based on processing accounts which posit greater difficulties with less transparent details of grammar.

Keywords
Case marking, Hungarian, morphology, production, specific language impairment

Corresponding author:
Ágnes Lukács, BME Department of Cognitive Science, Budapest and Research Institute of Linguistics, Hungarian Academy of Sciences, Egry J. u, 1, H-1111, Budapest, Hungary.
Email: alukacs@cogsci.bme.hu

Problems with grammatical morphology are widely documented in children with specific language impairment (SLI) in many languages.[1] Ample attention has been devoted to examining problems with verb morphology, especially with tense and agreement morphemes, and there are studies on the grammatical encoding of aspect as well (e.g., Fletcher, Leonard, Stokes, & Wong, 2005). The severity of problems with verb morphology seems to be dependent on language type: the grammatical marking of tense and agreement has been shown to be extremely difficult for English-speaking children with SLI (e.g., Bedore & Leonard, 1998; Rice & Wexler, 1996, 2001), but it is relatively easier in languages with a rich system of morphemes (Spanish, Italian, Hebrew, Hungarian, e.g., Bedore & Leonard, 2001; Bortolini, Caselli, & Leonard, 1997; Leonard, Caselli, & Devescovi, 2002; Leonard & Dromi, 1994; Lukács, Leonard, Kas, & Pléh, 2009).

Much less is known about noun morphology in SLI. This may be because the majority of studies on SLI have focused on English and, in this language, verb morphology problems appear to be more striking than errors in noun morphology. However, in recent years, studies of SLI in languages with a rich noun morphology have begun to appear. Hungarian – the focus of the present investigation – is one such language. The present investigation examines the use of case markers by children with SLI speaking Hungarian. This language is an agglutinative language where multiple suffixation is possible. Studying case marking in Hungarian-speaking children with SLI is motivated by several factors. First, one of the earliest case studies of language impairment in Hungarian by Vinkler and Pléh (1995) identified case marking as a problematic area, and thus a potential marker of language impairment in Hungarian. The child examined by Vinkler and Pléh often substituted target suffixes with a more frequently occurring suffix (e.g., replacing the instrumental with the accusative). Second, Hungarian as an agglutinative language with a rich case morphology offers a unique opportunity for testing sensitive areas within noun morphology generally and within case marking more specifically. Third, accounts of grammatical deficits in SLI have devoted relatively little attention to case morphology and, as will be seen, this attention has focused on but one type of case, leaving much unexplained. For these reasons, we examined the use of case markers in spontaneous speech and in an elicited production task in Hungarian children with SLI. We also compared knowledge of case markers both in their spatial/semantically transparent and nonspatial/semantically opaque use to compare how function influences the use of the same form in typical development and in SLI.

Case marking in SLI: Data and theories

Studies of case marking by children with SLI have largely concentrated on structural case, that is, case assigned by particular positions in the syntactic structure. The widest documentation of problems with structural case in SLI comes from studies of subject pronoun errors by English-speaking children with SLI. These children tend to make more errors than typically developing (TD) peers, but the pattern of error types is the same: they overapply accusative forms for nominative forms (as in **Me drink it all*). In contrast, accusative forms (as in *Mommy hugged me*) are produced mostly correctly (Clahsen, Bartke, & Göllner, 1997; Leonard, 1982, 1995; Loeb & Leonard, 1988, 1991; Radford & Ramos, 2001; Schütze, 1997; Wexler, Schütze, & Rice, 1998). Subject case marking in English is often suggested to correlate with the development and marking of subject–verb agreement in typical development and in SLI as well. Indeed, Wexler et al. (1998), in their formulation of the Agreement and Tense Omission Model (ATOM), suggest that case marking difficulties of children with SLI are a corollary of their primary deficit in the use of tense and agreement marking of the verb. That is, in English, nominative case is licensed by the presence of the functional category of Agreement (AGR); when this functional category is not projected, nominative case is not possible and the default case (which is accusative in English) is selected.

The overapplication of the accusative is not observed in every language: Clahsen (1991) documents errors in German-speaking children with SLI, for whom the majority of errors consisted of substitutions of the nominative for the accusative and dative forms in NPs and pronouns (but other substitutions also occurred). However, in German, unlike English, the default case is nominative, not accusative.

Although errors of nominative case were the most prominent among German-speaking children with SLI, the fact that case errors in accusative and dative contexts were observed prompted Clahsen (1991) to propose that these children have a broader agreement deficit – one that adversely affects not only subject–verb agreement, but also the assignment of other types of case (see also Roberts & Leonard, 1997). Later studies of German by Clahsen and his colleagues (Eisenbeiss, Bartke, & Clahsen, 2005; Rothweiler & Clahsen, 1994) have led to a modification of the original agreement deficit account of SLI. These investigators compared performance on structural (nominative and accusative) and lexically bound case marking in children with SLI and controls matched for mean length of utterance (MLU). Children with SLI tended to mark structural case (nominative and accusative) correctly, but they incorrectly overapplied these to other lexically bound cases such as the obligatory dative. For example, **helf den Frau* was produced, where instead of the accusative (and masculine) *den* determiner the correct choice would have been the dative *der* (correct: *helf der Frau* 'helps the woman'). Another example was **ich bin kalt*, where the nominative pronoun was incorrectly used instead of the dative experiencer (correct: *Mir ist kalt* 'I'm cold'). In spite of frequent substitution errors, children with SLI were not selectively impaired on lexical case either: they performed at the same level as their MLU controls. So German children with SLI did not show an impairment of either structural or lexical case relative to their MLU controls, although they had relatively low scores on agreement marking. The authors concluded (modifying the previous general agreement deficit account including case) that these results argue against a broad agreement deficit and support models of SLI that propose a syntactic deficit in subject–verb agreement and not in areas of grammar such as case or tense marking.

An important set of data comes from Turkish, a language that is similar to Hungarian in that case marking and not word order is the core marker of grammatical functions. Results from monolingual and bilingual children with SLI suggest that in Turkish, noun morphology is more vulnerable than verb morphology. Çavuş (2009) tested structural and semantic case marking in Turkish bilingual children with SLI and did not find difficulties with structural (dative and accusative) case (in line with Eisenbeiss et al., 2005). She also points out that the bilingual SLI group produced utterances with fewer semantic case (semantic dative, locative, genitive, ablative) contexts, but they did not make more errors than TD children when they used these case forms. Results of De Jong, Çavuş, and Baker (2010) from Turkish–Dutch bilingual children suggest that bilingual children with SLI have greater problems with noun morphology in Turkish (case marking, accusative: 56% correct) than with verb morphology (89% correct). Rothweiler et al. (2010) also found case marking deficits in four Turkish-German successive bilingual children with SLI, but impairments of verb morphology were also evident. Importantly, regardless of whether they have deficits in verb morphology, children with SLI who speak Turkish seem to have problems in the area of case marking.

Lukács, Leonard, and Kas (2010) tested knowledge of accusative forms (together with plurals) in noun morphology in Hungarian-speaking children with SLI. There was no evidence of a special difficulty marking the accusative in SLI, and when children made simplification errors in forms where multiple suffixation was required (plural+accusative), the children were more likely to produce the accusative only, suggesting that structural case marking is not impaired in Hungarian SLI.

Although not a study of either children with SLI or case marking, Friederici's (1982) results might also be relevant to our questions. She tested knowledge of German prepositions in their semantic versus syntactic function in patients with aphasia, and found that in Broca's aphasia prepositions that appear in their spatial or semantic use were easier to produce than prepositions that only have a syntactic function (even if they have the same form), while she found the reverse pattern in Wernicke's aphasia. Based on Friederici's observations, we would expect to find the pattern observed in Broca's aphasia to characterize SLI performance: suffixes in their spatial meaning should be easier than suffixes that only have a syntactic function. Friederici (1982, pp. 251–252) argues that,

> The German language allows for investigation of the processing of structural and semantic information using the *same form* of a closed class item by only varying its functional role. For example, a preposition can be used, first, as a lexical preposition, that is, as a freely substitutable form of a preposition that bears at least some semantic information. … Second, the same form can also be used as an obligatory preposition that is lexically dependent on the preceding verb, that is, the verb is subcategorized for a particular preposition. … These obligatory prepositions bear virtually no semantic meaning by themselves, but are nonetheless a structural requirement.

The same is true in other languages, including English prepositions and, most relevant to our study, Hungarian case markers.

Most accounts of the grammatical deficits of SLI appear to be silent with regard to distinctions such as that made by Friederici (1982). The Representational Deficit for Dependent Relations (RDDR) account proposed by van der Lely and her colleagues (van der Lely, 1994; van der Lely & Stollwerck, 1997) proposes that children with SLI have a broad syntactic computational deficit that leads to weaknesses in structure-dependent relationships. This account has since been expanded and is termed the Computational Grammatical Complexity Hypothesis (CGC, Marshall & van der Lely, 2007). According to this expanded account, children with SLI have a deficit in structural complexity that extends beyond syntax to include morphology and phonology as well. The CGC hypothesis also argues that not all kinds of dependent relations are expected to be impaired in SLI, only complex ones involving movement chain formations. To test the predictions of CGC and contrast them with those of other theories, Stavrakaki and van der Lely (2010) tested production and comprehension of clitics and anaphors in Greek-speaking children with SLI, and found that SLI performance was only significantly worse with object clitics. Object clitics solely rely on syntactic dependencies for their interpretation (as opposed to pronouns with an independent inherent semantic reference). Importantly, children with SLI performed well on anaphors, which are similar to object clitics in that they are non-salient and they rely on syntactic information, but they only require a core grammatical operation of feature checking within spec-head agreement, and not complex dependencies requiring movement. However, it is not clear whether this assumption of a structural complexity deficit applies to non-structural case suffixes that are dictated by the characteristics of the verb.

One account that allows for a distinction between semantically based and non-semantically based forms in SLI is the morphological richness account proposed by Leonard and his colleagues (Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999; Leonard, 1998, pp. 255–257; Leonard, Sabbadini, Leonard, & Volterra, 1987). According to this account, children with SLI have a limited processing capacity and devote their limited processing resources to the dominant features of the language they are acquiring. In English, children with SLI devote resources to word order and remaining resources are devoted to grammatical morphology. However, in a morphologically rich language (especially when word order is not rigid), resources are devoted first to grammatical morphology. This state of affairs means that differences between children with SLI and typically developing peers in the use of grammatical morphology will be smaller in a morphologically rich language than in a morphologically sparse language.

However, given the limitations in processing capacity in children with SLI, the benefits of a rich morphology will have limits. Grammatical morphemes with relatively transparent functions will be acquired first, placing morphemes with less transparent functions at risk for being processed incompletely due to limited resources. As a result, it will take a greater number of encounters with these less transparent forms before children with SLI acquire them. As will be seen, this distinction between transparent and opaque functions of morphemes has particular relevance in Hungarian.

Case morphology in Hungarian

In a non-configurational language like Hungarian, where word order is relatively free, morphology is the core marker of grammatical functions. Hungarian has a very rich system of suffixes both in the verbal and the nominal paradigm. Suffix combinations are possible and frequent; the order of suffixes within a word is fixed. Theoretically, there are 756 different forms in which a noun can appear, taking all possible suffixes and their well-formed combinations into account. If the nominative is assumed to have a zero case marker, there are 18 cases. Case suffixes can combine with the plural and with possessive markers, in a fixed order: the case marker is always word-final, and all nouns have to end in a case marker. Like most suffixes, case markers can have several allomorphs and the choice of the allomorph is determined by the stem and the rules of vowel harmony: the suffix agrees with the stem vowels in frontness (and in some cases in roundness as well).[2]

In line with Chomsky's (1981) distinction between structural and lexical case and earlier proposals for other languages, Kiefer (2000) and Bartos (2000) describe the nominative, the accusative and the dative as syntactic (or structural) cases (assigned by particular positions of the sentence structure), and all the others as lexical (assigned by lexical specifications of predicates) or inherent (associated with specific thematic roles) cases. According to this approach, the cases we examine in the study presented below are all lexical or inherent cases. None of these cases requires agreement with a *grammatical* feature of any other element in the sentence. All of them require agreement with semantic (of both the verb and the case-marked noun in the case of spatial meanings) or lexical (of the verb in the case of nonspatial meanings) features of other elements. Some syntactic accounts of non-structural cases propose a further distinction between lexical and inherent cases. Woolford (2006) argues that there are lexical cases, which are truly irregular and selected by individual verbs, and the more regular inherent cases. This subdivision within non-structural cases is justified by licensing differences, according to Woolford: lexical Case is restricted to themes or internal arguments (licensed by V inside the VP proper at vP structure), while inherent Case only appears with external arguments (licensed by so-called little/light verbs above the VP proper at vP structure).

Although application of this linguistic subdivision for Hungarian case markers awaits further study, it is true that even within the group of 'lexical' or 'inherent' cases, the selection of a case marker may be determined by one of two different processes in Hungarian, very much like the selection of prepositions in English or German. First, the choice of suffix may be governed directly by the idiosyncratic lexical specifications of the predicate. In this instance, the case marker 'loses' its spatial meaning, as in example (a) below. The other process involves indirect selection, where the predicate subcategorizes for an obligatory or optional argument of a certain thematic type, which may be marked by one of a set of suffixes (it constrains the path type of the suffix: whether it should be a Goal, Static or Source suffix). The choice of suffix from within this set is determined by the properties of the noun host (constraining the spatial relation type of the suffix: whether it is going to be a Container, Surface or Neighborhood suffix), see example (b) below.

(a) *Pisti tanult a baleset**ből***

Pisti learned the accident-FROM.

Pisti learned from the accident.

(b) *Az oroszlán megszökött a ketrec**ből**.*

The lion escaped the cage-FROM.

The lion escaped from the cage.

While in sentence (a), the elative suffix is selected on the basis of idiosyncratic lexical specification of the verb, in sentence (b) semantic restrictions of the verb and the noun cooperate: the verb *megszökik* 'escape' only requires that the noun have a Source-type suffix. This information combines with the specifications by the noun *ketrec* 'cage' which is a container, unambiguously specifying the elative as the right suffix choice.

Predictions of different accounts of SLI

Relatively few accounts have provisions for predicting the status of lexical or inherent case in children with SLI. Clahsen (e.g., Eisenbeiss et al., 2005) modified his account to emphasize subject–verb agreement limitations, and Rice and Wexler's (1996, 2001) account focuses on tense and subject–verb agreement. The computational grammatical complexity hypothesis may be relevant to lexical or inherent case, but thus far van der Lely and her colleagues have not yet outlined how such a weakness would be treated in their account.

On the other hand, the morphological richness account has provisions for expecting a milder deficit in case use by children with SLI acquiring Hungarian relative to those acquiring a morphologically sparse language. Furthermore, and, importantly, within Hungarian, children with SLI should lag behind their typically developing peers to a greater degree in their use of those case markers with relatively opaque functions than in their use of case markers with relatively transparent functions. As operationalized in the present study, we assume that suffixes expressing spatial meanings would be relatively transparent, and hence less problematic, and suffixes expressing nonspatial meanings would be more opaque and reveal larger differences between children with SLI and their typically developing peers.

We tested this hypothesis in two studies, applying different methods. In Study 1, we analyzed corpora of narrative language samples from children with SLI and two groups of TD children: one group matched on chronological age and the other on vocabulary size; Study 2 was an elicited production task disguised in the form of a sentence repetition task with masked inflections.

Study 1: Analysis of narratives

Method

Participants. We analyzed narrative samples from 16 children with SLI. All 16 of these participants met exclusionary and inclusionary criteria for SLI. Each child had normal intelligence (above 85 on the Raven Colored Progressive Matrices; Raven, Court, & Raven, 1987), passed a hearing screening, and had no history of neurological impairment. Each child in the SLI group scored at least 1.5 SDs below the mean for his or her age on two or more of four language tests administered, two of which assessed receptive skills, and two evaluated expressive skills. The receptive tests were the Hungarian standardizations of the Peabody Picture Vocabulary Test (PPVT) and the Test for the Reception of Grammar (TROG). The expressive tests were the Hungarian Sentence Repetition Test, and a nonword repetition test.

The Hungarian adaptation of the PPVT is modeled closely after its English equivalent (Csányi, 1974; Dunn & Dunn, 1981). The Hungarian adaptation of the original TROG (Bishop, 1983, 2012; Lukács et al., 2012) has been standardized for the age range 4-12 years.[3] The test assesses the children's understanding of grammatical forms in increasing difficulty. It consists of 20 blocks, each with four items of the same grammatical construction (e.g., sentences with comparatives, postmodified subjects and embedded clauses). The child must point to a picture (from an array of four pictures) that matches the sentence spoken by the experimenter. A block is scored as complete if the child responds correctly to all four pictures in the block. We used both the number of blocks correctly completed (max. 20), and the number of items correctly answered (max. 80).

The Hungarian Sentence Repetition Test (*Magyar Mondatutánmondási Teszt*; Kas & Lukács, in press) contains 40 sentences, distributed evenly across five types of grammatical constructions. These are simple Subject-Verb-Object (SVO) sentences, simple OVS sentences, and complex sentences containing SS relative clauses, SO relative clauses, and OS relative clauses. The sentences vary in length from eight to 15 syllables within each type. The task is to immediately and accurately repeat each sentence presented by the experimenter. Accuracy is measured in terms of the number of correctly repeated sentences.

The nonword repetition test (Racsmány, Lukács, Németh, & Pléh, 2005) consists of 36 nonwords ranging from one to nine syllables in length. Four nonwords are used for each length. All nonwords conform to the phonotactic patterns of Hungarian. The particular phonological sequences used in the nonwords do not reflect the frequency distribution of Hungarian phoneme sequences. However, sequences that are articulatorily difficult for speakers are not employed. The child is asked to repeat each nonword. The child's score is defined as the length at which the child successfully repeated at least two of the four nonwords.

We compared results of the SLI group to two control groups matched individually to the children with SLI: a group matched on chronological age and gender (hereafter, the CA group) and a group matched on receptive vocabulary raw scores (PPVT and gender). Vocabulary controls were chosen to test whether case marking deficits in SLI (if they exist) exceeded children's vocabulary limitations. These children scored above −1 SD on each of the four language tests. Hereafter, these children are referred as the vocabulary control (VC) children. Details of the groups are shown in Table 1.

Procedure. We elicited narrative language samples using Mayer's (1969) Frog Story. All participants were tested individually in a quiet room. Children were allowed to go through the pages of the picture book first, and were then prompted to tell the story in their own words. Neither questions nor directives were given during the task, to promote continuous narratives from the children.

Measures. Several measures were counted in the samples of each child: the total number of case marker suffixes (tokens), the number of different case markers (types) and the total number of case marking errors (note that nouns in the nominative do not have an overt suffix; therefore, we did not include nominative case in the scoring). Two kinds of case marking errors were differentiated: (1) omission of a case marker in an obligatory context, and (2) substitution of an obligatory case marker with another. Erroneous case marking was only considered as a substitution or an omission when the verb frame requiring a specific case marker was fully identifiable.

Table 1. Mean age and receptive vocabulary raw scores (with standard deviations) by language group and age.

	SLI	VC	CA
Mean age (SD)	6.4 (.7)	5.4 (1.0)	6.3 (.7)
Mean receptive vocabulary score (PPVT, SD)	78.8 (16.5)	79.1 (15.7)	89.3 (11.2)

SLI: children with SLI, VC: vocabulary control group, CA: chronological age control group.

Results

Because of the relatively large number of case marker types in Hungarian, the number of tokens of any particular type was small, often zero. Therefore, instead of analyzing case marker types separately we only analyzed group effects on the total numbers of case marker tokens, types and errors. First, we examined whether vocabulary level and general length differed among groups. According to a one-way ANOVA there were no significant differences among groups in the total number of words, $F(2, 45) = 2.14$, n.s., and the total number of nominative nouns, $F(2, 45) = .40$, n.s. Thus, general length and vocabulary did not differentiate the groups.

Next, we compared the groups' productions of case marker suffixes and different case markers. We compared the performance of children with SLI with the CA and the VC group in a one-way multivariate ANOVA conducted on the total number of case marker suffixes (tokens) and on the number of different case markers (types). Multivariate tests showed a significant main effect of Group, $F(4, 88) = 3.20$, $p < .05$, $\eta_p^2 = .127$. Univariate tests revealed significant effect of Group for both the total number of case marker suffixes (tokens), $F(2, 45) = 4.37$, $p < .05$, $\eta_p^2 = .163$, and the number of different case markers (types), $F(2, 45) = 6.22$, $p < .01$, $\eta_p^2 = .217$. Bonferroni-corrected pairwise comparisons (post-hoc tests) showed that children with SLI only differed from the CA group in both respects, that is, they used a significantly smaller number of case marker suffix tokens and also a smaller number of different case markers (types) than the CA group, while they did not differ from the VC group in either respect.

Finally, we analyzed the grammaticality of children's use of case markers by comparing the number of case marking errors among the groups. A one-way ANOVA conducted on the number of case marking errors showed no significant differences, $F(2, 45) = .52$, n.s. The main results for the different measures in the three groups are presented in Table 2.

Summary of results from the narrative analysis

In a narrative task, children with SLI used significantly fewer case-marked nouns (case marker tokens) and fewer different types of case marker suffixes than their age-equivalent peers. However, comparisons with younger children with the same level of receptive vocabulary showed no such difference: children with SLI used the same number of case-marked nouns (tokens) and the same number of different case markers (types) as typically developing children matched on vocabulary. Their level of accuracy of case marker production was comparable to VC children, with very few case marking errors observed.

Table 2. Means (and standard deviations) for total number of words, number of different nominative case nouns, total number of case marker suffixes (tokens), number of different case markers (types) and total number of case marking errors in the narrative samples by group.

	SLI	VC	CA
Total number of words	226.56 (107.54)	303.5 (129.98)	307.31 (134.73)
Number of different nominative case nouns	16.2 (9.48)	18.69 (7.15)	17.88 (7.44)
Total number of case marker suffixes (tokens)	19.18_{a} (7.49)	$25.56_{a,c}$ (10.31)	29.12_{c} (10.75)
Number of different case markers (types)	5.88_{a} (1.54)	$6.93\ (2.11)_{a,c}$	8.00_{c} (1.36)
Total number of case marking errors	.38 (.61)	.37 (.81)	.69 (1.40)

SLI: children with SLI, VC: vocabulary control group, CA: chronological age control group.
Means within a row lacking a common subscript differ significantly, $p < .05$.

Thus, in the narrative task, the overall case marking performance of children with SLI matches levels expected based on their vocabulary size.

Study 2: Elicited production with masked inflections

Method

Participants. Ninety-two children participated in the second study. Forty-six children met the criteria for SLI and 46 were developing language in a typical manner. Children with SLI were recruited in two age groups. The younger group consisted of 17 participants, ranging in age from 4;10 to 7;2 (M = 6;0), the older group consisted of 29 children ranging in age from 7;11 to 11;4 (M = 9;10). The younger participants were selected from two special kindergarten classes for children with language impairment. The older participants were selected from two special schools for children with language impairment. All 46 children with SLI met exclusionary and inclusionary criteria for SLI described in Study 1.

The 46 typically developing children were selected for two vocabulary control groups (VC) matched individually on receptive vocabulary raw scores (based on their raw scores on the PPVT, and criteria described in Study 1 above). Seventeen of the children (younger VC group) were matched according to PPVT raw scores with the younger SLI group, ranging in age from 3;3 to 6;2 (M = 5;1). The remaining 29 children (older VC group) were matched according to PPVT scores with the older SLI group, and ranged in age from 4;4 to 8;2 (M = 6;3). It can be seen that these two groups of VC children overlapped in age; for this reason the designation 'younger' and 'older' should be read as 'matched to the younger' and 'matched to the older', respectively. A VC child was considered a match if his or her PPVT raw score was within 4 points of the PPVT raw score of a child in the SLI group. Means and ranges for age, PPVT raw score, nonword repetition span, TROG raw score and Sentence Repetition raw score for the four groups are provided in Table 3.

Procedure. We employed a structured elicited production task in which children were instructed to repeat sentences spoken by a female speaker, heard from a loudspeaker. The

Table 3. Means and ranges for age and PPVT raw score for the four groups.

	Age	PPVT raw score	Nonword span	TROG raw score	Sentence repetition
SLI young	5.97 (4.83–7.16)	74.23 (45–110)	3.26 (0–7)	60.82 (46–73)	9.25 (0–21)
VC young	5.07 (3.25–6.16)	74.29 (48–106)	5.14 (4–7)	66.5 (51–80)	28.42 (11–40)
SLI old	9.84 (7.9–11.33)	90.48 (63–124)	3.51 (0–7)	68.1 (59–78)	25.2 (9–39)
VC old	6.28 (4.33–8.16)	90.93 (62–126)	5.36 (3–7)	69.82 (48–80)	31.5 (16–40)

VC: vocabulary control, SLI: specific language impairment.

sentences were digitized, with coughs inserted to replace the inflections only, as in the example below, where *XXX* marks the cough:

*Az oroszlán megszökött a ketrec***XXX.**

Target: *Az oroszlán megszökött a ketrec**ből**.*

The lion escaped the cage-FROM.

The lion escaped from the cage.

Since participants heard the whole sentence up to the suffix, including the final noun stem, and they only had to supply the suffix, they could rely on the combined information from the verb and the noun. All sentences were normalized for a length between eight and 12 syllables. The nouns used with the spatial and nonspatial meanings of the suffixes were different, but they were all one- to three-syllable nouns that the children were familiar with. Participants did not have a problem repeating the nouns and verbs or any other unmasked part of the sentence.[4] This method was based on the phoneme restoration effect demonstrated by Warren (1970), which also works at the level of morphemes (e.g., for affixes in Hungarian; see Dankovics & Pléh, 2001). We have already successfully exploited this method in an earlier developmental study as an elicited production method testing knowledge of agreement in Hungarian children with SLI (Lukács et al., 2009) and in testing early morphological productivity in young typically developing children (Gábor & Lukács, 2012). Children were tested individually in a quiet room.

We used a stimulus design that was originally developed to test the use of spatial language without the confounding factor of spatial cognition (Lukács, Pléh, & Racsmány, 2007). The target sentences included spatial and nonspatial meanings of all suffixes. Each of nine suffix types was represented by three sentences, for a total of 27 sentences (see Table 4 for examples). Spatial case markers can be viewed from two perspectives. First, they can be viewed from the perspective of Path type (Goal, Source and Static). Second, they can be viewed from the perspective of Relation type (Surface, Container, Neighborhood) (Table 5). Path type and Relation type descriptors derive from the different uses when the case markers are used in the spatial sense. However, nonspatial markers also pattern in the same way based on formal grounds. Specifically, when a prefix associated with container/goal is present, a Container/Goal case suffix is required even though a literal spatial relationship is not involved (e.g. *Ella teljesen* bele*ásta magát T. S.*

Table 4. Examples of sentences used in the sentence completion task.

	Spatial	Nonspatial
ban/ben	*A kismadarak ott vannak a fészek**ben**.* The birds are there in the nest.	*Kristóf hisz az angyalok**ban**.* Kristóf believes in angels.
ba/be	*Nagyi elment a templom**ba**.* Grandma went to church.	*A nagynéni szerelmes a király**ba**.* Auntie is in love with the king.
ból/ből	*Az oroszlán megszökött a ketrec**ből**.* The lion escaped from the cage.	*A tanárnak elege lett a sajt**ból**.* The teacher got tired of the cheese
on/en/ön	*Az autó átment a híd**on**.* The car crossed the bridge.	*Ildikó meglepődött az ajándék**on**.* Ildikó was surprised at the present.
ra/re	*A kertész felállt a létrá**ra**.* The gardener stepped up the ladder.	*Pisti emlékezett a kirándulás**ra**.* Pisti remembered the trip.
ról/ről	*A cserepek leestek a tető**ről**.* The tiles fell off the roof.	*Mindenki hallott már a delfinek**ről**.* Everybody has heard of dolphins.
nál/nél	*A busz megállt a piros lámpá**nál**.* The bus stopped at the red light.	*A nyúl gyorsabban fut a csigá**nál**.* The rabbit runs faster than the snail.
hoz/hez/höz	*Péter elment a fogorvos**hoz**.* Péter visited the dentist.	*Károly csatlakozott a kirándulás**hoz**.* Károly joined the trip.
tól/től	*Nagyi visszajött az orvos**tól**.* Grandma came back from the doctor's.	*A húgom nagyon fél a halak**tól**.* My sister is very much afraid of fish.

Table 5. Target suffixes in the sentence completion task.

	Static	Goal	Source
Container	*-ban/ben* in Inessive	*-ba/be* into Illative	*-ból/ből* out of Elative
Surface	*-on/en/ön* on Superessive	*-ra/re* onto Sublative	*-ról/ről* off Delative
Neighborhood	*-nál/nél* at Adessive	*-hoz/hez/höz* to Allative	*-tól/től* from Ablative

*Eliot költészeté**be***.- Ella completely into-dig-3SgPastDef herself-Acc T. S. Eliot poetry-Poss-Illative, 'Ella completely immersed herself in T. S. Eliot's poetry'). In our detailed analysis of performance patterns, we examine the results from both the perspective of Path type and of Relation type, for both spatial and nonspatial meanings.

Scoring. Although children were asked to repeat the entire sentence they heard, given our focus on case marking, only the suffixes on the nouns were scored as correct or incorrect. All correct answers were given a score of 1, incorrect answers were scored as 0. Correct answers included target suffixes, and also included some deviations from target answers. That is, with some sentences and structures two suffixes may be in free variation (or in some cases dialectal variation) with the meaning of the structure preserved (*A mamut hasonlít az elefánt**ra*** vs. *elefánt**hoz***. 'The mammoth resembles the elephant', where elephant-ONTO and elephant-TO are both acceptable). Substitutions that resulted in a slight change in sentence meaning relative to the target were also accepted, provided that the child's sentence described the situation appropriately (*A katona hátralépett a kapu**tól** vs kapu**ból***. 'The soldier stepped back from the gate' where both gate-FROM and gate-OUT OF are acceptable.

Results

Two sets of analyses were performed, to accommodate the fact that Path type and Relation type represented two different dimensions that could not be combined in a single analysis due to too few items for each Path type–Relation type permutation.

Analyses employing path type. Correct scores were first analyzed in a 2 (Group: SLI, VC) × 2 (Age) × 2 (Suffix meaning: Spatial, Nonspatial) × 3 (Path type: Static, Source, Goal) design. All four main effects were significant: Group, $F(1, 88) = 28.79$, $p < .001$, $\eta_p^2 = .247$; Age, $F(1, 88) = 14.05$, $p < .001$, $\eta_p^2 = .137$; Suffix meaning, $F(1, 88) = 89.30$, $p < .001$, $\eta_p^2 = .504$; and Path type $F(2, 176) = 28.60$, $p < .001$, $\eta_p^2 = .24$. There were three significant two-way interactions: Suffix meaning × Group, $F(1, 88) = 6.79$, $p < .05$, $\eta_p^2 = .72$; Suffix meaning × Age $F(1, 88) = 5.92$, $p < .05$, $\eta_p^2 = .063$; and Suffix meaning × Path type, $F(2, 176) = 14.45$, $p < .001$, $\eta_p^2 = .141$. However, these interactions were subsumed by two significant three-way interactions: Suffix meaning × Group × Age: $F(1, 88) = 3.97$, $p < .05$, $\eta_p^2 = .043$; and Suffix meaning × Path type × Age: $F(2, 176) = 3.74$, $p < .05$, $\eta_p^2 = .041$.

The main effects reflected differences that were quite straightforward. The VC children were more accurate than the children with SLI; older children were more accurate than younger children, and Spatial meanings were more accurate than Nonspatial meanings. Pairwise comparisons for Path type revealed that Goal suffixes were more accurate than Source suffixes which, in turn, were more accurate than Static suffixes. However, a more complete understanding of these findings requires inspection of the observed interactions. We focus here on the two three-way interactions.

Because both three-way interactions involved Group, we examined the effects for each age level separately. When the Suffix meaning × Group × Age interaction was broken down by age level, we found that Suffix meaning and Group interacted only in the younger children, $F(1, 32) = 5.69$, $p < .05$, $\eta_p^2 = .151$. Within this younger group, the advantage of Spatial meanings over Nonspatial meanings was present in both the SLI and TD groups, but stronger among the children with SLI, $t(28) = 4.05$, $p < .001$ than among the TD children, $t(28) = 3.76$, $p < .01$.

We also examined the Suffix meaning × Path type × Age interaction in greater detail by computing the effects for each age level separately. The interaction of Suffix meaning and Path type was significant in the younger group, $F(2, 64) = 8.76$, $p < .001$, $\eta_p^2 = .215$, as well as in the older group, $F(2, 112) = 6.53$, $p < .01$, $\eta_p^2 = .105$. For the younger children, when the suffixes had a Spatial meaning, Goal suffixes were marginally easier than both Source and Static suffixes ($p < .10$ for both). When the suffixes had a Nonspatial meaning, Goal and Source suffixes did not differ, and they were both more accurate than Static suffixes ($p < .001$). For the older children, when the suffixes had a Spatial meaning, performance on Goal suffixes was superior to performance on Static suffixes ($p < .01$). There was a tendency for Goal suffixes to be more accurate than Source suffixes, but the difference did not reach significance ($p = .156$). Source and Static suffixes did not differ. When the suffixes had a Nonspatial meaning, Goal suffixes were more accurately produced than Source suffixes ($p < .05$), which, in turn, were used more accurately than Static suffixes ($p < .05$).

Analyses employing relation type. As noted earlier, separate analyses for Path type and Relation type were required. For the latter, we performed an ANOVA with a 2 (Group: SLI, VC) × 2 (Age) × 2 (Suffix meaning: Spatial, Nonspatial) × 3 (Relation type: Container, Surface, Neighborhood) design. Because Group, Age and Suffix meaning did not differ from the analyses employing Path type, we do not repeat these significant main effects here. We limit our presentation to new effects and interactions that surfaced as a result of including Relation type.

The main effect of Relation type was significant, $F(2, 176) = 14.69$, $p < .001$, $\eta_p^2 = .143$. In addition, there was a two-way interaction of Suffix meaning × Relation type, $F(2, 176) = 41.83$, $p < .001$, $\eta_p^2 = .322$, as well as a three-way interaction of Suffix meaning × Relation type × Group, $F(2, 176) = 9.86$, $p < .001$, $\eta_p^2 = .101$.

The Suffix meaning × Relation type × Group interaction was examined further by testing the effects on the SLI and VC groups separately. For the VC children, the effect of Relation type was significant both when the suffixes conveyed a Spatial meaning, $F(2, 90) = 20.24$, $p < .001$, $\eta_p^2 = .310$, and a Nonspatial meaning, $F(2, 90) = 4.41$, $p < .05$, $\eta_p^2 = .089$. Similar results obtained for the SLI group. The Relation type effect was seen for both Spatial, $F(2, 90) = 30.88$, $p < .001$, $\eta_p^2 = .407$, and Nonspatial meanings, $F(2, 90) = 4.95$, $p < .01$, $\eta_p^2 = .099$.

Given the effects for Relation type, we employed pairwise comparisons to determine how the three Relation types might have differed. For VC children, when the suffixes had a Spatial meaning, Surface and Container relations were significantly easier than Neighborhood relations ($p < .001$ in both cases). When the suffixes had a Nonspatial meaning, Surface relations were easier than Container relations ($p < .01$) and Neighborhood relations ($p < .05$) but the latter two did not differ. For children with SLI, when the suffixes conveyed a Spatial meaning, the pattern was similar to that of VC children: Surface and Container relations were expressed more accurately than Neighborhood relations ($p < .001$ in both cases), while for Nonspatial meanings, Surface and Container relations did not differ, but both were expressed less accurately than Neighborhood relations ($p < .01$ in both cases).

Summary of results of analysis of correct responses

The VC group showed a significantly larger overall proportion of correct responses than the children with SLI. Older children showed a better performance than younger children within both SLI and VC groups. Sentences with suffixes in their spatial meanings were easier for both SLI and VC children, older and younger, with a bigger advantage for spatial items in the younger SLI than the younger VC group. Performance along Path type differed significantly, but it showed largely the same pattern across the SLI and VC groups, older and younger: Goal suffixes were easiest, followed by Source suffixes, while children found sentences with Static suffixes the most difficult. Path type affected Spatial and Nonspatial meanings differentially: the effect was stronger for Nonspatial suffixes than for Spatial suffixes, and the pattern was also different. For Spatial suffixes, Goal was easiest, while Static and Source did not differ. For Nonspatial meanings, younger children found Goal suffixes easier than Static suffixes which they found easier than Source suffixes. For older children, there was no difference between Goal and Source suffixes, which they found easier than Static suffixes.

The effect of Relation type was stronger in younger than in older children, but did not differ between VC and SLI groups. Overall, Surface and Container relations did not differ, but performance on both was significantly better than on suffixes coding Neighborhood relations. Contrary to the path effect pattern, Relation type had a stronger effect on Spatial than on Nonspatial suffixes, where there was also a difference between groups, with children with SLI showing a stronger effect. For Spatial meanings, Surface and Container relations did not differ, and they were both easier than Neighborhood relations for both the VC and SLI groups. With Nonspatial meanings, VC children found Surface relations easier than both Container and Neighborhood relations, which did not differ. Children with SLI found Surface and Container relations equally more difficult than Neighborhood relations.

Error analysis

The great majority of errors were commission errors, i.e., children tended to use another suffix in place of the target form instead of omitting the suffix or producing other answers. This was true for all groups except for the older SLI group, where only a little more than half of the responses were substitutions with another suffix. For more detailed analysis of error patterns, we only analyzed commission errors further; results are shown in Table 7.

As Table 6 shows, children in all groups mostly substituted target suffixes with another one of the nine suffixes under investigation,[5] and the majority of these substitutions were near-miss errors sharing either Path type or Relation type with the target suffix. From within the two-dimensional (Path type × Relation type) matrix of the nine spatial suffixes, for each target suffix, there are four that are errors on only one dimension (two agreeing in Path type but differing in Relation type, and two agreeing in Relation type and differing in Path type), and four that differ from the target on both dimensions, so the chance of producing an error on only one dimension by selecting one of the nine suffixes at random is 50% (4/8), which is greatly exceeded by the actual percentages in all four groups. As the distribution of same Path type versus same Relation type errors within errors on only one dimension shows (last two rows of Table 6), the majority of responses showed the correct Relation type but incorrect Path type.

Table 7 shows the distribution of errors between Spatial and Nonspatial targets (adding up to 100% in every case). While the difference between the ratio of substitution errors for Spatial and Nonspatial target suffixes did not differ greatly, there were some more specific differences. Suffix substitutions with errors on both dimensions were more

Table 6. Distribution of different error types as a function of language ability and age.

	VC young	SLI young	VC old	SLI old
Errors (%)	26.76	41.07	19.25	29.7
%Commission errors/all errors	88.27	72.79	82.98	52.12
%Commission errors using one of the 9 suffixes under study/commission errors	70.12	85.90	78.77	77.77
%Errors on only one dimension/using one of the 9 suffixes under study	72.19	75.57	76.71	78.57
%Same Relation type/Errors only on Path type	63.47	59.47	71.95	76.82
%Same Path type/Errors only on Relation type	36.52	40.52	28.04	23.17

VC: vocabulary control, SLI: specific language impairment.

Table 7. The distribution of errors between spatial and nonspatial targets (percentages).

		VC young	SLI young	VC old	SLI old
Commission errors using one of the 9 suffixes under study	Spatial	56.88	46.91	51.94	46.07
	Nonspatial	43.13	53.09	48.06	53.93
Commission errors using one of the 9 suffixes under study, error in one dimension	Spatial	66.09	54.74	56.71	52.32
	Nonspatial	33.91	45.26	43.29	47.68
Error only on Path type	Spatial	75.34	63.72	59.32	56.90
	Nonspatial	24.66	36.28	4.68	43.10
Error only on Relation type	Spatial	50.00	41.56	5.00	37.14
	Nonspatial	50.00	58.44	5.00	62.86
Error on both dimensions	Spatial	28.57	24.56	33.33	18.42
	Nonspatial	71.43	75.44	66.67	81.58
Accusative	Spatial	30.23	30.43	31.58	26.32
	Nonspatial	69.77	69.57	68.42	73.68

VC: vocabulary control, SLI: specific language impairment.

common with Nonspatial targets, as was replacing the target suffix with the accusative. Errors sharing the Relation type with the target suffix (thus, constituting an error of Path type) were more frequent with Spatial targets.

General discussion

The purpose of this investigation was to determine whether Hungarian-speaking children with SLI have difficulty with non-structural case marking, and, if so, whether the difficulty is comparable for Spatial and Nonspatial meanings conveyed by these markers. In Study 1, we used the Frog Story task (Mayer, 1969) to elicit narratives from the children. Analysis of narratives showed that children with SLI have problems with case marking, although this deficit was not manifest in the number of case marking errors (relative to either chronological age-matched or vocabulary controls). Differences were seen in the total number of case-marked nouns and the number of different case marker suffix types, but only relative to age-equivalent peers, not to younger children with the same level of receptive vocabulary. Their accuracy levels were comparable to both age and vocabulary controls, all producing very few case marking errors. This pattern of results suggests that although case marking is problematic for Hungarian children with language impairment relative to age-matched peers, case marking performance in spontaneous language production is largely influenced by lexical abilities and that the acquisition of the case marking system of Hungarian does not pose *selective* difficulties for children with SLI.

In Study 2, we employed an elicited production task. VC groups matched on vocabulary size showed a significantly larger overall proportion of correct responses than children with SLI. Older children showed a better performance than younger children within both SLI and VC groups. Sentences with suffixes in their Spatial meanings were easier for both SLI and VC children, older and younger, with a bigger advantage for Spatial items in the older SLI than older VC group. Performance along Path type differed significantly, but showed largely the same pattern across the SLI and VC groups, older and younger: Goal suffixes were easiest, followed by Source suffixes, while children found sentences with Static suffixes the most difficult. Path type affected Spatial and Nonspatial meanings differentially: the effect was stronger for Nonspatial suffixes than for Spatial suffixes, and the pattern was also different. For Spatial suffixes, Goal was easiest, while Static and Source did not differ. For Nonspatial meanings, older children found Goal suffixes easier than Static suffixes which they found easier than Source suffixes. For younger children, there was no difference between Goal and Source suffixes, which they found easier than Static suffixes.

The effect of Relation type was stronger in older than in younger children, but did not differ between VC and SLI groups. Overall, Surface and Container relations did not differ, but performance on both was significantly better than on suffixes coding Neighborhood relations. Contrary to the Path type pattern, Relation type had a stronger effect on Spatial than on Nonspatial suffixes, where there was also a difference between groups, with children with SLI showing a stronger effect. For Spatial meanings, Surface and Container relations did not differ, and they were both easier than Neighborhood relations for both the VC and SLI groups. With Nonspatial meanings, VC children found Surface relations easier than both Container and Neighborhood relations, which did not differ. Children with SLI found Surface and Container relations equally more difficult than Neighborhood relations.

Performance of the SLI group was significantly weaker on Nonspatial items. This pattern suggests a level of language development in SLI where many of the verbs tested here, together with their argument structure containing the suffix, are missing from the lexicon. The dissociation between Spatial and Nonspatial items was smaller in the VC group Also, we found a bigger group difference for Nonspatial meanings, i.e., for suffix occurrences that have to be lexically learnt one by one for each verb. Accurate performance on Nonspatial items requires the acquisition of argument structures or complex lexical representations of verbs; it seems that it is this aspect of lexical acquisition that is especially difficult in SLI.

Results of the error analysis showed that when children were making an error, they did not resort to using some kind of default case such as the accusative, or to using the most frequent cases, showing that they encoded some part of the suffix representation. For most children, the majority of the errors were substitutions with another of the nine suffixes tested in the experiment. For Spatial meanings, errors usually shared either Relation type or Path type with the target suffix. Errors were mostly errors of Path type; this is the information that is encoded in the argument structure of the verb. Spatial relations, determined by the noun, were easier to select. For Spatial meanings, errors on both dimensions (Relation type and Path type) were a minority; this reflects the fact that Spatial meanings semantically code information along two dimensions, combining Relation type and Path type, determined by the noun and the verb, respectively. It is possible that when processing resources are not fully available, only one of these dimensions (mostly the easier one determined by the noun) gets encoded. As suffixes in their Nonspatial meanings are not transparent, they do not combine these two kinds of information from two sources; rather, they represent labels determined only by the verb. For this reason, systematic errors such as selecting only one of the dimensions would not be expected.

Taken together, results from narratives and an elicited production task suggest that case marking performance is more a function of vocabulary size than of grammatical knowledge, and follows the same pattern as in typical development. It would seem, therefore, that similar semantic and pragmatic influences determine case suffix acquisition in typically developing children and children with SLI. These findings argue that there is no selective deficit of case marking per se in Hungarian-speaking children with SLI. As the difficulties can be explained by lack of lexical knowledge, or by semantic complexity and transparency of the suffixes, there is no need to posit a selective case marking deficit. Indeed, the specific pattern with better performance on systematic Spatial than on idiosyncratic Nonspatial items would be difficult to interpret on a selective difficulty account.

As discussed in the introduction, most accounts of grammatical deficits in SLI do not address the issue of lexical and inherent case use. The most recent modification of the agreement deficit account of Clahsen and his colleagues (e.g., Eisenbeiss et al., 2005) holds that the deficit lies in subject–verb agreement relations. The Rice and Wexler (1996) account focuses on inconsistent projection of the functional category AGR that prevents the licensing of nominative case in a language such as English. The latest version of the account by van der Lely and colleagues (e.g., Marshall & van der Lely, 2007) – that complexity of structure is the source of difficulty – has thus far not been applied to the use of lexical or inherent case by children with SLI.

However, one processing account of the grammatical deficits of SLI – the morphological richness account (e.g., Leonard, 1998, pp. 255–257) – does provide a basis for predicting the major findings reported here. In particular, this account predicts less difficulty with morphology (and thus with case marking) in a morphologically rich language such as Hungarian. Further, although we observed poorer performance by the children with SLI on the elicited production task, performance patterns matched those of typical development with regard to the Spatial–Nonspatial distinction, and the relative difficulty of Relation type and Path type. The finding that the children with SLI found Nonspatial meanings especially difficult relative to typically developing children further supports predictions of this account, as the most vulnerable suffixes are expected to be those that pose a processing difficulty. Suffixes in their Nonspatial meanings are assumed to be difficult to process and acquire because they are semantically non-transparent, not systematically associated with a thematic role, and vary greatly in their frequency of use. The finding that most errors, especially with Spatial meanings, were errors in which the children could rely on the spatial relation information provided by the noun (thereby getting only the Path type wrong), also suggests difficulties with lexically more complex representations of verbs or integrating lexical information from verbs in the sentence.

As a final remark, we would like to add that discovering deficits that seem attributable to processing difficulty does not imply that diagnostically accurate clinical markers of SLI cannot emerge from these efforts. On the contrary, the discovery of areas of grammar that constitute the greatest processing challenges for children with SLI may prove to be an especially good way of identifying language impairment. Furthermore, the advantage of an approach such as the morphological richness account is that the areas of weakness can be seen through an interaction of a processing capacity limitation and the properties of the particular language being acquired. Such an approach possesses the means of explaining why the potential markers of SLI seem to differ so widely across different languages. We strongly suspect that the potential clinical marker observed in the present investigation – difficulty with case suffixes in their Nonspatial meanings – represents one such example of a problem that emerged as a consequence of processing limitations coupled with special complexity within one part of Hungarian grammar.

Acknowledgements

The authors are grateful to the children in the Dr Nagy László Institute of Special Education in Kőszeg, in the ELTE Special Preschool and Early Intervention Centre, and in the Zölderdő Preschool for Speech Therapy and Nature Preservation for their participation. We also thank the speech therapists in all institutions for their help with screening and organization.

Funding

This research was supported by research grant R01 DC00458 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (USA) to Laurence B. Leonard and by OTKA TS 049840 to Csaba Pléh and OTKA K 83619 to Ágnes Lukács from the Hungarian National Science Foundation. Ágnes Lukács was a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science.

Notes

1. Following current practice, we use the term 'specific language impairment'. However, we recognize, as do other scholars, that these children with SLI often show subtle weaknesses in nonlinguistic processing tasks and, as a group, may be somewhat slow in their motor development. Nevertheless, language constitutes their most conspicuous difficulty.
2. Further details of Hungarian morphophonology are not discussed here; the interested reader is referred to Siptár and Törkenczy (2000) for details. Also, we do not discuss the agglutinative versus fusional elements of the paradigms, as we do not focus on suffix combinations, and the target suffixes in our study are homogeneous in this regard.
3. We are grateful to Professor Dorothy Bishop for providing us with the TROG for this purpose. the time of testing, 600 typically developing children had been tested as part of the standardization process. The scores of the children with SLI in this study were compared against the

values obtained for the TD children participating in the standardization. TROG has a discontinue procedure after five blocks are failed, but this discontinue rule was not followed either in the standardization or in testing children in this study.
4. In our first study using this method (Lukács et al., 2009) we wanted to ensure that the inserted coughs were sufficient to obscure the inflection, and that there were no anticipatory coarticulatory cues on the stem to provide the children with an indication of the inflection that was masked. We asked 15 adult listeners to guess which inflection was used with the stem for all the verb forms with coughs. They guessed correctly on 5.6% of the items, significantly below performance of either group of children on the task. We take these findings as indication enough that the method in general ensures that the stimuli do not contain unintended cues that could lead to correct performance without knowing the appropriate inflection.
5. Note that here the term 'spatial suffix' does not specify whether it was used in its spatial or nonspatial meaning; rather, it is shorthand for the nine cases examined in the study.

References

Bartos, H. (2000). Az inflexiós jelenségek szintaktikai háttere. In F. Kiefer (Ed.), *Strukturális magyar nyelvtan 3. – Morfológia* (pp. 653–762). Budapest, Hungary: Akadémiai.

Bedore, L., & Leonard, L. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research*, *41*, 1185–1192.

Bedore, L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 905–924.

Bishop, D. V. M. (1983). *Test for the reception of grammar*. Manchester, UK: Medical Research Council.

Bishop, D.V.M. (2012). TROG - Test for Reception of Grammar. Handbook. OS Hungary Tesztfejlesztő Kft. 47–86.

Bortolini, U., Caselli, M. C., & Leonard, L. (1997). Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *40*, 809–882.

Çavuş, N. (2009). *In search of clinical markers in Turkish SLI; an analysis of the nominal morphology in Turkish bilingual children* (Master's thesis). General Linguistics, Faculty of Humanities, University of Amsterdam, the Netherlands.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures*. Holland, the Netherlands: Foris Publications.

Clahsen, H. (1991). *Child language and developmental dysphasia*. Amsterdam, the Netherlands: John Benjamins.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics*, *10*, 151–171.

Clahsen, H., Eisenbeiss, S., & Vainikka, A. (1994). The seeds of structure: A syntactic analysis of the acquisition of case marking. In T. Hoekstra, & B. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp. 85–118). Amsterdam, the Netherlands: John Benjamins.

Csányi, F. I. (1974). *Peabody szókincs-teszt*. Budapest, Hungary: Bárczi Gusztáv Gyógypedagógiai Főiskola.

Dankovics, N., & Pléh, C. (2001). Hangrestaurációs jelenségek és alaktani feldolgozás a magyarban: Azt halljuk-e, amit várunk? In C. Pléh, & Á. Lukács (Eds.), *A magyar morfológia pszicholingvisztikája* (pp. 55–83). Budapest, Hungary: Osiris.

de Jong, J., Çavuş, N., & Baker, A. (2010). Language impairment in Turkish-Dutch bilingual children. In S. Topbaş, & M. Yavaş (Eds.), *Communication disorders in Turkish in monolingual and multilingual settings* (pp. 290–303). Clevedon, UK: Multilingual Matters.

Dromi, E., Leonard, L., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research*, *42*, 1414–1431.

Dunn, L., & Dunn, L. (1981). *Peabody picture vocabulary test – Revised*. Circle Pines, MN: American Guidance Service.

Eisenbeiss, S., Bartke, S., & Clahsen, H. (2005). Structural and lexical case in child German: Evidence from language-impaired and typically-developing children. *Language Acquisition*, *13*(1), 3–32.

Fletcher, P., Leonard, L., Stokes, S., & Wong, A. M.-Y. (2005). The expression of aspect in Cantonese-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*, 621–634.

Friederici, A. D. (1982). Syntactic and semantic processes in aphasic deficits: The availability of prepositions. *Brain and Language*, *15*, 249–258.

Gábor, B., & Lukács, Á. (2012). Early morphological productivity in Hungarian: Evidence from sentence repetition and elicited production. *Journal of Child Language*, *39*, 411–442.

Kas, B., & Lukács, Á. (in press). *Magyar Mondatutánmondási Teszt*.

Kiefer, F. (2000). *A ragozás* [Inflectional morphology]. In K. Ferenc (Ed.), *Strukturális magyar nyelvtan. 3. Alaktan* (pp. 569–618). Budapest, Hungary: Akadémiai Kiadó.

Leonard, L. (1982). The nature of specific language impairment in children. In S. Rosenberg (Ed.), *Handbook of applied psycholinguistics* (pp. 295–327). Hillsdale, NJ: Lawrence Erlbaum.

Leonard, L. (1995). Functional categories in the grammars of children with specific language impairment. *Journal of Speech and Hearing Research*, *38*, 1270–1283.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L., Caselli, M. C., & Devescovi, A. (2002). Italian children's use of verb and noun morphology during the preschool years. *First Language*, *22*, 287–304.

Leonard, L., & Dromi, E. (1994). The use of Hebrew verb morphology by children with specific language impairment and children developing language normally. *First Language*, *14*, 283–304.

Leonard, L., Sabbadini, L., Leonard, J., & Volterra, V. (1987). Specific language impairment in children: A crosslinguistic study. *Brain and Language*, *32*, 233–252.

Loeb, D., & Leonard, L. (1991). Subject case marking and verb morphology in normally-developing and specifically-language-impaired children. *Journal of Speech and Hearing Research*, *34*, 340–346.

Loeb, D., & Leonard, L. B. (1988). Specific language impairment and parameter theory. *Clinical Linguistics & Phonetics*, *2*, 317–327.

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

Lukács, Á., Leonard, L. B., & Kas, B. (2010). The use of noun morphology by children with language impairment: The case of Hungarian. *International Journal of Language & Communication Disorders*, *45*, 145–161.

Lukács, Á., Pléh, C., & Racsmány, M. (2007). Spatial language in Williams syndrome: Evidence for a special interaction? *Journal of Child Language*, *34*, 311–343.

Lukács, Á., Győri, M., & Rózsa, S. (2012). A TROG pszichometriai jellemzőinek magyar vizsgálata, a normák kialakítása [Psychometric characteristics of TROG in Hungarian, norms]. In.: Bishop, D.V.M. (2012). TROG - Test for Reception of Grammar. Handbook. OS Hungary Tesztfejlesztő Kft. 47–86.

Marshall, C. R., & van der Lely, H. K. J. (2007). The impact of phonological complexity on past tense inflection in children with grammatical-SLI. *Advances in Speech Language Pathology*, *9*, 191–203.

Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Book for Young Readers.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle*, *4*, 479–505.

Radford, A., & Ramos, E. (2001). Case, agreement and EPP: Evidence from an English-speaking child with SLI. *Essex Research Reports in Linguistics*, *36*, 42–81.

Raven, J., Court, J., & Raven, J. (1987). *Raven's progressive matrices and Raven's colored matrices*. London, UK: H.K. Lewis.

Rice, M., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, *39*, 1239–1257.

Rice, M., & Wexler, K. (2001). *Rice/Wexler test of early grammatical impairment*. San Antonio, TX: Psychological Corporation.

Roberts, S., & Leonard, L. (1997). Grammatical deficits in German and English: A crosslinguistic study of children with specific language impairment. *First Language*, *17*, 131–150.

Rothweiler, M., Chilla, S., & Babur, E. (2010). Specific language impairment in Turkish: Evidence from case morphology in Turkish-German successive bilinguals. *Clinical Linguistics & Phonetics*, *24*, 540–555.

Rothweiler, M., & Clahsen, H. (1994). Dissociations in SLI children's inflectional systems: A study of participle inflection and subject-verb-agreement. *Journal of Logopedics and Phoniatrics*, *18*, 169–179.

Schütze, C. (1997). *INFL in child and adult language: Agreement, case and licensing* (PhD dissertation). Cambridge, MA: MIT (Distributed by MITWPL).

Siptár, P., & Törkenczy, M. (2000). *The phonology of Hungarian*. Oxford, UK: Oxford University Press.

Stavrakaki, S., & van der Lely, H. (2010). Production and comprehension of pronouns by Greek children with specific language impairment. *British Journal of Developmental Psychology*, *28*, 189–216.

van der Lely, H. K. J. (1994). Canonical linking rules: Forward versus reverse linking in normally developing and specifically language impaired children. *Cognition*, *51*, 29–72.

van der Lely, H., & Stollwerck, L. (1997). Binding theory and specifically language impaired children. *Cognition*, *62*, 245–290.

Vinkler, Z., & Pléh, C. (1995). A case of a specific language impaired child in Hungarian. In M. Kovačević (Ed.), *Language and language communication barriers* (pp. 131–158). Zagreb, Republic of Croatia: Hrvatska Sveucilisna Naklada.

Warren, R. M. (1970). Perceptual restorations of missing speech sounds. *Science*, *167*, 392–393.
Wexler, K., Schütze, C., & Rice, M. (1998). Subject case in children with SLI and unaffected controls: Evidence for the Agr/Tns omission model. *Language Acquisition*, *7*, 317–344.
Woolford, E. (2006). Lexical case, inherent case, and argument structure. *Linguistic Theory*, *37*, 111–113.

Study 7.

Zakariás, L., Keresztes, A., Demeter, Gy. & Lukács, Á. (2013). A specific pattern of executive dysfunctions in transcortical motor aphasia. *Aphasiology*, 27:12, 1426-1439, doi: 10.1080/02687038.2013.835783

Thesis 11. In aphasia, the acquired language impairment is not specific to language: it is often accompanied by the impairment of nonverbal executive functions (7)

Aphasiology, 2013
Vol. 27, No. 12, 1426–1439, http://dx.doi.org/10.1080/02687038.2013.835783

A specific pattern of executive dysfunctions in transcortical motor aphasia

Lilla Zakariás[1,2], Attila Keresztes[3], Gyula Demeter[3], and Ágnes Lukács[3]

[1]Faculty of Education and Psychology, Eötvös Loránd University, Budapest, Hungary
[2]Neurology and Rehabilitation Centre, Flór Ferenc Hospital, Kistarcsa, Hungary
[3]Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary

Background: Recent studies imply that executive functions (EF) are closely related to our ability to comprehend and produce language. A number of findings suggest that functional communication and language recovery in aphasia depend not only on intact language abilities but on EF as well. Some patients with transcortical motor aphasia (TMA) show language deficits only in tasks in which conflicting representations must be resolved by executive processes. In line with these results, others have proposed that TMA should be referred to as "dysexecutive aphasia". EF in aphasia have mostly been studied using neuropsychological tests, therefore there is a need for systematic experimental investigations of these skills.
Aims: 1. To investigate EF in TMA, and to test whether executive dysfunctions are specific to TMA. 2. To experimentally measure different components of EF: updating working memory representations and inhibition of prepotent responses.
Methods & Procedures: Five individuals with TMA, five patients with conduction aphasia and ten healthy controls participated. We designed four nonverbal tasks: to measure updating of working memory representations, we used a visual and an auditory *n*-back task. To assess inhibition of prepotent responses, we designed a Stop-signal and a nonverbal Stroop task. All tasks involved within-subject baseline conditions.
Outcomes & Results: We found certain EF deficits in both groups of individuals with aphasia as compared to healthy controls. Individuals with TMA showed impaired inhibition as indexed by the Stop-signal and the nonverbal Stroop tasks, as well as a deficit of updating of working memory representations as indexed by the auditory *n*-back task. Participants with conduction aphasia had difficulties in only one of the tasks measuring inhibition, but no clear evidence for impairment of updating of working memory representations was found.
Conclusions: Although the results show different patterns of EF deficits in the groups with aphasia, the findings clearly demonstrate that EF deficits are not specific to participants with TMA. Based on these results, and on earlier data highlighting the role of executive

Address correspondence to: Lilla Zakariás, Faculty of Education and Psychology, Eötvös Loránd University, Budapest 1136, Hungary. E-mail: lilla.zakarias@gmail.com

We thank Klára Marton for discussions on the study and her comments on a previous version of the manuscript, and László Jenei for his help in analysing neurological data.

processes in functional communication and language recovery, we suggest that tests of EF should be an inherent part of clinical aphasia assessment.

Keywords: Executive functions; Cognitive control; Working memory; Conduction aphasia; *n*-back; Aphasia assessment.

Recent studies imply that executive functions (EF) are closely related to our ability to comprehend and produce language (Novick, Trueswell, & Thompson-Schill, 2010). EF are generally defined as a range of cognitive processes that enable us to control and regulate various cognitive processes and thereby behaviour (e.g., Miyake et al., 2000). These functions do not add up to a unitary construct, and are always considered as a set of functions or components, such as shifting between tasks or mental sets, updating and monitoring of working memory representations and inhibition of prepotent responses (Miyake et al., 2000).

More and more studies investigating aphasia emphasise the role of EF in successful communication, particularly in conversation (Alexander, 2006; Frankel, Penn, & Ormond-Brown, 2007; Green et al., 2010; Penn, Frankel, Watermeyer, & Russell, 2010; Purdy, 2002; Ramsberger, 2005). These findings suggest that conversational success depends not only on language ability but on EF as well. Based on the model of Barkley (1997), Penn et al. (2010) suggested separate roles for inhibition and working memory in discourse features: these EF components seem to be important in maintaining focus, initiating new topics, planning and monitoring our communicative performance, including shifting between communication strategies to successfully convey information (Ramsberger, 2005) or effectively generating self-repair to error correction (Penn et al., 2010).

Others have suggested that EF also play a role in recovery from aphasia (Green et al., 2010; Penn et al., 2010; Ramsberger, 2005). For instance, differences in executive abilities may account for different language recovery patterns in bilingual aphasia (Green et al., 2010). Moreover, it seems that executive deficits also have an influence on therapy outcome, because the ability to generate, select and apply strategies is important in utilising trained methods (Keil & Kaszniak, 2002).

In line with these results, some studies have investigated the influence of cognitive therapy, particularly EF training, on language improvement (Hardin & Ramsberger, 2004, cited by Helm-Estabrooks & Albert, 1991; Helm-Estabrooks, Connor, & Albert, 2000; Ramsberger, 2005). After delivering attention-training programme, Helm-Estabrooks et al. (2000) revealed improved performance on tasks measuring auditory comprehension and visual analytic reasoning. Similarly, Hardin and Ramsberger (2004) noted that attention/executive training can lead to improvement of transactional success in conversation.

In brief, understanding executive processes is relevant to understanding aphasia (e.g., Code, Tree, & Dawe, 2009; Green et al., 2010), and although EF and their relationship to certain language symptoms have already been investigated in aphasia (e.g., Alexander, 2006; Penn et al., 2010), research specifically addressing the relationship between executive deficits and different types of aphasia is scarce (Keil & Kaszniak, 2002).

TMA: A special case of executive dysfunctions?

According to theories linking EF with language symptoms, transcortical motor aphasia (TMA) seems to be of outstanding relevance (Alexander, 2006; Ardila, 2010;

Luria, 1973; Novick, Kan, Trueswell, & Thompson-Schill, 2009; Robinson, Blair, & Cipolotti, 1998). In classical terminology, TMA, one of the eight aphasia syndromes, is characterised by nonfluent output, anomia, good auditory comprehension, and relative to spontaneous speech, almost preserved repetition. However, the characteristics of TMA, and associated lesion sites can vary greatly. Based on such variations, some authors suggested that that there are different forms of TMA (e.g., the distinction between TMA and dynamic aphasia, a type of aphasia first described by Luria, 1973). The nature of symptoms and the overlap between brain regions affected in TMA and those associated with EF have led some researchers to propose that language symptoms in TMA arise due to executive dysfunctions (Alexander, 2006; Luria, 1973).

TMA can involve many different brain regions overlapping to a large extent with regions associated with EF. Reviewing clinical–anatomical studies, Alexander (2003) pointed out that patients diagnosed with TMA had diverse lesions in many different areas of the left frontal lobe and in structures deep into them. The most common lesion sites were in the dorsolateral frontal cortex (BA 45, 46, 9), typically extending into the deep white matter, the ventrolateral (BA 44, 45, 47) and medial frontal lobe (BA 24, 32), including the supplementary motor area (BA 6, 32). Given the large overlap of these areas with neural networks involved in EF (for localisation of EF, see Botvinick, Cohen, & Carter, 2004; Milham et al., 2001; Miyake et al., 2000; Smith, Jonides, Marshuetz, & Koeppe, 1998), Alexander (2006) posited two different executive processes that subserve complex language use. According to his account, the left medial frontal lobe is critical for the activation of language responses, whereas the left lateral frontal areas are necessary to exert control (e.g., inhibition, suppressing, sustaining and monitoring) over procedures implementing syntax and narrative discourse. Disruption of these control processes will, depending on the site and the extension of the lesion, manifest in language use in aphasia to different degrees. More specifically, the above-mentioned EF disturbances might lead to the impairment of complex syntax implementation, lexical selection and difficulties in narrative discourse in TMA. In the framework proposed by Alexander (2006), the level of control procedure impairment, partly associated with lesions in different frontal loci, determines the level of language impairment, leading to hierarchically organised types of aphasia related to TMA.

In a similar vein, Ardila (2010) proposed that TMA is not a primary aphasic syndrome in terms of the underlying impairment. Rather, TMA is "an executive function defect specifically affecting language use" (Ardila, 2010, p. 374–375). He argued that TMA patients' primary language skills are intact, but demonstrate the characteristics of dysexecutive syndrome specifically with regard to verbal processes. Hence, he also has proposed that TMA should be referred to as "dysexecutive aphasia".

Convergent evidence for TMA language symptoms as manifestations of an executive deficit is provided by a few case studies. Some authors, investigating the purest form of TMA, dynamic aphasia, reported patients whose language deficits arose only under certain conditions (Luria, 1973; Novick et al., 2009; Robinson et al., 1998). Following injury to the left inferior frontal gyrus, these patients presented a conspicuous lack of verbal fluency only on tasks in which conflicting representations had to be resolved by executive processes. They have been characterised by an almost complete lack of spontaneous speech in contrast to well-preserved naming, comprehension, repetition and reading skills.

Robinson et al. (1998) described the case of A. N. G., who presented extremely reduced speech during conversation but had no difficulty in a confrontation-naming task or when she had to generate sentences from a pictorial scene. Moreover, the authors revealed that A. N. G. had difficulties on the structurally analogous tasks depending on how many verbal responses could be activated. For example, in a sentence-completion task, she appropriately completed almost all the sentences where only few continuations were possible, whereas her performance was significantly impaired when trying to complete sentences with more response options. The authors concluded that difficulties of patients with dynamic aphasia can be explained by the inability to select from competing response options. According to Robinson and colleagues, this might also explain the nonfluency of A. N. G.'s spontaneous speech.

Novick et al. (2009) also investigated a single patient's, I.G.'s, performance on several conflict-resolution tasks. These included a proactive-interference task using letter stimuli, a picture-naming task using stimuli of varying name agreement (low-agreement stimuli depicting objects with multiple names, e.g., couch, sofa, loveseat vs. high-agreement stimuli depicting objects with a unique name, e.g., apple), a verbal-fluency task and a comprehension task involving syntactic ambiguity. Similar to Robinson et al. (1998), Novick and colleagues concluded that "I. G. had a general conflict resolution impairment which affects his ability to produce and comprehend language under specific conditions" (Novick et al., 2009, p. 528).

Measuring EF in TMA

Despite the growing interest in the relationship between EF and aphasia, and in particular TMA, "pure" nonverbal executive skills have not yet been systematically investigated in this type of aphasia. Following a review of studies focusing on EF in aphasia Keil and Kaszniak (2002) concluded that performance on most of the widely used EF tests require language-related processes, which poses serious limitations on their use in populations with aphasia. In addition, they suggested that tests meant to measure EF in individuals with aphasia should mitigate psychomotor slowing and avoid motor processing speed confound (Keil & Kaszniak, 2002).

As a further step in the understanding of the exact nature of executive processes in TMA, our study aimed to assess two different components of EF which are crucial for language abilities like lexical selection, successful conversation and narrative discourse, in a group of individuals with TMA. Based on the framework of Penn et al. (2010) we focused on updating and monitoring of working memory representations and inhibition of dominant responses. Working memory processes have been proposed to support shifting, maintaining topics during conversation, integrating new information with current communicative content and organising communicative behaviour across time (Frankel et al., 2007; Penn et al., 2010). Inhibition of dominant responses, according to this framework involves two different types of inhibition processes that are involved in different aspects of language processing and production. The ability to stop a prepotent response is proposed to be necessary to recognise and to stop ineffective communicative strategies and to shift to an effective one. Inhibition-based interference control, on the other hand, would make it possible to sustain the topic of a conversation, and the communicative goal in the face of distractors, and inhibiting irrelevant information. This type of inhibition is also important for selecting appropriate lexical and syntactic representations in cases of competition (e.g., Novick, Trueswell, & Thompson-Schill, 2005; Schnur, Lee, Coslett, Schwartz, & Thompson-Schill, 2005).

Similar distinctions have been made by Novick et al. (2005) who suggested different inhibitory processes for the resolution of response-based and representational conflict (i.e., response inhibition versus inhibition-based interference control in the Penn et al., 2010 framework). Recently, it has been suggested that two classical inhibitory paradigms, the Stop-signal task and the Stroop task not only differ in their complexity but also in the type of conflict that has to be resolved during performing them. Whereas the Stop-signal task is supposed to tap the resolution of response-based conflict, the Stroop task is more likely to tap resolution of representational conflict. In line with these suggestions, successful performance on the Stop-signal task have been correlated with activations in medial frontal areas, whereas success in the Stroop task have been shown to correlate with ventrolateral frontal activity (Milham et al., 2001; Novick et al., 2005). Therefore, in assessing inhibitory functions in aphasia, we used both the Stroop and the Stop-signal paradigm.

To see whether any pattern of impairment found is specific to TMA, we also included a group of patients with conduction aphasia as controls. Two major reasons motivated our choice of a group with conduction aphasia. First, we intended to include a control group with different neural networks underlying symptoms and possible cognitive dysfunctions, but with similar level of auditory comprehension necessary to perform the experimental tasks. Second, although working memory, and in particular verbal working memory deficits have been reported in both types of aphasia, the way these deficits manifest in language seems to be different. In addition, in conduction aphasia, these deficits have been primarily related to an impaired storage capacity of the phonological loop (e.g., Friedmann & Gvion, 2003; Gvion & Friedmann, 2012), disturbances in TMA have been associated also to impaired manipulation of representations stored in working memory. Taken together, we expected to find different patterns of performance on tasks measuring EF.

We designed four nonverbal tasks. In order to avoid confounds summarised by Keil and Kaszniak (2002), all tasks involved within-subject baseline conditions. The four tasks were variations of widely used EF tasks: a nonverbal Stroop and a Stop-signal task measuring inhibition of prepotent responses (Logan, 1994; Milham et al., 2001; Novick et al., 2005; Stroop, 1935) and two variations (one in the auditory and one in the visual modality) of the *n*-back task to measure updating of working memory representations (Miyake et al., 2000). Based on earlier findings we expected to detect a specific pattern of executive deficits in TMA that would be clearly different from that observed among healthy controls and in conduction aphasia.

METHODS

Participants

A total of five individuals with TMA (age: M = 58 years, SD = 13.60; 1 female, 4 males; education: M = 12.6, SD = 2.6; $\text{RAVEN}_{\text{age corrected}}$: M = 48.92, SD = 13.81) participated. As controls, a group with conduction aphasia (n = 5; age: M = 53 years, SD = 4.84; 1 female, 4 males; education: M = 11.2, SD = 2.05; $\text{RAVEN}_{\text{age corrected}}$: M = 33.60, SD = 7.56) and a group of healthy participants (n = 10; age: M = 59.5 years, SD = 12.26; 6 female, 4 males; education: M = 12.9, SD = 2.96; $\text{RAVEN}_{\text{age corrected}}$: M = 53.77, SD = 11.69) were recruited. Healthy controls were matched in age and education. All participants with aphasia have had a single left hemisphere infarct, confirmed by CT or structural MRI, except one, who has had a traumatic injury, also to the left hemisphere. The mean time post-onset was 8.4 months for the group with TMA and 8.6 months for the group with conduction aphasia. All of them spoke Hungarian as their primary language and were right-handed. They had been recruited and tested at two rehabilitation centres in Budapest, Hungary:

the Flór Ferenc Hospital and the National Institute for Medical Rehabilitation. Their language impairment was classified by the Western Aphasia Battery (Kertész, 1982; Hungarian adaptation: Osmánné Sági, 1991) complemented with the Boston Naming Test (Kaplan, Goodglass, Weintraub, & Segal, 1983). Neurological assessment showed no visual problems for any of the patients, and all of them reported hearing within normal limits. Kruskal–Wallis tests showed that there were no significant differences between the subject groups in terms of age ($H(2) = 3.71$, ns.), education ($H(2) = 0.93$, ns.) and intelligence ($H(2) = 4.79$, ns.).

Table 1 summarises the characteristics of all participants with aphasia. In the current study, we used the diagnostic label "transcortical motor aphasia" to refer to patients with aphasia whose language output was nonfluent, extremely reduced, fragmentary echoic and perseverative after 1 month post-onset. Their performance was also impaired on picture-naming, but to a remarkably lower extent than in spontaneous speech. Naming reflected word-finding difficulties, most frequently hesitations, pauses and perseverative errors. Comprehension at the word level as well as at the level of one-part commands was intact, but showed problems at the level of two-part commands. Repetition was good, nearly normal for all participants. During therapy, output became more fluent but still anomic, especially in conversation. In naming, they demonstrated only a milder anomia with long latencies and hesitations. Comprehension and repetition developed to a normal level.

Participants with conduction aphasia showed good auditory comprehension, fluent spontaneous speech interrupted by phonemic paraphasias and self-correction of errors. Compared to spontaneous speech, repetition was severely impaired. Word-finding problems were prominent in naming, coupled with phonemic paraphasias and pauses.

Materials, designs and procedures of the EF tasks

To assess EF, we designed four nonverbal tasks that reduced the influence of impaired linguistic ability on task performance. We focused on two major processes related to EF, updating of working memory representations and inhibition of prepotent responses. All experiments were run by E-Prime (Psychology Software Tools, Pittsburgh, PA, Version 1.2) except for the auditory *n*-back task that was programmed and run by Presentation® software (Version 14.1) on an IBM T40p thinkpad. Participants used the buttons on the keyboard to respond. All participants completed the experiments in two sessions, each lasting 1–1.5 hours, depending on the length of self-paced pauses between the experimental tasks.

Tasks measuring updating of working memory representations

We designed two modified *n*-back tasks, one relying on auditory processing and the other relying on visual processing. The *n*-back task is generally used to index updating of information maintained in working memory (e.g., Miyake et al., 2000).

Auditory n-back task

Participants were exposed to a stream of tones. One tone was presented on each trial and participants had to respond when the stimulus presented was identical to the one appearing in n trials before. We varied n within subjects, and all participants performed the *n*-back task with $n = 1$, then with $n = 2$. In both conditions, the task

TABLE 1
Background description of the participants with aphasia

Participant	Age	Gender	Education	Aetiology	Lesion	Time post-onset (months or days)	Aphasia type (WAB profile)	Aphasia quotient (WAB AQ)	WAB subtests' scores: Spontaneous speech IC	WAB subtests' scores: Spontaneous speech Flu	WAB subtests' scores: Comprehension	WAB subtests' scores: Repetition	WAB subtests' scores: Naming	Boston scores
H. J.	63	M	17	CVA	Infarct of the left MCA	7.5 m	TMA	76.1	6	5	9.85	10	7.2	29
T. I.	66	M	11	CVA	Left frontal infarct	12 m	TMA	85.4	9	8	7.9	10	7.8	41
K. J.	69	M	11	CVA	Left fronto-parietal infarct	9.5 m	TMA	90.1	9	9	9.05	10	8	43
N. F.	57	M	11	CVA	Left frontal infarct	12 m	TMA	94	9	9	10	10	9	57
P. T.	35	F	13	CVA	Left frontal and insular infarct	1 m	TMA	68.4	7	5	9.6	8	4.6	29
V. V.	55	M	11	CVA	Infarct of the left MCA	2 m	Conduction	63.6	7	6	8.2	5.2	5.4	22
O. Gy.	48	M	13	TBI	Bilateral parietal contusion	6 m	Conduction	44.8	5	6	7.6	1.6	2.2	4
Sz. I.	59	M	13	CVA	Left parietal infarct	14 m	Conduction	62	7	6	8.3	4.7	5	18
K. J.	48	M	11	CVA	Left fronto-temporal infarct	21 m	Conduction	43.1	5	5	7.35	1	3.2	5
Cz. J.	55	F	8	CVA	Left fronto-temporal infarct	5 d	Conduction	77.4	8	8	9.3	6.4	7	28

CVA: cerebrovascular accident, TBI: traumatic brain injury, MCA: middle cerebral artery, TMA: transcortical motor aphasia, IC: information content, Flu: fluency.

consisted of 5 blocks of 30 trials. Blocks were separated by self-paced resting periods. The first blocks in both conditions were used as practice blocks. The results show data from Blocks 2–5 in both conditions. On each trial, a sound was sampled from a pool of eight pure frequency sounds (ca. half sounds starting from the standard musical note A5: 440 Hz, 490 Hz, 540 Hz, 590 Hz, 640 Hz, 690 Hz, 740 Hz and 790 Hz). Sampling was fully randomised so that on each trial the chance of sampling a sound that was presented *n* trials before was one to four. In each trial, the sound was presented for 300 ms, followed by a silent period of 1000 ms, during which participants had time to respond. Trials were separated by a 500 ms intertrial interval. In the practice blocks, all trials were followed by a 1000 ms feedback trial if the participant pressed the response button. No feedback was provided in Blocks 2–5.

Visual n-back task

Participants were exposed to a stream of pictures from 14 different semantic categories (e.g., dogs, windows). One picture appeared on each trial and participants had to respond by pressing the ENTER on the keyboard when the stimulus presented was from the same semantic category as the one presented *n* trials before. We varied *n* within subjects, and all participants performed the *n*-back task with $n = 1$, then with $n = 2$. In both conditions, the task consisted of 60 trials. On each trial, a picture was sampled from a pool of pictures of a given semantic category. Sampling was pseudorandomised so that in both conditions for all participants, 10 trials required a hit response. In each trial, the picture was presented in the middle of the screen for 2500 ms. Trials were separated by a 500 ms intertrial interval.

Tasks measuring inhibition

We used two modified inhibition tasks to measure different types of conflict resolution (Lukács, Kemény, Fazekas, Ladányi, & Németh, Unpublished manuscript). The Stop-signal task is generally used to index the ability to resolve response-based conflict through inhibition (Logan, 1994; Milham et al., 2001), while the Stroop task is generally used to assess the ability to resolve representational conflict through inhibition (Novick et al., 2005; Stroop, 1935).

Stop-signal task

On each trial, a stimulus (either a circle or a square) appeared in the middle of the screen for 2000 ms, and participants had to respond as fast as possible by pressing the corresponding button on the keyboard ("c" for circle, "b" for square). Trials were separated by a 250 ms fixation trial (a fixation cross was presented in the middle of the screen and the participant had to fixate on it until the next trial was presented). On some trials, a loud tone was presented after the onset of the stimulus that signalled to the participants that they should refrain from responding (stop trials). Delay of tone onset (Stimulus onset Asynchrony—SOA) was varied within subjects, so that it was increased from 50 ms to 350 ms by steps of 50 ms, through seven blocks. Each block consisted of 60 trials, with 15 stop trials and 45 go trials (trials where the loud tone was absent) randomly intermixed. The seven blocks were separated by self-paced resting periods. Following previous work by Logan (1994), we used proportion of correct rejections (not pressing any button on a stop trial) as a measure of inhibition.

Nonverbal Stroop task

On each trial an arrow was presented on the screen for 3000 ms, with four possible directions (left, right, up, down) and four possible positions relative to the centre of the screen (left, right, over, below). We varied congruency of position and direction within subjects. In the congruent condition, direction matched position (e.g., an arrow pointing to the left, presented on the left side of the screen). In the incongruent condition, direction did not match position (e.g., an arrow pointing to the left, presented on the right side of the screen). Participants had to press the arrow button on the keyboard corresponding to the direction of the arrow on the screen as fast as possible.

This experiment started with 60 control trials that also helped participants to familiarise with matching directions to buttons. In each control trial, an arrow appeared in the middle of the screen (pointing to either of the four possible directions) and participants had to press the corresponding arrow button on the keyboard. (Note that there were no congruent or incongruent trials during these control trials.) Congruent and incongruent trials were blocked so that each participant performed a block of 60 congruent trials followed by a block of 60 incongruent trials. The two blocks were separated by a self-paced resting period. As a measure of inhibition, we used the difference between proportions of errors in the incongruent vs. congruent conditions.

RESULTS

Data were screened for outliers. We excluded all data points that were more than two standard deviations from the group mean (we performed this analysis separately for healthy controls and participants with aphasia). Altogether, less than 3% of the data were removed. For each experimental task, we compared the performance of participants with TMA and conduction aphasia to that of healthy controls. For all comparisons reported we used the nonparametric equivalent of the independent *t*-test, the Mann–Whitney test.

Auditory *n*-back task

Average hit rates for the one-back condition and the two-back condition in the three groups are presented in Figure 1(A). Mann–Whitney tests showed that TMA participants' hit rate in the two-back condition was significantly lower than that of healthy controls, $U = 5.0$, $p = .05$. This occurred despite the fact that their performance did not differ in the one-back condition, $U = 15$, $p = .64$. That is, impaired updating performance was accompanied by an intact ability to discriminate between sounds. Hit rates of patients with conduction aphasia in the two-back condition also differed from that of healthy controls, $U = 5.5$, $p = .05$. However, their hit rates were already lower (at the level of tendency) than that of healthy controls in the one-back condition, $U = 7.5$, $p = .1$.

Follow-up Mann–Whitney tests showed that false alarm rates did not differ significantly from those of healthy controls, in either the one-back ($U = 9$, $p = .20$ for both TMA vs. controls, and conduction vs. controls contrasts) or the two-back condition ($U = 11$, $p = .33$ for the TMA vs. controls contrast, and $U = 16$, $p = .83$ for the conduction vs. controls contrast).

dc_1365_16

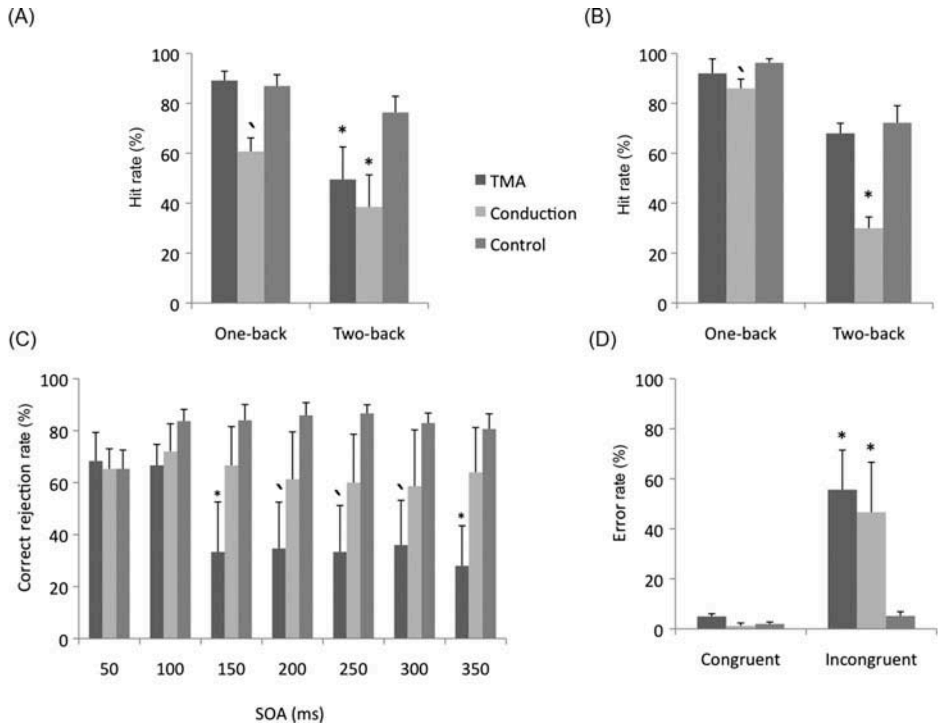


Figure 1. Indicators of executive functions measured in four tasks for the three experimental groups. Asterisks indicate significant differences between the patient and the control groups (*$p < .05$), primes indicate tendencies for differences between the patient and the control groups (‘$p < .1$). (A) Auditory n-back task, with $n = 1$ and $n = 2$. Updating of working memory is assessed by hit rates in the two-back condition. (B) Visual n-back task, with $n = 1$ and $n = 2$. Updating of working memory is assessed by hit rates in the two-back condition. (C) Stop-signal reaction time task, with seven levels of stimulus onset asynchrony (SOA) defined as the onset time of the Stop-signal minus the time onset of the target stimulus. Resolution of response-based conflict through inhibition is assessed by the rate of correct rejections on trials where a Stop-signal occurred (stop trials). (D) Nonverbal Stroop task, with a congruent and an incongruent condition. Resolution of representational conflict through inhibition is assessed by the difference in error rates between congruent and incongruent conditions. Asterisks in this panel mean that this difference in both patient groups was significantly larger than in the control group. Error bars represent standard error of the mean.

Visual *n*-back task

Average hit rates for the one-back condition and the two-back condition in the three groups are presented in Figure 1(B). Mann–Whitney tests showed that performance of patients with TMA did not differ from that of healthy controls in either conditions, $U = 18$, $p = .83$ for the one-back, and $U = 20.5$, $p = .79$ for the two-back conditions. In contrast, patients with conduction aphasia performed worse than healthy controls in the one-back condition at the level of tendency, $U = 7$, $p = .065$, and their performance was significantly worse than that of healthy controls in the two-back condition, $U = 1$, $p = .002$.

Again, follow-up Mann–Whitney tests showed that false alarm rates did not differ significantly from those of healthy controls, in either the one-back ($U = 14$, $p = .80$ for the TMA vs. controls contrast, and $U = 9.5$, $p = .13$ for the conduction vs. controls contrasts) or the two-back condition ($U = 15$, $p = .52$ for the TMA vs. controls contrast, and $U = 11.5$, $p = .46$ for the conduction vs. controls contrast).

Stop-signal task

We plotted the percentage of correct rejections at all seven SOAs for the three groups in Figure 1(C). This measure shows how often participants could successfully stop responding on stop trials, i.e., refrain from pressing any response button, when a Stop-signal required them to do so.

As can be seen in Figure 1(C), the pattern of performance changed, as a function of SOAs, in different ways in the three groups. Mann–Whitney tests confirmed this pattern. At SOAs 150 through 350 TMA patients refrained from stopping their answer on stop trials less often than healthy controls, although this difference did not reach the level of significance at all SOAs ($U = 8.5$, $p = .039$ at SOA = 150 ms, $U = 8.0$, $p = .051$ at SOA = 200 ms, $U = 8.5$, $p = .059$ at SOA = 250 ms, $U = 8.5$, $p = .061$ at SOA = 300 ms, and $U = 6$, $p = .019$ at SOA = 350 ms.) The same comparisons between the conduction aphasia and the control groups did not yield any significant differences (all Us > 13, ns.).

Nonverbal Stroop task

We plotted the error rates in both the congruent and the incongruent conditions for the three groups in Figure 1(D). A larger difference in error rates between the congruent and the incongruent condition indicates a lower degree of representational conflict resolution through inhibition.

Mann–Whitney tests showed that difference in error rates between the congruent and the incongruent condition among TMA participants was significantly higher than among healthy controls, $U = 0.0$, $p = .003$. Similarly, patients with conduction aphasia also produced significantly more errors in the incongruent than in the congruent condition compared to healthy controls, $U = 3.5$, $p = .024$.

DISCUSSION

In the current study, we used four nonverbal tasks to investigate different components of EF in TMA and in conduction aphasia in an attempt to test whether executive dysfunctions are specific to TMA. Our results demonstrate executive deficits among individuals with transcortical motor *and* with conduction aphasia. Importantly, our data revealed different patterns of performance in the two aphasia types.

Deficits of EF among individuals with TMA was evident on several measures: compared to healthy controls, these participants were impaired in resolving response-based conflict (as shown by their performance on the Stop-signal task), in resolving representational conflict through inhibition (as shown by results of the nonverbal Stroop task) and in updating working memory representations (as shown by results of the auditory n-back task).

Importantly, these results are in line with theories (Alexander, 2006; Ardila, 2010; Luria, 1973) predicting that TMA patients will present extensive EF deficits. Earlier, it has been shown that impaired ability to resolve representational conflict results in difficulties in lexical selection and in impairment of syntax (e.g., Alexander, 2006;

Novick et al., 2009). These in turn might explain word-finding difficulties, hesitations and reduced grammar in a nonfluent spontaneous speech in TMA, and in dynamic aphasia. Poor ability to resolve representational conflicts can also lead to unwanted interruptions, sudden topic changes and a general difficulty to stay on topic in narrative discourse (Penn et al., 2010). On the other hand, the inability to resolve response-based conflict can lead to perseverations of communicative strategies. In addition, the disrupted ability to update working memory representations can disturb management of the temporal integration of conversations (Penn et al., 2010). Although this updating deficit was not observed in our visual *n*-back task, it is possible that the two-back condition of the task was not demanding enough to tap the differences between healthy controls and participants with TMA, and that administering a three-back condition might reveal significant differences. Taken together, we suggest that TMA patients have deficits in both inhibition and updating which might explain a range of narrative discourse problems often observed in TMA (e.g., Alexander, 2006).

Our findings also clearly demonstrate that TMA is not the only type of aphasia exhibiting executive dysfunctions: individuals with conduction aphasia also performed poorer than healthy controls on several EF measures.

First, as evidenced by results of the Stroop task, compared to healthy controls, these participants were impaired in resolving representational conflict. Second, they performed generally worse than healthy controls on the auditory *n*-back task. Their performance was already below that of healthy controls in the one-back condition which might be attributed to deficits of both auditory discrimination and working memory. For the interpretation of these results, it is important to note that (see Figure 1(A)) the pattern of performance in the two groups was similar, i.e., increasing working memory load (from one-back to two-back) did not decrease performance in conduction aphasia more than in healthy controls, as one would expect in case of a marked deficit of updating functions (Vasic, Walter, Sambataro, & Wolf, 2009; Waltz et al., 2004). Unfortunately, the fact that the possible deficits in working memory and auditory discrimination are confounded in this task puts limitations on the use of auditory *n*-back task in investigating EF in conduction aphasia. Third, in the visual modality, we observed a clear deficit of updating working memory representations in conduction aphasia. Because the pictures used in this task depicted everyday objects for which both object names and category names are easy to verbalise, participants could lean in part on subvocal rehearsal (Baddeley, Eysenck, & Anderson, 2009). Subvocal rehearsal is known to be affected in conduction aphasia (Buchsbaum et al., 2011), and this might have resulted in their lower performance in the two-back condition of this task.

In brief, the poor level of performance on the *n*-back tasks might be caused by different impairments in the two aphasia types. Whereas in TMA, the deficit of updating seems to be the main factor explaining the results of the *n*-back task, in conduction aphasia, deficits in subvocal rehearsal (in the case of the visual *n*-back task) and auditory discrimination (in the case of the auditory *n*-back) may also contribute to poor performance.

The different patterns of performance displayed in the two types of aphasia on the two inhibitory tasks (with both groups showing deficits on the nonverbal Stroop task, but only the TMA group showing a deficit on the Stop-signal task) support the view (Milham et al., 2001; Novick et al., 2005) that these tasks indeed measure distinct components of inhibitory executive processes. Novick et al. (2009, 2010) suggested that the resolution of representational conflict, conventionally indexed by the Stroop task, is associated with language abilities. Accordingly, the same authors as well as others (Robinson et al., 1998) have suggested that this type of conflict resolution is associated with TMA. Our results provide support for this suggestion but also show that the impairment of representational conflict resolution is present in conduction aphasia as well. The results of the Stop-signal task used in our study provide evidence for the involvement of response-based conflict resolution in TMA, but not in conduction aphasia. Based on Penn et al. (2010) we suggest that these inhibitory deficits might have separable contributions to narrative discourse impairment in TMA, and lead together to the overall pattern of language symptoms in TMA.

Our study is the first systematic assessment of EF in aphasia to demonstrate clear and extensive executive deficits among TMA patients. We hope that it might serve as a starting point for future research addressing the exact relationship between different EF components and language abilities. The most intriguing question for this line of research is the causal relationship between EF deficits and language disorders. Earlier results suggest that EF deficits might manifest in language symptoms in TMA. In conduction aphasia, however, it is possible that the observed EF deficits are only associated with language symptoms, but are not causal in their development. Whatever the answer to these questions, examination of EF components and training of EF should be an essential part of clinical aphasia assessment and rehabilitation.

Manuscript received 30 March 2013
Manuscript accepted 14 August 2013
First published online 14 September 2013

REFERENCES

Alexander, M. P. (2003). Transcortical motor aphasia: A disorder of language production. In M. D'Esposito (Ed.), *Neurological foundations of cognitive neuroscience* (pp. 165–174). London: MIT.

Alexander, M. P. (2006). Impairments of procedures for implementing complex language are due to disruption of frontal attention processes. *Journal of the International Neuropsychological Society*, *12*, 236–247.

Ardila, A. (2010). A proposed reinterpretation and reclassification of aphasic syndromes. *Aphasiology*, *24*, 363–394.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Hove: Psychology Press.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*, 539–546.

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory – an aggregate analysis of lesion and fMRI data. *Brain and Language, 119*, 119–128.

Code, C., Tree, J. J., & Dawe, K. (2009). Opportunities to say "yes": Rare speech automatisms in a case of progressive nonfluent aphasia and apraxia. *Neurocase*, *15*, 445–458.

Frankel, T., Penn, C., & Ormond-Brown, D. (2007). Executive dysfunction as an explanatory basis for conversation symptoms of aphasia: A pilot study. *Aphasiology*, *21*, 814–828.

Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language*, *86*, 23–39.

Green, D. W., Grogan, A., Crinion, J., Ali, N., Sutton, C., & Price, C. J. (2010). Language control and parallel recovery of language in individuals with aphasia. *Aphasiology*, *24*, 188–209.

Gvion, A., & Friedmann, N. (2012). Does phonological working memory impairment affect sentence comprehension? A study of conduction aphasia. *Aphasiology*, *26*, 494–535.

Hardin, K. H., & Ramsberger, G. (2004). *Treatment of attention in aphasia: A case study*. Poster presented at Clinical Aphasiology Conference, Park City, UT.

Helm-Estabrooks, N., & Albert, M. L. (1991). *Manual of aphasia therapy*. Austin, TX: Pro-Ed.

Helm-Estabrooks, N., Connor, L. T., & Albert, M. L. (2000). Treating attention to improve auditory comprehension in aphasia. *Brain and Language*, *74*, 469–472.

Kaplan, E., Goodglass, H., Weintraub, S., & Segal, O. (1983). *Boston naming test*. Philadelphia, PA: Lea & Febiger.

Keil, K., & Kaszniak, A. W. (2002). Examining executive function in individuals with brain injury: A review. *Aphasiology*, *16*, 305–335.

Kertész, A. (1982). *The Western aphasia battery*. New York, NY: Grune-Stratton.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Lukács, Á., Kemény, F., Fazekas, K., Ladányi, E., & Németh, K. Verbal and nonverbal executive functions in Specific Language Impairment. Manuscript submitted for publication.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York, NY: Basic Books.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*, 467–473.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, *26*, 527–567.

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

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

Osmánné Sági, J. (1991). Az afázia klasszifikációja és diagnosztikája. *Ideggyógyászati Szemle*, *44*, 339–362.

Penn, C., Frankel, T., Watermeyer, J., & Russell, N. (2010). Executive function and conversational strategies in bilingual aphasia. *Aphasiology*, *24*, 288–308.

Purdy, M. (2002). Executive function ability in persons with aphasia. *Aphasiology*, *16*, 549–557.

Ramsberger, G. (2005). Achieving conversational success in aphasia by focusing on non-linguistic cognitive skills: A potentially promising new approach. *Aphasiology*, *19*, 1066–1073.

Robinson, G., Blair, J., & Cipolotti, L. (1998). Dynamic aphasia: An inability to select between competing verbal responses? *Brain*, *121*, 77–89.

Schnur, T. T., Lee, E., Coslett, H. B., Schwartz, M. F., & Thompson-Schill, S. L. (2005). When lexical selection gets tough, the LIFG gets going: A lesion analysis study of interference during word production. *Brain and Language*, *95*, 12–13.

Smith, E. E., Jonides, J., Marshuetz, C., & Koeppe, R. A. (1998). Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences*, *95*, 876–882.

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

Vasic, N., Walter, H., Sambataro, F., & Wolf, R. C. (2009). Aberrant functional connectivity of dorsolateral prefrontal and cingulate networks in patients with major depression during working memory processing. *Psychological Medicine*, *39*, 977.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Back-Madruga, C., McPherson, S., . . . Miller, B. L. (2004). Relational integration and executive function in Alzheimer's disease. *Neuropsychology*, *18*, 296–305.

Study 8.

Lukács Á., Kemény F., (2014). Domain general sequence learning deficit in Specific Language Impairment. *Neuropsychology*, 28(3), 472-483. doi: 10.1037/neu0000052

Thesis 6. Cognitive impairments in SLI do not selectively target language; deficits also occur in skill learning outside the language domain, most prominently for sequentially organized stimuli (3, 8.)

Thesis 7. In concert with the Procedural Deficit hypothesis, procedural learning is vulnerable in SLI, while processes of declarative learning and retention are relatively intact (3, 8, 12)

Neuropsychology
2014, Vol. 28, No. 3, 472–483

0894-4105/14/$12.00 DOI: 10.1037/neu0000052

Domain-General Sequence Learning Deficit in Specific Language Impairment

Ágnes Lukács and Ferenc Kemény
Budapest University of Technology and Economics

Objective: Grammar-specific accounts of specific language impairment (SLI) have been challenged by recent claims that language problems are a consequence of impairments in domain-general mechanisms of learning that also play a key role in the process of language acquisition. Our studies were designed to test the generality and nature of this learning deficit by focusing on both sequential and nonsequential, and on verbal and nonverbal, domains. ***Method:*** Twenty-nine children with SLI were compared with age-matched typically developing (TD) control children using (a) a serial reaction time task (SRT), testing the learning of motor sequences; (b) an artificial grammar learning (AGL) task, testing the extraction of regularities from auditory sequences; and (c) a weather prediction task (WP), testing probabilistic category learning in a nonsequential task. ***Results:*** For the 2 sequence learning tasks, a significantly smaller proportion of children showed evidence of learning in the SLI than in the TD group (χ^2 tests, $p < .001$ for the SRT task, $p < .05$ for the AGL task), whereas the proportion of learners on the WP task was the same in the 2 groups. The level of learning for SLI learners was comparable with that of TD children on all tasks (with great individual variation). ***Conclusions:*** Taken together, these findings suggest that domain-general processes of implicit sequence learning tend to be impaired in SLI. Further research is needed to clarify the relationship of deficits in implicit learning and language.

Keywords: implicit learning, sequence learning, language impairment, probabilistic categorization

Children with specific language impairment (SLI) have a primary deficit in language abilities. SLI involves a significant delay in the acquisition of language in the absence of any hearing deficits, neurological disorders, emotional and social problems, environmental deprivation, or mental retardation that could account for their language problems. Although there are claims that, in SLI, language is selectively impaired in an otherwise intact cognitive system (as the definition and the term implies; Clahsen & Hansen, 1997; Gopnik & Crago, 1991; Rice, Wexler, & Cleave, 1995; van der Lely, 1997; van der Lely & Stollwerck, 1996), research in the past two decades has shown that the impairment in many (if not in most) cases turns out to be not as specific to language as originally claimed.

Results show that SLI is often associated with impairments in several nonlinguistic functions, but the nature, extent, and generality of these deficits is yet unclear, as is their relationship with language abilities. Most often, impairments are reported in the domains of coordination of oral and fine movements, hypothesis testing and categorization (of linguistic and nonlinguistic materials), mental rotation, sequencing, parallel tasks, and executive functions (Bishop & Edmundson, 1987; Bishop & Norbury, 2005; Henry, Messer, & Nash, 2012; Hill, 2001; Im-Bolter, Johnson, & Pascual-Leone, 2006; Powell & Bishop, 1992; Zelaznik & Goffman, 2010; for a review, see Leonard, 1998). Some propose that the core impairment concerns more fundamental nonlinguistic mechanisms, like the processing of rapidly changing auditory stimuli; others argue that the main cause of language impairment is the reduced capacity of verbal short-term memory (STM) or some kind of processing limitation (Bishop, 1992; Gathercole & Baddeley, 1990, 1993; Graf Estes, Evans, & Else-Quest, 2007; Joanisse & Seidenberg, 1998; Leonard et al., 2007; Montgomery, 2000, 2002; Tallal & Piercy, 1973; Tallal et al., 1996; for reviews, see Bishop, 2006; Leonard, 1998). On these accounts, higher order language impairment at the grammatical level is a consequence of such lower level deficits.

Several accounts point out that SLI also involves a learning deficit, but studies addressing the processes and basic mechanisms

This article was published Online First March 3, 2014.

Ágnes Lukács and Ferenc Kemény, Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary.

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. During the period of research and writing the article, Ágnes Lukács was a grantee of the János Bolyai Research Scholarship of the Hungarian Academy of Science. This work is connected to the scientific program of the "Talent care and cultivation in the scientific workshops of BME (Budapest University of Technology and Economics, Hungary)" project and is supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.2/B-10/1-2010–2009).

This article benefited a lot from discussions with Vinciane Gaillard, Rebecca Gómez, Darlene V. Howard, James H. Howard, Karolina Janacsek, Beat Meier, Dezső Németh, and Kathleen M. Thomas. We are grateful for all their helpful comments. We would also like to thank Kornél Németh, Gábor Szvoboda, Kata Fazekas, Enikő Ladányi, Emese Lukács, Borbála Győri, and psychology students of the Péter Pázmány Catholic University who were instrumental in collecting the data.

Correspondence concerning this article should be addressed to Ágnes Lukács, Budapest University of Technology and Economics Department of Cognitive Science, Egry J. u, 1, H-1111 Budapest, Hungary. E-mail: alukacs@cogsci.bme.hu

of learning are difficult to find. We find reference to learning deficits in different forms in both linguistic and nonlinguistic accounts. Gopnik and Crago (1991) argue that "dysphasics do not have the normal language acquisition mechanism described by Pinker (1984) that would allow, or, perhaps, even compel them to construct inflectional paradigms on the basis of regularities hypothesized on the basis of observed linguistic evidence" (p. 47). The authors describe language impairment as a deficit of implicit grammatical rule acquisition (compensated by explicit acquisition of grammatical rules and productive morphological forms). Leonard's surface (Leonard, 1989; Leonard, Eyer, Bedore, & Grela 1997; Leonard, McGregor, & Allen, 1992) and morphological richness (Dromi, Leonard, Adam, & Zadoneisky-Ehrlich, 1999; Leonard, 1998, pp. 255–257; Leonard, Sabbadini, Leonard, & Volterra, 1987) accounts build on a general processing limitation; he describes the learning problem stemming from this processing limitation as a "problem of slow intake of relevant data due to the reduced speed of processing" (Leonard, 1998, p. 249). Bates and Goodman (2001; see also Marchman & Bates, 1994) also consider slower learning to be central to SLI (and other developmental disorders): The deficit in working memory leads to slower learning of items and, this way, also to a difficulty in reaching the "critical mass" of lexical representations required for grammatical generalizations.

Implicit Learning in SLI

The aforementioned proposals for a learning problem imply, but do not specify, a problem with implicit learning (IL) in SLI. There are also suggestions that explicitly focus on the nature of the learning deficit in language impairment; they differ in their details, but they share the view that the problem in SLI concerns IL, a form of learning that is associated mainly with the acquisition of motor, cognitive, and social skills. This type of learning takes place slowly and gradually over many trials, with an output that is manifest in performance, but is not available to conscious introspection (e.g., Cleeremans, Destrebecqz, & Boyer, 1998; Frensch & Ruenger, 2003; Reber, 1967, 1993).

Two of the approaches with an emphasis on learning specifically target SLI. The procedural deficit hypothesis (PDH; Ullman & Pierpont, 2005) suggests a central impairment in the procedural system that—among other symptoms—presents itself as a grammatical deficit within the domain of language. Others argue that difficulties in extracting statistical regularities and abstracting away from them is the core problem that can present itself in any area of language, not just grammar, and in nonlinguistic domains as well (Evans, Saffran, & Robe-Torres, 2009; Hsu & Bishop, 2011). A third relevant theory highlights the overlap of IL of sequential information and linguistic processes, and expects different forms of language impairment to involve deficits in sequential learning as well (e.g., Christiansen, Conway, & Onnis, 2012; Christiansen, Kelly, Shillcock, & Greenfield, 2010). We now turn to presenting each proposal, with empirical support for them provided in detail.

The PDH (Ullman & Pierpont, 2005) tries to incorporate linguistic and nonlinguistic impairments into a unified theory and a neurobiological model. On this view, language impairment is a result of abnormal development of brain structures underlying the procedural memory system responsible for learning cognitive and motor skills, and, among them, grammar. Developmental disorders of such a system should result in deficits of skills based on procedural learning within both the linguistic and nonlinguistic domains. Theoretically, any part of the network can be impaired, which explains the heterogeneity of SLI. Because parts of the network can be impaired to a different extent, and because they have neural connections to other areas as well, SLI is not necessarily a selective impairment. The PDH also predicts that the better declarative memory abilities (associated more with the lexicon and with explicit learning) are, the less conspicuous SLI is, because of compensatory mechanisms like storing information in chunks and learning explicit rules.[1]

Several observations are in concert with this hypothesis. As discussed in the introductory section of this article, language disorders are often accompanied by attention deficits, motor problems, or working memory impairment, which are also subserved by structures constituting the procedural system. The association between language impairment and weak (sequential and nonsequential) motor skills seems to be strong, and the same is true for the relationship of working memory and language (see previous references). The PDH predicts that the brain structures involved are part of the procedural system. Neurobiological research supports this hypothesis: When compared with typical development, the greatest differences are found in the basal ganglia (especially in the caudate nucleus) and in the Broca region of the frontal lobe (Belton, Salmond, Watkins, Vargha-Khadem, & Gadian, 2002; Gauger, Lombardino, & Leonard, 1997; Vargha-Khadem et al., 1998; although the planum temporale is also involved: Gauger et al., 1997; Plante, Swisher, Vance, & Rapcsak, 1991).

Most relevantly for the proposal of a learning deficit, there are results showing impaired IL of both specific and abstract sequences in SLI. In the earliest study motivated by the PDH, Tomblin, Mainela-Arnold, and Zhang (2007) tested motor-sequence learning abilities in 85 adolescents with SLI using a version of the serial reaction time task (SRT; Nissen & Bullemer, 1987, described in detail in the Method section). In this task, a target stimulus appears in one of four locations, and the role of the participant is to press the key that corresponds to the stimulus location. Unknown to the participant, the appearance of the stimulus is determined by a preset sequence. Reaction times (RTs) decrease as long as the sequence is present, but as soon as the appearance becomes random, RTs increase (Nissen & Bullemer, 1987). This task is used by several studies because it is a well-established procedural learning task, and because it is a visuomotor learning task not involving linguistic knowledge. This ensures that when poor sequence learning performance is observed in SLI, it cannot be due to deficits in speech perception or to reduced phonological STM. Tomblin et al. (2007) observed slower learning rates in those adolescents with SLI who had a grammatical im-

[1] The explicit–implicit distinction is based on recollection, whereas the declarative–procedural distinction is rooted in representational differences. Although there is a difference between the distinctions, most studies treat the notions of implicit and procedural as identical, just like the notions explicit and declarative (e.g., Price, 2009). Others, on the other hand, suggest that the two distinctions should not be mapped onto each other (e.g., Berry & Dienes, 1993). Trying to separate the processes involved in statistical/implicit/procedural learning tasks is beyond the scope of the article, and we will not differentiate between them here, admitting that the theoretical questions these concepts raise is a very important one.

pairment in kindergarten, but not in individuals with a primary lexical impairment.

These findings were seemingly countered by results from a probabilistic version of the SRT task by Gabriel, Maillart, Guillaume, Stefaniak, and Meulemans (2011), who argue against a general procedural learning deficit in SLI, based on finding the same level of sequence learning in 16 children with SLI between the ages of 7 and 13 years as in age-matched typically developing (TD) children. This apparent contradiction may have been due to age and task differences, and a relatively small number of participants varying greatly in age in Gabriel et al.'s study. Gabriel et al. explained it in terms of differences in sequence complexity. In a following study (Gabriel et al., 2013) with 21 children with SLI using more complex sequences with second-order probabilities, they found results supporting the PDH: There was no evidence of sequence-specific learning in the SLI group.

An extensive study by Lum, Conti-Ramsden, Page, and Ullman (2012) found further evidence supporting the PDH by testing 51 children with language impairment and TD children matched on chronological age. Together with procedural learning, the authors also covered working memory and declarative memory functions to test whether the deficit selectively impairs procedural functions. The procedural deficit in SLI was evident in less successful sequence learning on the SRT task (even after controlling for working memory differences). Declarative memory functions were spared in both the visual and verbal domains (although in the verbal domain, this was only true after controlling for working memory and language level). The SLI group also displayed a deficit in verbal, but not in visuospatial, working memory. Importantly for the PDH, declarative memory measures were associated with lexical abilities in SLI and TD, but whereas grammar was associated with procedural performance in TD, it correlated with declarative memory in SLI, suggesting that, in line with PDH, spared declarative memory functions support grammar and are used as a compensating mechanism for the grammatical impairment resulting from the procedural deficit.

Besides testing initial learning on a procedural task, Hedenius et al. (2011) expanded testing to more than one day to learn whether children with SLI show consolidation deficits in procedural learning. Participants were tested with the alternating serial reaction time (ASRT) task in four sessions on Day 1, and they were tested again in a fifth session, an average of 3 days later, to see whether they showed effects of consolidation and preserved sequence-learning effects in the long term as well. Grammar-impaired and TD children did not differ in their measures of initial learning, and both groups showed the same amount of sequence-specific learning. Notably, in contrast to TD children, children with a grammatical impairment did not show a consolidation effect, and by Session 5, they lost the sequence-specific knowledge they gained in the initial four learning sessions. Hedenius and colleagues also tested whether individual differences in vocabulary, nonverbal IQ, and grammatical knowledge were associated with procedural learning (in the entire sample of the two groups combined), and only found an association with grammatical knowledge. These conclusions have to be taken with caution, though, because, in fact, no significant group difference is reported; the group difference after consolidation was only marginally significant, as were interactions with group.

A different set of studies tested implicit statistical learning with verbal stimuli. These results warrant further controls, as the deficient learning in the verbal domain can also be caused by well-known deficits in speech perception or phonological STM. Evans and colleagues (2009) compared performance of children with SLI (35 children between 6.5 and 14.4 years) on statistical learning tasks of word and tone segmentation. In a continuous stream of syllables or tones, words were defined by transitional probabilities, which are higher for units that are part of the same word than for units spanning a word boundary. After 21 min of exposure, children with SLI performed at chance on the word segmentation task (on all words), compared with 58% correct answers in the control group, which was significantly better than chance (this was true for all words). Children with SLI did show evidence of word segmentation with a double exposure time of 42 min, as did the controls, and the two groups' performance did not differ with this longer training (SLI = 56.2%; control = 64.4%). This suggests that, besides problems with grammar or serial recall, children with SLI might already have a deficit at the earliest stages of language learning, even with discovering word boundaries relying on co-occurrence information. Group differences were even more evident in the nonlinguistic tone-segmentation task: Chance performance (48%) in the SLI group was outperformed by significant learning at 66% in the control group. Based on impaired learning for both speech and nonspeech stimuli, Evans et al. (2009) argue for deficits in domain-general IL abilities in children with SLI, and they also suggest that this IL deficit has consequences not just for grammar but also for vocabulary.

Evans and colleagues (2009) concentrated on learning specific sequences in a task that parallels word segmentation. Not surprisingly, language impairment is also associated with a deficit in implicit statistical learning when the task goes beyond mapping statistical distributions of adjacent and specific stimuli, and requires the extraction of more abstract patterns and nonadjacent dependencies. Plante, Gómez, and Gerken (2002) found evidence of impaired artificial grammar learning (AGL) in adults with language-based learning disabilities (LLD; chance performance). In a later study, Grunow, Spaulding, Gómez, and Plante (2006) tested the learning of nonadjacent dependencies in a similar group (11 college students with a history of LLD who also met the criteria of SLI). They used an AGL task in which sentences were composed of three nonsense words, and tested how learning of nonadjacent dependencies is altered by the variability of the middle ("noise" element, AXB, e.g., *pel* wadim *jic*; *pel* kicey *jic*). In two conditions, the middle element was picked from either 12 (low variability) or 24 (high variability) possibilities. The LLD group performed at chance in both conditions, whereas the control group was able to learn nonadjacent dependencies in the high variability condition, and was also able to generalize the pattern to new intervening items. Although (perhaps because of low number of participants) group differences are not statistically significant, Grunow et al. argue that "a reduced sensitivity to information that can be used to enhance learning may explain why language acquisition is slow in children with language impairments" (p. 168).

In a recent article, Hsu and Bishop (2011) aim to give a theoretical integration of different results on, and related to, IL in SLI. They propose that, in SLI, the learning system is "a system biased toward memorization of exemplars, and is poor at extracting statistical dependencies from the input" (p. 264). In the past two decades constructivist approaches to language acquisition (e.g., Tomasello, 2000, 2003) has shown that children's initial learning is item-based, and that the pool of item-specific constructions then forms the basis of processes of schematization and generalization, helped by processes of statistical learning, pattern extraction, and distributional analysis. Extensive research on statistical learning abilities of infants demonstrated that basic and domain-general learning skills help both word segmentation and grammar learning (e.g., Gómez & Gerken, 1999, 2000; Saffran, 2002; Saffran, Aslin, & Newport, 1996). Based on these results, Hsu and Bishop argue that children with SLI might have a central problem with statistical learning: Their difficulties concern exploiting the statistical regularities in the input and abstracting and schematization away from them; as a result, they stick more to item-specific representations than TD peers of the same age. This general impairment may affect both procedural and declarative learning.

Hsu and Bishop (2011) enumerate evidence pointing in this direction from different sources. The first is a set of results showing that children with SLI rely on item-based representations, whereas TD peers rely on abstract rules, most frequently demonstrated in the domain of regular–irregular inflection (by regulars also showing a frequency effect: Oetting & Horohov, 1997; Oetting & Rice, 1993; Ullman & Gopnik, 1999; van der Lely & Ullman, 2001). A more item-based nature of linguistic representations in SLI is also supported by analyses of spontaneous speech samples, which demonstrate that SLI children tend to follow adult utterances more closely, and their verb use is less productive and more input- and context-dependent than that of TD children (Jones & Conti-Ramsden, 1997; Skipp, Windfuhr, & Conti-Ramsden, 2002; Stokes & Fletcher, 2000). Another set of evidence comes from AGL studies demonstrating no or diminished learning in SLI (Grunow et al., 2006; Plante et al., 2002; see above). Third, they review findings of deficits in nonverbal sequence learning supporting the PDH (also reviewed here earlier: Lum et al., 2012; Tomblin et al., 2007).

Hsu and Bishop (2011) try to make a distinction between the statistical learning deficit hypothesis and the PDH. They argue that although the PDH reserves the procedural deficit's involvement only to grammar, a statistical learning deficit may extend to other aspects of language as "learning arbitrary associations between sounds and referent . . . and generalizing names for solid objects by shape in learning vocabulary" (p. 9). They go on to argue that the well-documented working memory deficits may also contribute to SLI impairments in statistical learning, especially and centrally to SLI, when long-distance co-occurrence patterns have to be detected (as in the case of A-X-B, A-Y-B, A-Z-B sequences in Plante et al.'s, 2002, study). As a result, they say, children with SLI need more exposure for developing abstract patterns, potentially because of a difficulty in reaching the "critical mass" (although a basic problem in statistical learning itself would seem to be a good candidate for explaining such difficulties).

In a comment on Hsu and Bishop (2011), Dąbrowska (2011) suggests that children with SLI do not have problem with abstraction and generalization per se (this is argued by pointing out that SLI children do demonstrate relatively high levels of productivity, even with problematic morphemes; see, e.g., Marchman, Wulfeck, & Ellis Weismer, 1999; Rice, Wexler, & Hershberger, 1998). According to Dabrowska, the central deficit concerns the slow proceduralization of such knowledge, due to a problem with consolidation necessary for developing fast and automatic retrieval of the learned representations.

As pointed out earlier, AGL studies in the acquisition literature stress the importance of implicit statistical learning in word segmentation and grammar acquisition. Results arguing for common mechanisms involved in *sequence* learning and language processing also imply that IL of sequences is impaired in SLI (Conway, Bauernschmidt, Huang, & Pisoni, 2010; Conway, Pisoni, & Kronenberger, 2009). There is evidence for a connection between implicit sequence learning and language abilities, but the specifics of the domains and mechanisms involved on both sides require further exploration. Conway and colleagues (Conway, Karpicke, & Pisoni 2007; Conway et al., 2010) found a correlation between IL of visual sequences (location sequence learning) and speech perception in noise (participants had to identify last words in sentences that varied in their predictability), but only when the visual task contained easily verbalizable items (when different locations were also associated with different colors). They concluded that implicit sequence learning, when it involves stimuli that are easy to encode verbally, is an important factor in speech and language processing, and it plays a primary role in building "more detailed and robust representations of the word order probabilities in spoken language" (Conway et al., 2010, p. 365).

In another exploration of the association between implicit statistical learning and language, Kidd (2012) tested 100 children between 4.5 and 6.11 years with a version of SRT task adapted for children, explicit declarative learning, and a syntactic priming task. He found that children who showed effects of syntactic priming also showed significantly better IL, whereas no parallel differences in explicit learning were observed. Kidd claims that these findings argue for a direct association between syntax acquisition and IL. Such an association is also confirmed by Misyak and Christiansen's (2012) results: In their study, statistical learning scores on two tasks (one involving the learning of nonadjacent dependencies) remained significant predictors of the individual variation in sentence processing (as measured by a self-paced reading task with sentences involving different types of syntactic complexities), even after controlling for verbal working memory, motivation, vocabulary, and fluid intelligence. Christiansen et al. (2012) further support reliance on common background mechanisms by observing a neural overlap of processes involved in the processing of linguistic and nonlinguistic sequences. In a within-subject design, they observed a P600 with similar distributions for incongruencies in nonlinguistic sequences and for syntactic violations. Based on this similarity, they interpret the P600 as reflecting integration cost for expectations in sequentially organized stimuli.

Aims

These three approaches to the nature of the learning deficit in SLI have different predictions, but although the methods used thus far have proven very useful in finding support for certain aspects of the hypotheses, they cannot clearly distinguish between them. Most studies use sequential learning tasks in either the nonverbal (SRT; Nissen & Bullemer, 1987) or the verbal (versions of AGL tasks, also with nonverbal stimuli; Reber, 1967; also referred to as statistical learning, e.g., Conway & Christiansen, 2005) domain to test IL. Results on these tasks imply that language impairment is often associated with deficits in implicit *sequence* learning, in both the linguistic and nonlinguistic domains. On the other hand, little is known about IL of nonsequential information. To our knowledge, there is only one earlier study presenting evidence for deficits in that domain as well: Children with SLI lagged behind TD peers in the weather prediction (WP) task relying on probabilistic categorization (Kemény & Lukács, 2010).

To get a more precise nature of the IL deficit in SLI, the current study was designed to test both sequential and nonsequential IL, and also to cover both verbal and nonverbal domains. With that aim in mind, three different learning tasks were chosen: an SRT (Nissen & Bullemer, 1987), an AGL task (Saffran, 2002), and the WP task, a probabilistic category learning task (Knowlton et al., 1994). The SRT and AGL are sequence learning tasks; the SRT tests the learning of specific motor sequences and the AGL task tests the extraction of regularities from auditory sequences of verbal stimuli. The WP task is a nonsequential task involving IL through probabilistic categorization. The AGL and WP tasks require the extraction of probabilistic information, whereas the SRT relies on learning deterministic sequences. Although all three tasks are regarded as paradigms for testing IL, besides differences in task design and stimulus structure, neuropsychological evidence suggests that they rely on partially different neural structures. Parkinson's patients, for example, are impaired on the WP task (Knowlton, Mangels, & Squire, 1996) and the SRT (Siegert, Taylor, Weatherall, & Abernethy, 2006), but not on the AGL (Witt, Nühsman, & Deuschl, 2002).[2] These results are supplemented by imaging data, implying that although all tasks are nondeclarative, they may rely on different brain circuitry (Daselaar, Rombouts, Veltman, Raaijmakers, & Jonker, 2003; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999; Skosnik et al., 2002).

Taken together, the literature on the nature of the learning deficit in SLI suggests different predictions on how the performance profile of children with SLI will differ from TD children. If there is a general deficit in statistical learning in SLI, as argued by Hsu and Bishop (2011), the SLI group is expected to show weaker performance on the AGL and WP tasks, in which the *extraction of statistical information* is more difficult, in contrast to the SRT task, in which a second-order deterministic sequence is repeated over and over again. Depending on the severity of the impairment of the learning mechanism, a general implicit statistical leaning deficit is also compatible with a pattern in which all three tasks are affected.[3]

If children with SLI have a deficit in *procedural learning*, as suggested by the PDH (Ullman & Pierpont, 2005), we expect them to show impaired performance on the SRT and WP tasks, and also possibly, but not necessarily, on AGL (the prediction of the PDH is not entirely clear with regard to this task; procedural impairment has been found to be associated with impaired SRT and WP, but not AGL, performance in Parkinson's disease).

As domain-general sequential organization has been shown to be associated with language (Conway et al., 2010; Conway, Pisoni, & Kronenberger, 2009; Misyak & Christiansen, 2012), it is possible that children with SLI have a specific difficulty with *sequential organization*, which would be supported by finding impaired performance on the two sequential tasks (SRT and AGL) but not on the nonsequential WP task. It is also possible that tasks involving verbal stimuli pose extra difficulties for children with SLI; in this case, the deficit should be most prominent on the AGL task.

Method

Participants

Altogether, 29 children with SLI were involved in the study (7 girls, 22 boys). Their mean age was 9.10 years, with a standard deviation of 1.28 years. All children met the criteria for SLI. Each child scored above 85 on the Raven Colored Progressive Matrices (Raven, Court, & Raven, 1987), a measure of nonverbal intelligence. All children passed a hearing screening, and no child had a history of neurological impairment. Each child scored at least 1.5 standard deviations below age norms on at least two of four language tests administered. These four tests included two receptive tests and two expressive tests. The receptive tests were the Hungarian standardizations of the Peabody Picture Vocabulary Test (Csányi, 1974) and the Test for Reception of Grammar (TROG; Bishop, 1983, 2012; Lukács, Győri, & Rózsa, 2012). The expressive tests were the Hungarian Sentence Repetition Test (Magyar Mondatutánmondási Teszt [MAMUT]; Kas & Lukács, 2011) and a nonword repetition test (Racsmány, Lukács, Németh, & Pléh, 2005).

Performance of the SLI group was compared with that of a group of TD children. To reduce the effect of individual differences in the control group, each participant in the clinical group was matched in age with three control participants, yielding a TD group of 87 children (40 girls, 47 boys; mean age = 9.14 years, *SD* = 1.27 years). 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.

Stimuli

All participants completed three different IL tasks: the SRT, the AGL task, and the WP task. All three tasks were presented on a 640 × 480 display, on a computer using E-prime 1.2 (Psychology Software Tools, Inc., Pittsburgh, PA).

SRT task. The SRT task was an adaptation of the task used by Meulemans, Van der Linden, and Perruchet (1998). There were four circles on the screen (diameter of approximately 55 pixels); one of the circles was black and the other three were open, with a black contour. The circles were arranged horizontally in the center line of the screen. The distances between the circles were equal.

AGL task. Stimuli of the AGL task were adapted from the P language of Saffran (2002). The task consisted of a training phase and the test phase. In the training phase, participants heard 58 sentences twice in the same fixed pseudorandom order. The sentences were constructed to follow the rules presented in (1), taken from Saffran (2002). The vocabulary is provided in Table 1.

[2] Other studies fail to report deficient IL in Parkinson's patients, see e.g., Shohamy, Myers, Onlaor, and Gluck (2004), and Smith, Siegert, and McDowall (2001).

[3] As pointed out by one of the reviewers of the article, the predictions of the statistical learning approach for the SRT task are problematic and difficult to formulate precisely. Because the four recurring items create a second-order deterministic sequence, although no long-distance dependencies are involved in the task, the participants have to take into account the current and the preceding element together to be able to predict the following one.

Table 1
Words in the Artificial Language by Category (With Hungarian Transcription)

Category				
A	bif	hep	mib	rud
C	kav	lam	neb	szig
D	lor	gal		
F	dup	dók	rász	vot
G	tez	péf		

The test phase consisted of 24 items, following Saffran (2002). In each item, two sentences were provided, one after the other. One of the sentences was grammatical according to the rules of the grammar presented in point (1) below, while the other included one a violation, again following Saffran (2002) (e.g., *hep gal lam péf dók* [A-D-C-G-F] - * *hep péf lam gal dók* [A-G-C-D-F]; or *rud gal neb dup* [A-D-C-F] -: * *rud gal dup* [A-D-F]). There were six ungrammatical strings for each violation.

$$\begin{aligned} S &\rightarrow AP + BP + (CP) \\ AP &\rightarrow A + (D) \\ BP &\rightarrow CP + F \\ CP &\rightarrow C + (G) \end{aligned} \tag{1}$$

WP task. The WP task was the adaptation of Knowlton et al. (1994), adapted by Kemény and Lukács (2010) for testing children. One, two, or three out of four different cues appeared at the same time. Cue 1 was a square, Cue 2 was a triangle, Cue 3 was a pentagon, and Cue 4 was a rhombus. Each cue was fit into a 125 × 125 pixel square. Cues appeared 144 pixels from the top. The cue or combination appeared in the horizontal center line: If there was only one cue, it was in the center line. If there were two cues, there was a cue on each side of the center line; in the case of three simultaneous cues, the central cue was in the center line. The order of cues in combinations always appeared in the way that Cue 1, if present, was the leftmost cue, Cue 2 was always spatially before Cue 3 and Cue 4, but after Cue 1, and Cue 3 always preceded Cue 4.

A feedback display followed each prediction. The feedback was an 83 × 86 pixel icon of a sun or of a cloud with rain. The icon appeared in the horizontal center line, 343 pixels from the top.

Design

SRT task. The task for participants was to press the button that corresponds to the location of the black circle. The response buttons were arranged horizontally in the bottom line of the keyboard with one button between each. The response buttons were Y, C, B, and M (Hungarian standard QWERTZ keyboards were used). Each target item was on screen until one of the response keys was pressed. Participants rested fingers on the keys so that each key corresponded to a different finger. In the case of an incorrect response, a 560-ms tone was played. The response–stimulus interval was 250 ms.

There were 12 blocks in the task; each block was made up of 60 stimulus presentations. In Blocks 1 to 11, a 12-element-long sequence determined the appearance of the target stimulus (121423413243; numbers represent the position of the black circle). In Block 12, stimuli appeared in pseudorandom order, with the constraint that the target stimulus cannot appear in the same location twice in a row.

AGL task. The training phase consisted of 2 × 58 auditorily presented sentences. Children were asked to draw during the listening phase. At the end of the training session, participants were told that the sentences were in an unknown language. New instructions were given: Participants were asked to listen to pairs of sentences and decide which member of each pair was more similar to the previously heard language. The order of the pairs was random, but sentence pairs were preset. Participants had to respond verbally with “first” or “second.” Responses were recorded in E-prime by the experimenter.

WP task. The role of participants in the WP task is to guess the weather. They faced one, two, or three out of four different cues, and had to decide whether there would be sunshine or rain. They were asked to press ENTER for sunshine and SPACE for rain. There were four blocks with 50 items in each block. Items were presented in a preset pseudorandom order. The only constraint on the presentation order of items was that there were no two consecutive predictions using the same cues. That is, the combination of Cues 1 and 2 could not appear twice in a row, whereas this combination could be followed by, for example, the combination of Cues 1, 2, and 3.

Each cue had a preset predictive value. There were two cues predicting each outcome: Cues 1 and 2 predicted sunshine, and Cues 3 and 4 predicted rain. There was a strong and a weak cue for each outcome: Cues 1 and 4 predicted their outcome in 85.7% of their appearances, whereas the predictive value of Cues 2 and 3 was 70%. Table 2 summarizes the design.

Data Analysis

Due to technical problems, the data of one child with SLI was not registered in the SRT task. For another member of the SLI group, AGL results were lost, also due to technical problems. The

Table 2
Types and Occurrences of Cues or Cue Combinations per Blocks of 50 Trials

Cues	Frequency	P(SUN)
A	8	0.875
B	4	0.75
C	4	0.25
D	8	0.125
AB	8	0.875
AC	1	1
BC	2	0.5
BD	1	0
CD	8	0.125
ABC	2	1
ABD	1	1
ACD	1	0
BCD	2	0

Note. Column 1 (cues) identifies cues and combinations: A is Cue 1, B is Cue 2, C is Cue 3, and D is Cue 4. Column 2 (frequency) is the number of appearances within a block of 50 trials. Column 3 provides the probability that the given cue or combination leads to sunshine (which equals 1 minus the probability of rain).

data of these children were excluded task-wise; their data from other tasks were included in the analyses.

Three types of analyses were conducted on all three tasks. First, we compare overall group means of performance indexes on each of the three tasks. Second, participants were categorized as "learners," "nonlearners," or "other." The "other" group is made up of participants whose performance is difficult to interpret. These are participants who perform well below 45% in the AGL and WP tasks—that is, their performance is below chance level. In the SRT task, participants were categorized as "other" if their mean RT in the random block was lower than the last sequence block. Participants were categorized as nonlearners if their performance was interpretable as evidence for lack of learning. For the AGL and WP tasks, the criterion to be included among nonlearners was a chance-level performance between 45% and 55%. For nonlearners on the SRT task, the RT difference between the random and last sequence blocks was below 25 ms.[4] Participants with better performance than these were categorized as "learners." Their mean RTs on the random block were at least 25 ms higher than mean RTs on the last sequence blocks in the SRT. In the other two tasks, a participant was considered a learner if his or her performance was above 55%. After categorizing participants, chi-square tests were applied to determine whether the ratio of learners differed by group (SLI vs. TD).

In the third analysis, performance of "learners" was compared by group (SLI vs. TD). This analysis was conducted to test whether participants in the clinical group had a gradual impairment of IL, as suggested by the PDH (Ullman & Pierpont, 2005), or whether learners among children with SLI show a similar level of learning as TD learners.

Results

Comparison of Overall Group Means

A univariate ANOVA was conducted on results from the SRT task to test whether Block 12 (random) − Block 11 (sequential) raw RT differences, indexing sequence-specific learning, differ by group. The ANOVA revealed no difference between the groups ($p = .418$). As previous studies have pointed out the importance of differences in the baseline RTs between clinical and control groups, we applied individual z transformations: For each individual, mean and standard deviations were computed separately, then the response latency for each item was z transformed by subtracting the given participant's overall mean from the raw RT and dividing the value with the participant's standard deviation (Christ, White, Mandernach, & Keys, 2001). Next, the difference between the mean of z-transformed Block 11 (the last sequence block) RTs were extracted from the mean of z-transformed Block 12 (random block) RTs. This difference reflecting the size of sequence learning was compared by group, revealing a significant group main effect, $F(1, 113) = 5.888$, $p < .05$, $\eta_p^2 = 0.050$, with bigger learning effect in the control than in the SLI group.

A one-way ANOVA was also conducted to test whether there was an overall performance difference between the groups on the AGL task. The control group outperformed the clinical group, as revealed by a significant main effect of group, $F(1, 113) = 6.645$, $p < .05$, $\eta_p^2 = 0.056$.

We also analyzed data on the WP task with a 4 × 2 repeated measures ANOVA with block (Block 1 vs. Block 2 vs. Block 3 vs. Block 4) as a within-subject variable and group as a between-subject variable. The Huynh-Feldt corrected ANOVA revealed that neither the main effect of block ($p = .196$) nor the main effect of group ($p = .814$) was significant. The Block × Group interaction approached, but did not reach, significance, $F(2.502, 285.197) = 2.302$, $p = .089$.

Comparison of Ratio and Performance of Learners

For each task, Table 3 provides the percentage of participants in each category by group. In the SRT task, a chi-square test was applied on the number of participants in each learning-category (learners vs. nonlearners vs. other) by group (SLI vs. TD). Results revealed that the distribution of category membership differed by group, $\chi^2(N = 115, df = 2) = 19.083$, $p < .001$, with a significantly higher proportion of learners in the TD group.

RT differences between the z-transformed RTs of Block 12 and Block 11, reflecting the size of sequence learning, were compared by group for learners. The ANOVA revealed that the main effect of group was not significant, $F(1, 79) = 1.308$, $p = .256$, $\eta_p^2 = 0.017$.

Testing the distribution of category membership by group (SLI vs. TD) in the AGL task revealed a significant difference, $\chi^2(N = 115, df = 2) = 8.319$, $p < .05$. Results again show that the proportion of TD learners is higher than the proportion of SLI learners in the AGL task. In the second analysis, the accuracy rates of learners were compared by group. The ANOVA revealed that the difference between the two groups was not significant, $F(1, 54) = 1.290$, $p = .261$, $\eta_p^2 = 0.023$. Z-transformed performance of learners on the three tasks by group is shown in Figure 1.

In the WP task, the distributions of category memberships were not different in the SLI and TD groups, as indicated by a nonsignificant χ^2 test, $\chi^2(N = 116, df = 2) = 2.000$, $p = .368$. As in the other two tasks, performance of the learning participants was compared by group. The ANOVA revealed no significant difference, $F(1, 56) = 0.922$, $p = .341$, $\eta_p^2 = 0.015$.

To test relationships between different forms of skill learning, participants' performance on the three tasks was correlated. Results revealed no significant correlations, all $ps > 0.407$. Controlling for age did not change this pattern (all $ps > 0.580$); neither did separate analysis for the SLI (all $ps > 0.162$) and control groups (all $ps > 0.462$). Language measures (performance on the screening tests) were only available for the SLI group. These tests were the TROG, sentence repetition, nonword repetition, and PPVT tasks. The only significant correlation was a negative one between PPVT scores and performance on the AGL task ($r = -0.436$, $N = 28$, $p < .05$). No other significant correlations were observed. Controlling for age did not affect this pattern (PPVT-AGL: $r = -0.580$, $df = 25$, $p < .01$).

[4] This arbitrary criterion for learning on the SRT task was set on the basis of our previous developmental study showing that the mean RT difference between Blocks 11 and 12 under 11 years is 47 ms (Lukács & Kemény, in press). Based on this result, we decided to use 25 ms as the arbitrary criterion, as it is high enough to say that participants above this difference are true learners, and low enough not to categorize too many learning participants as nonlearners.

Table 3
The Percentage and Number of Participants in Each Category by Group in the Three Tasks

		%			*n*		
Task	Group	Learners	Nonlearners	Other	Learners	Nonlearners	Other
SRT task	SLI	32.14%	25.00%	42.86%	9	7	12
	CTRL	77.01%	9.20%	13.79%	67	8	12
WP task	SLI	48.28%	41.38%	10.34%	14	12	3
	CTRL	57.47%	27.59%	14.94%	50	24	13
AGL task	SLI	25.00%	53.57%	21.43%	7	15	6
	CTRL	56.32%	31.03%	12.64%	49	27	11

Discussion

This study was based on the premise that IL is vulnerable in SLI. By using three different IL tasks, it provided an opportunity to test whether this vulnerability (a) affects the *extraction of statistical information* (Hsu & Bishop, 2011), (b) is present as a general deficit in *procedural learning* (Ullman & Pierpont, 2005), or (c) primarily affects performance on tasks with *sequentially organized* stimuli (Conway et al., 2010; Conway, Pisoni, & Kronenberger, 2009; Misyak & Christiansen, 2012). We found deficits in the SLI group in the domain of sequence learning both for motor sequences on the SRT task and for auditory sequences on the AGL tasks, but no deficits were observed on the nonsequential probabilistic categorization task. Taken together, findings of a sequential deficit in both the verbal and nonverbal domains show that domain-general processes of IL tend to be vulnerable in SLI, and this deficit primarily affects performance on tasks with sequentially organized stimuli.

Keeping in mind that even TD children tend to differ greatly in their individual learning abilities, and that this variance is even more accentuated in SLI, we took a novel approach in analyzing the data. Besides comparing overall group performance levels, we compared the proportion of learners in the SLI and TD groups, and then contrasted the level of learning only for learners as well. This method allowed us to get a more precise picture of impairments in IL in SLI and to select between different possible scenarios behind an overall SLI deficit at the group level. We wanted to see if such a deficit, when present, is caused by (a) a general impairment affecting all children in the SLI group, manifest as a lower level of learning at the individual level; (b) a smaller proportion of learning children in the SLI group, with the same level of individual learning in the SLI and TD groups; or (c) a smaller proportion of learning children in the SLI group, also with a lower level of individual learning.

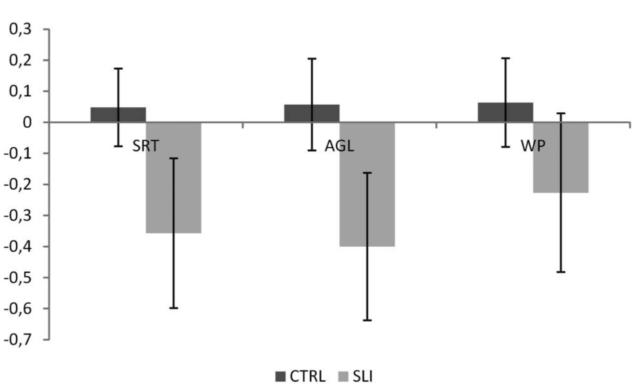


Figure 1. Z-transformed performance of learners on the SRT, AGL, and WP tasks by group. Error bars indicate the standard error.

In the two implicit sequence learning tasks—an SRT task testing the learning of specific motor sequences and an AGL task testing the extraction of regularities from auditory sequences of verbal stimuli—a significantly smaller proportion of children showed evidence of learning in the SLI than in the TD group (32.14% learners on the SRT, 25% learners on the AGL task in the SLI group; 77.01% learners on the SRT, 56.32% learners on the AGL task in the TD group), whereas the proportion of learners on the nonsequential WP task requiring probabilistic categorization did not differ. The amount of learning for those children with SLI who did show evidence of learning was overall comparable with the level of learning TD children on all three tasks (with great individual variation). These similar performance levels across groups are not explained by a uniform level of learning in every learner: There was great individual variation even among learners in both groups for all tasks. Finding evidence of deficits in sequence learning for some, but not for all, children with SLI, together with similar levels of performance among learners in the two groups, makes implicit sequence learning a good candidate for establishing a subgroup of children with SLI.

In line with previous studies on motor sequence learning (Lum et al., 2012; Tomblin et al., 2007), we found a larger ratio of nonlearners on the SRT task in the SLI than in the TD group. This group difference was evident even though, instead of testing probabilistic motor sequences, we used a deterministic version of the SRT task. Finding difficulties with motor sequences based on visuospatial information in itself contradicts language-specific accounts of SLI and suggests that children with SLI often have a problem in nonverbal domains of learning as well.

Our results from AGL also have precedents in studies of verbal sequence learning. Although these earlier studies used different paradigms and had a different focus, they demonstrate that verbal sequence learning poses difficulties for children both in a task that models the extraction of word boundaries from the speech stream (Evans et al., 2009) and when they have to extract long-distance dependencies between specific items (Grunow et al., 2006; Plante et al., 2002). Based on these findings, it is not surprising that we also found deficient learning in SLI in a more complex task which requires children to extract regularities from auditory sequences, and this way, models grammar acquisition, in which transitional probabilities show a more complex pattern and sequences are defined at the level of categories instead of items.

Although these results parallel the majority of earlier results, showing that procedural/IL deficits in SLI extend to domains outside language (Hsu & Bishop, 2011; Ullman & Pierpont, 2005), some of the theoretical proposals on IL do not find support in our findings. First, as reviewed in the introduction, Hsu and Bishop claim that the central deficit in SLI involves an impairment in extracting *nonadjacent* dependencies (due to which children with SLI fail to develop abstract grammatical representations; instead, they rely on item-based ones). In line with the evidence reviewed in the introduction (also by Hsu and Bishop) showing results contradicting this claim (findings of Evans et al., 2009; Lum et al., 2012; Tomblin et al., 2007 show that children with SLI do not only have problems with nonadjacent dependencies), we argue that children with SLI also have problems with adjacent dependencies, and even with deterministic sequences. As Dabrowska (2012) also points out, bigger problems with nonadjacent dependencies can stem from task difficulty and taxing working memory more. Second, although SLI does seem to involve a consolidation problem, as shown by Hedenius et al. (2011), most other studies suggest that the problem is already evident at the earlier stage of acquiring the representation through IL. In this regard, our results also confirm an earlier problem with sequence learning, which could be associated with a later consolidation problem, but this has to be confirmed by future research.

We also extended previous studies on IL in SLI by including a task that focuses on the acquisition of nonsequential information. Although SLI deficits were apparent in sequential learning tasks, we did not find evidence for problems with IL when stimuli are not sequentially organized: There were no group differences in probabilistic categorization on the WP task. This contrasts with our earlier results on the same task (Kemény & Lukács, 2010). Although the current pattern of results clearly argues for greater problems in the sequential domain, the lack of a group difference for the WP task is not conclusive evidence for intact nonsequential IL in SLI. We can only speculate on the reason for the discrepancy between the two studies. Children in our earlier study were an average of 2 years older in both groups; it is possible that the deficit of the SLI group only becomes evident at a later age, when WP performance develops further and at a larger rate, but performance in SLI, at least for some children, could be stuck at this lower level. Another possible explanation is the great individual variation for SLI children (and even for TD children at this age): The ratio of learners versus nonlearners can change through different samples. Nonsequential implicit statistical learning can be vulnerable in SLI, but as our findings suggest, to a lesser extent than sequence learning; detecting this smaller deficit might be more sensitive to individual variation and changes in samples. This variation is not surprising, given the heterogeneity of children with SLI in both the pattern and the severity of the symptoms, as well as in the etiology of the language impairment (see, e.g., Bishop, 2006; Fattal, Friedmann, & Fattal-Valevski, 2011; Leonard, 1998), and can also explain contrasting findings in the earlier literature.

Finding impaired IL for some children with SLI, but not for others (as shown by distribution of learners and nonlearners on the three tasks), is probably also explained by the heterogeneity of SLI, and further supports the existence of impairments in different background mechanisms behind the language symptoms. A deficit in IL, especially of sequences, can be the problem for some children with SLI that may contribute or lead to language deficits. If a causal connection exists, we expect to find associations between IL and language domains. The PDH connects procedural learning abilities to grammar, and previous studies (Lum et al., 2012; Tomblin et al., 2007) found associations between motor sequence learning and grammar. As Evans and colleagues (2009) pointed out, an IL deficit can already contribute to language impairment at the lexical level, and may extend to other domains of language that build on the extraction of statistical information (Hsu & Bishop, 2011). It is one of the limitations of our study that we did not include complex language tasks that are good candidates for involving sequence learning (and we did not find meaningful correlations with the screening tests). Although sequence learning/IL/statistical learning is established as a domain-general skill, and a basic learning process involved in language acquisition, we do not yet have a clear picture of what aspects of language are affected by it impairments to such processes. An important direction for future research is set by that question and by the few studies that show an association between nonlinguistic (or linguistic) sequence learning and language acquisition/language processing.

References

Bates, E., & Goodman, J. C. (2001). On the inseparability of grammar and the lexicon: Evidence from acquisition. In M. Tomasello & E. Bates (Eds.), *Language development: The essential readings* (pp. 134–162). Oxford, UK: Blackwell.

Belton, E., Salmond, C., Watkins, K., Vargha-Khadem, F., & Gadian, D. (2002). Bilateral grey matter abnormalities in a family with a mutation in FOXP2. *Neuroimage, 16,* 10144.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, UK: Erlbaum.

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

Bishop, D. V. M. (1992). The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry, 33,* 3–66. doi:10.1111/j.1469-7610.1992.tb00858.x

Bishop, D. V. M. (2006). What causes specific language impairment in children? *Current Directions in Psychological Science, 15,* 217–221. doi:10.1111/j.1467-8721.2006.00439.x

Bishop, D. V. M. (2012). *TROG – Test for Reception of Grammar. Handbook*. Budapest, Hungary: OS Hungary.

Bishop, D. V. M., & Edmundson, A. (1987). Specific language impairment as a maturational lag: Evidence from longitudinal data on language and motor development. *Developmental Medicine & Child Neurology, 29,* 442–459. doi:10.1111/j.1469-8749.1987.tb02504.x

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology: I. Generativity. *Autism: The International Journal of Research and Practice, 9,* 7–27. doi:10.1177/1362361305049027

Christ, S. E., White, D. A., Mandernach, T., & Keys, B. A. (2001). Inhibitory control across the life span. *Developmental Neuropsychology, 20,* 653–669.

Christiansen, M. H., Conway, C., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes, 27,* 231–256. doi:10.1080/01690965.2011.606666

Christiansen, M. H., Kelly, M. L., Shillcock, R. C., & Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition, 116,* 382–393. doi:10.1016/j.cognition.2010.05.015

Clahsen, H., & Hansen, D. (1997). The grammatical agreement deficit in specific language impairment: Evidence from therapy experiments. In M. Gopnik (Ed.), *The inheritance and innateness of grammars* (pp. 141–160). Oxford, UK: Oxford University Press.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences, 2,* 406–416. doi:10.1016/S1364-6613(98)01232-7

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 114,* 356–371. doi:10.1016/j.cognition.2009.10.009

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31,* 24–39.

Conway, C. M., Karpicke, J., & Pisoni, D. B. (2007). Contribution of implicit sequence learning to spoken language processing: Some preliminary findings with normal-hearing adults. *Journal of Deaf Studies and Deaf Education, 12,* 317–334. doi:10.1093/deafed/enm019

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science, 18,* 275–279. doi:10.1111/j.1467-8721.2009.01651.x

Csányi, I. (1974). *Peabody szókincs-teszt* [Peabody vocabulary test]. Budapest, Hungary: Bárczi Gusztáv Gyógypedagógiai Főiskola.

Dąbrowska, E. (2010). Productivity, proceduralisation and SLI: Comment on Hsu and Bishop. *Human Development, 53,* 276–284.

Daselaar, S. M., Rombouts, S. A., Veltman, D. J., Raaijmakers, J. G., & Jonker, C. (2003). Similar network activated by young and old adults during the acquisition of a motor sequence. *Neurobiology of Aging, 24,* 1013–1019. doi:10.1016/S0197-4580(03)00030-7

Dromi, E., Leonard, L. B., Adam, G., & Zadoneisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1414–1431.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52,* 321–335. doi:10.1044/1092-4388(2009/07-0189)

Fattal, I., Friedmann, N., & Fattal-Valevski, A. (2011). The crucial role of thiamine in the development of syntax and lexical retrieval: A study of infantile thiamine deficiency. *Brain: A Journal of Neurology, 134,* 1720–1739. doi:10.1093/brain/awr068

Frensch, P. A., & Ruenger, D. (2003). Implicit learning. *Current Directions in Psychological Science, 12,* 13–18. doi:10.1111/1467-8721.01213

Gabriel, A., Maillart, C., Guillaume, M., Stefaniak, N., & Meulemans, T. (2011). Exploration of serial structure procedural learning in children with language impairment. *Journal of the International Neuropsychological Society, 17,* 336–343. doi:10.1017/S1355617710001724

Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013). Procedural learning in specific language impairment: Effects of sequence complexity. *Journal of the International Neuropsychological Society, 19,* 264–271. doi:10.1017/S1355617712001270

Gathercole, S., & Baddeley, A. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29,* 336–360. doi:10.1016/0749-596X(90)90004-J

Gathercole, S., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.

Gauger, L. M., Lombardino, L., & Leonard, C. (1997). Brain morphology in children with specific Language impairment. *Journal of Speech & Hearing Research, 40,* 1272–1284.

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70,* 109–135. doi:10.1016/S0010-0277(99)00003-7

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4,* 178–186. doi:10.1016/S1364-6613(00)01467-4

Gopnik, M., & Crago, M. B. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39,* 1–50. doi:10.1016/0010-0277(91)90058-C

Graf Estes, K., Evans, J., & Else-Quest, N. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50,* 177–195. doi:10.1044/1092-4388(2007/015)

Grunow, H., Spaulding, T. J., Gómez, R. L., & Plante, E. (2006). The effects of variation on learning word order rules by adults with and without language-based learning disabilities. *Journal of Communication Disorders, 39,* 158–170. doi:10.1016/j.jcomdis.2005.11.004

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., . . . Ullman, M. T. (2011). Grammar predicts procedural learning and consolidation deficits in children with specific language impairment. *Research in Developmental Disabilities, 32,* 2362–2375. doi:10.1016/j.ridd.2011.07.026

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry, 53,* 37–45. doi:10.1111/j.1469-7610.2011.02430.x

Hill, E. L. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders, 36,* 149–171. doi:10.1080/13682820010019874

Hsu, H. J., & Bishop, D. V. (2011). Grammatical difficulties in children with specific language impairment: Is learning deficient? *Human Development, 53,* 264–277. doi:10.1159/000321289

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77,* 1822–1841. doi:10.1111/j.1467-8624.2006.00976.x

Joanisse, M. F., & Seidenberg, M. S. (1998). Specific language impairment in children: An impairment in grammar or processing? *Trends in Cognitive Sciences, 2,* 240–247. doi:10.1016/S1364-6613(98)01186-3

Jones, M., & Conti-Ramsden, G. (1997). A comparison of verb use in children with specific language impairment and their younger siblings. *Journal of Speech & Hearing Research, 40,* 1298–1313.

Kas, B., & Lukács, Á. (2011). *Magyar Mondatutánmondási Teszt* [Hungarian sentence repetition test]. Manuscript in preparation.

Kemény, F., & Lukács, Á. (2010). Impaired procedural learning in language impairment: Results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology, 32,* 249–258. doi:10.1080/13803390902971131

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology, 48,* 171–184. doi:10.1037/a0025405

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399–1402. doi:10.1126/science.273.5280.1399

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory, 1,* 106–120.

Leonard, L. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10,* 179–202.

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

Leonard, L. B., Ellis Weismer, S., Miller, C., Francis, D., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50,* 408–428. doi:10.1044/1092-4388(2007/029)

Leonard, L. B., Eyer, J., Bedore, L., & Grela, B. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40,* 741–753.

Leonard, L. B., McGregor, K., & Allen, G. (1992). Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech & Hearing Research, 35,* 1076–1085.

Leonard, L. B., Sabbadini, L., Leonard, J., & Volterra, V. (1987). Specific language impairment in children: A cross-linguistic study. *Brain and Language, 32,* 233–252. doi:10.1016/0093-934X(87)90126-X

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

Lukács, Á., & Kemény, F. (in press). Development of different forms of skill learning throughout the lifespan. *Cognitive Science*.

Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 48,* 1138–1154. doi:10.1016/j.cortex.2011.06.001

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development – A test of the critical mass hypothesis. *Journal of Child Language, 21,* 339–366. doi:10.1017/S0305000900009302

Marchman, V. A., Wulfeck, B., & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research, 42,* 206–219.

Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology, 69,* 199–221. doi:10.1006/jecp.1998.2442

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62,* 302–331. doi:10.1111/j.1467-9922.2010.00626.x

Montgomery, J. (2000). Relation of working memory to offline and real-time sentence processing in children with specific language impairment. *Applied Psycholinguistics, 21,* 117–148. doi:10.1017/S0142716400001065

Montgomery, J. W. (2002). Information processing and language comprehension in children with specific language impairment. *Topics in Language Disorders, 22,* 62–84. doi:10.1097/00011363-200205000-00007

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Oetting, J. B., & Horohov, J. (1997). Past tense marking in children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 40,* 62–74.

Oetting, J. B., & Rice, M. (1993). Plural acquisition in children with specific language impairment. *Journal of Speech & Hearing Research, 36,* 1236–1248.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Plante, E., Gómez, R. L., & Gerken, L. (2002). Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders, 35,* 453–462. doi:10.1016/S0021-9924(02)00094-1

Plante, E., Swisher, L., Vance, R., & Rapcsak, S. (1991). MRI findings in boys with specific language impairment. *Brain and Language, 41,* 52–66. doi:10.1016/0093-934X(91)90110-M

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13,* 564–574. doi:10.1037/0894-4105.13.4.564

Powell, R. P., & Bishop, D. V. M. (1992). Clumsiness and perceptual problems in children with specific language impairment. *Developmental Medicine & Child Neurology, 34,* 755–765. doi:10.1111/j.1469-8749.1992.tb11514.x

Price, A. L. (2009). Distinguishing the contributions of implicit and explicit processes to performance of the weather prediction task. *Memory & Cognition, 37,* 210–222. doi:10.3758/MC.37.2.210

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai [Hungarian methods for studying verbal working memory]. *Magyar Pszichológiai Szemle, 60,* 479–505. doi:10.1556/MPSzle.60.2005.4.3

Raven, J., Court, J., & Raven, J. (1987). *Raven's progressive matrices and Raven's colored matrices*. London. UK: H. K. Lewis.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior, 6,* 855–863. doi:10.1016/S0022-5371(67)80149-X

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York, NY: Oxford University Press.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech & Hearing Research, 38,* 850–863.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1412–1431.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47,* 172–196. doi:10.1006/jmla.2001.2839

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928. doi:10.1126/science.274.5294.1926

Shohamy, D., Myers, C. E., Onlaor, S., & Gluck, M. A. (2004). Role of the basal ganglia in category learning: How do patients with Parkinson's disease learn? *Behavioral Neuroscience, 118,* 676–686. doi:10.1037/0735-7044.118.4.676

Siegert, R. J., Taylor, K. D., Weatherall, M., & Abernethy, D. A. (2006). Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology, 20,* 490–495. doi:10.1037/0894-4105.20.4.490

Skipp, A., Windfuhr, K., & Conti-Ramsden, G. (2002). Children's grammatical categories of verb and noun: A comparative look at children with specific language impairment (SLI) and normal language (NL). *International Journal of Language & Communication Disorders, 37,* 253–271. doi:10.1080/13682820110119214

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., & Reber, P. J. (2002). Neural correlates of artificial grammar learning. *Neuroimage, 17,* 1306–1314. doi:10.1006/nimg.2002.1291

Smith, J., Siegert, R. J., & McDowall, J. (2001). Preserved implicit learning on both the serial reaction time task and artificial grammar in patients with Parkinson's disease. *Brain and Cognition, 45,* 378–391. doi:10.1006/brcg.2001.1286

Stokes, S. F., & Fletcher, P. (2000). Lexical diversity and productivity in Cantonese-speaking children with specific language impairment. *International Journal of Language & Communication Disorders, 35,* 527–541. doi:10.1080/136828200750001278

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., . . . Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271,* 81–84. doi:10.1126/science.271.5245.81

Tallal, P., & Piercy, M. (1973). Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality. *Neuropsychologia, 11,* 389–398. doi:10.1016/0028-3932(73)90025-0

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4,* 156–163. doi:10.1016/S1364-6613(00)01462-5

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development, 3,* 269–293. doi:10.1080/15475440701377477

Ullman, M. T., & Gopnik, M. (1999). The production of inflectional morphology in hereditary specific language impairment. *Applied Psycholinguistics, 20,* 51–117. doi:10.1017/S0142716499001034

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex: A Journal Devoted to the Study of the Nervous System, 41,* 399–433. doi:10.1016/S0010-9452(08)70276-4

van der Lely, H. K. J. (1997). Language and cognitive development in a grammatical SLI boy: Modularity and innateness. *Journal of Neurolinguistics, 10,* 75–107. doi:10.1016/S0911-6044(97)00011-0

van der Lely, H., & Stollwerck, K. (1996). A grammatical specific language impairment in children. *Brain and Language, 52,* 484–504. doi:10.1006/brln.1996.0026

van der Lely, H. K. J., & Ullman, M. (2001). Past tense morphology in specific language impaired and normally developing children. *Language and Cognitive Processes, 16,* 177–217. doi:10.1080/01690960042000076

Vargha-Khadem, F., Watkins, K. E., Price, C. J., Ashburner, J., Alcock, K. J., Connelly, A., . . . Passingham, R. E. (1998). Neural basis of an inherited speech and language disorder. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 95,* 12695–12700. doi:10.1073/pnas.95.21.12695

Witt, K., Nühsman, A., & Deuschl, G. (2002). Intact artificial grammar learning in patients with cerebellar degeneration and advanced Parkinson's disease. *Neuropsychologia, 40,* 1534–1540. doi:10.1016/S0028-3932(02)00027-1

Zelaznik, H. N., & Goffman, L. (2010). Generalized motor abilities and timing behavior in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53,* 383–393. doi:10.1044/1092-4388(2009/08-0204)

Received May 22, 2013
Revision received November 21, 2013
Accepted November 21, 2013 ■

Study 9.

Lukács Á., Kemény F., (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science*, 39(2) 383-404. doi: 10.1111/cogs.12143

Thesis 10. Age-related changes in different forms of skill learning with potential roles in language acquisition argue against the existence of a critical period for these learning mechamisms (9)

dc_1365_16

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2014) 1–21

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12143

Development of Different Forms of Skill Learning Throughout the Lifespan

Ágnes Lukács,[a] Ferenc Kemény[a,b]

[a]*Department of Cognitive Science, Budapest University of Technology and Economics*
[b]*Research Institute for Linguistics, Hungarian Academy of Sciences*

Received 5 March 2013; received in revised form 16 October 2013; accepted 13 November 2013

Abstract

The acquisition of complex motor, cognitive, and social skills, like playing a musical instrument or mastering sports or a language, is generally associated with implicit skill learning (SL). Although it is a general view that SL is most effective in childhood, and such skills are best acquired if learning starts early, this idea has rarely been tested by systematic empirical studies on the developmental pathways of SL from childhood to old age. In this paper, we challenge the view that childhood and early school years are the prime time for skill learning by tracking age-related changes in performance in three different paradigms of SL. We collected data from participants between 7 and 87 years for (1) a Serial Reaction Time Task (SRT) testing the learning of motor sequences, (2) an Artificial Grammar Learning (AGL) task testing the extraction of regularities from auditory sequences, and (3) Probabilistic Category Learning in the Weather Prediction task (WP), a non-sequential categorization task. Results on all three tasks show that adolescence and adulthood are the most efficient periods for skill learning, since instead of becoming less and less effective with age, SL improves from childhood into adulthood and then later declines with aging.

Keywords: Motor sequence learning; Artificial grammar learning; Probabilistic categorization; Development across the lifespan

1. Introduction

The experience of learning is fundamentally associated with childhood and young adulthood. These are the years that, supported by society's institutions and educational system, are devoted to accumulating knowledge and to acquiring skills that prepare us for our adult life. It is a widely accepted view that learning is most effective at these ages

Correspondence should be sent to Ágnes Lukács, BME Department of Cognitive Science, Egry J. u, 1, H-1111 Budapest, Hungary. E-mail: alukacs@cogsci.bme.hu

and that skills like playing a musical instrument, mastering sports, or learning a language are best acquired if learning starts early in childhood with a lot of practice. Yet it is definitely possible to learn language, music, to drive a car, engage in sport or dancing, and many other new activities requiring complex skills in adulthood, although generally not with an aspiration to excel at something, just for the pleasure of learning and practicing the skill. There are examples of late bloomers and late starters even among the best: Joseph Conrad did not start to learn English until the age of 21, and he first published his writings at 37; Raymond Chandler's first short story appeared when he was 45; Leonard Cohen learned to play the guitar in his teens, and his first album came out when he was 32. Lifelong learning is something that gains more and more attention in everyday life and in research as well. People of all ages go to dancing schools, do yoga, learn to cook or sew, develop a new hobby, or learn languages for fun or because they move to a new country. Recently, NYU psychologist Gary Marcus (2012) devoted his sabbatical to learning to play the guitar to demonstrate adult plasticity of the brain. It seems that skill learning does take place at all ages, but relatively little is known about differences in its effectiveness as a function of age.

Children learn increasingly more complex materials in and outside class as they grow older, supported by multiple learning systems. The learning of complex motor, cognitive, and social skills is assumed to be at least partly associated with a set of mechanisms usually referred to as implicit learning[1] (IL; e.g., Reber, 1967, 1993; Cleeremans, Destrebecqz, & Boyer, 1998; Frensch & Rünger, 2003). As will be clear from results reviewed below, there are more and more data on the developmental pathways of such skills, but we are far from fully understanding age-related differences in complex skill learning across the lifespan. This paper aims to take a step in that direction by examining skill learning in three different traditional paradigms of skill learning (SL) between 7 and 87 years. We collected data from all participants using (a) a serial reaction time task (SRT) testing the learning of motor sequences, (b) an artificial grammar learning (AGL) task testing the extraction of regularities from auditory sequences, and (c) probabilistic category learning in the weather prediction task (WP), a non-sequential categorization task. Besides learning about developmental differences in these different forms of SL, plotting development on performance on all three tasks provides an opportunity to see whether they show similar developmental trends. If developmental trends diverge, that can be taken as an indication that these tasks tap into different forms of SL.

The three tasks were chosen because they are all known as traditional paradigms of SL. Taking that as our starting point, we also have to point out in the beginning that they differ in several important respects: They present stimuli in different modalities, stimuli of different complexity and organization, and they vary in their working memory load and even in the degree of involvement of explicit learning processes. In this study, we did not intend to test awareness, complexity, or working memory load differences between the tasks. Although these differences are important confounds that make it hard to determine if the age-related differences we observe are due to differences in IL, in explicit learning, or in working memory, since everyday SL also relies on a mixture of these processes, these tasks may serve as good models of such learning processes. We regard tracking developmental differences in these tasks as an important contribution to our knowledge about changes in the effectiveness of SL.

1.1. Models of age-related changes in skill learning

A considerable body of research has accumulated in the past two decades on age-related differences in implicit functions, yielding three models of implicit SL. The developmental invariance model argues for age independence of SL based on results that (a) do not show differences in the rate of learning between children and adults (Meulemans, Van Der Linden, & Perruchet, 1998), (b) demonstrate the existence of the same learning abilities in infants already (Gomez & Gerken, 1999; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; see also Krogh, Vlach, & Johnson, 2013), (c) show robustness of implicit SL even in neurological impairments like amnesia (Cohen & Squire, 1980), and (d) show that SL is associated with brain regions (basal ganglia and cerebellum) that mature early and are evolutionarily older (Reber, 1992, 1993). Other models argue for age-related changes in SL. One of them argues for improvement throughout childhood to young adulthood, stable performance in adulthood, and decline with aging.[2] This model is based on results showing that (a) older children and adults demonstrate better learning than younger children, and the elderly show poorer performance than younger adults (e.g., Fletcher, Maybery, & Bennett, 2000; Howard & Howard, 1997; Howard, Howard, Dennis, & Yankovith, 2007; Maybery, Taylor, & O'Brien-Malone, 1995; Thomas et al., 2004), and (b) fronto-striatal regions (among others, see below) play an important role in SL and these regions go through considerable development well into adolescence, and also show significant changes in aging. A third model (e.g., Janacsek, Fiser, & Németh, 2012) is a parallel of models claiming the existence of critical periods for language acquisition (an example of SL) and argues that there is a period from birth to adolescence, when SL is especially effective, which is the best time to acquire new skill. After adolescence, SL is less effective but still takes place until around 60 years, after which it declines considerably.

1.2. Empirical findings on skill learning throughout the lifespan

Empirical results on the developmental stability of IL are controversial. Meulemans et al. (1998) used alternating repeating and random sequences of stimuli in a SRT task and observed the same rate of sequence learning (in reaction times [RTs]) in 6–7-year-olds, 10–11 year-olds, and adults. Using a complex design of contingency learning, Amso and Davidow (2012) also argue for age-independent statistical learning. Maybery et al. (1995), on the other hand, found age-related differences in both implicit and explicit learning of covariation rules on the location of objects in a matrix: 10–12-year-olds showed better explicit learning but also more effective IL evidenced by better prediction of locations than 5–7-year-olds. Fletcher et al. (2000) supplemented these findings with results from gifted children and children with intellectual disability showing that mental age, that is, developmental level, has a critical influence on IL, contradicting proposals for IQ and age independence of IL by Reber (1992).

Age-dependent differences have also been attested in an fMRI study comparing implicit sequence learning in adults versus 7–11-year-old children on the SRT task (Thomas et al., 2004). Adults showed better sequence learning (bigger difference between random-sequence RTs, computed on z-scores) than children, who also needed a longer period of exposure for sequence learning to take effect at all. The bigger sequence learning effect in adults was associated with differential brain activation patterns.

In a recent paper, Weiermann and Meier (2012) compared sequence learning across the lifespan in a task sequence learning task sampling three age groups. All groups showed a sequence learning effect of the same magnitude, but for children (7–16 years) and adults above 65, only participants with explicit knowledge (measured by correct number of elements in a sequence generation task) of the sequence showed significant learning. In young adults, sequence learning was not tied to explicit knowledge. In the authors' view, this reflects less effective IL abilities in children and older adults, which are compensated by explicit learning in both groups, at least for the complex task with an abstract sequence used in this study.

Huang-Pollock, Maddox, and Karalunas (2011) compared implicit and explicit category learning in college students (18–25 years) and school-age children (8–12 years) using two information–integration paradigms, and the same approach was used to compare performance of older and younger adults (Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010). Young adults outperformed children and older adults; strategy analysis revealed that children and older adults tended to use explicit one-dimensional rule-based strategies, and they more often relied on an irrelevant dimension than adults. The authors argue (here and elsewhere as well in their COVIS model: Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & O'Brien, 2005) that different forms of categorization rely on different, though overlapping memory systems, and, that beyond the maturity and integrity of the implicit system, categorization performance in an information–integration paradigm is also a function of the effectiveness of inhibition of output from the explicit system, in which young adults are better.

Howard and Howard (2012) dissociate two different forms of IL (probabilistic) sequence learning (ASRT task) and spatial context learning (spatial contextual cueing task, SCCT) in aging and in several neuropsychological conditions. As described by Howard and colleagues, both tasks rely on associative and incidental learning, and both are implicit, but spatial contextual cueing is resistant to aging (although it does show improvement from childhood to young adulthood, reflecting the maturation of the medial temporal lobe; see Vaidya, Huger, Howard, & Howard, 2007), while sequence learning on the ASRT task becomes less effective with age (Howard and Howard 1997, 2012; Howard et al., 2007). There are several differences between the tasks accounting for the diverging pattern of age-related differences: The ASRT requires the integration of information across events distributed over time, while the SCCT does not. The SCCT is deterministic, while the ASRT requires the learning of probabilistic information. Howard et al. also cite imaging evidence showing that the two tasks rely on different neural structures: Gradual learning of probabilistic information is the function of the striatal system, which declines in function and structure after young adulthood (although MTL might also have a role in the early phases); summarizing findings from neuroimaging, Howard and Howard argue that IL is not a single unitary process, and different developmental patterns on the two tasks in aging reflect changes in their underlying neuroanatomical structures.

Similar learning processes have also been shown to operate in infants already. Most infant studies involve artificial language learning (Gomez & Gerken, 1999; Saffran et al., 1997). Such studies show that 8-month-olds are sensitive to transitional probabilities between items and are able to segment sequences of syllables and non-linguistic stimuli into "words" as well (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999). Infants as young as 12-month-olds also demonstrate learning of regularities when exposed to sequences of auditory non-sense syllables strung together according to the rules of a finite state grammar. Studies of statistical learning or AGL also found age-related differences related to the complexity of the grammar even in infants. Gómez and Maye (2005) found that although 15-month-old children were able to learn non-adjacent dependencies, these were too complex for 12-month-olds. Although there are many studies focusing on infants, far fewer results are available for older children. One notable exception is Saffran (2002) examining AGL in children between 7 and 9 years. After a 21-min exposure, children did show an effect of learning, but their learning performance was poorer than that of adults. Vinter, Pacton, Witt, and Perruchet (2010), on the other hand, cite evidence quoted by Reber (1993) from Roter (1985) showing the same level of learning in an artificial grammar task in 6–7, 9–11, and 12–15-year-olds.

The effect of aging on SL is also controversial. Salthouse, McGuthry, and Hambrick (1999) comparing young (18–39 years), middle-aged (40–59 years), and older (60–87 years) participants observed a small negative age effect on an SRT task. Some observe the same level of learning in the elderly as in younger adults (Howard & Howard, 1989, 1992). Gaillard, Destrebecqz, and Cleeremans (2009) also did not find differences in the amount of sequence learning and generation performance between 22, 45, and 71 years of age on the SRT task (although recognition performance was poorer in the older groups). Other studies, mainly those with more complex tasks with greater attention load in a dual task paradigm (Frensch & Miner, 1994; Nejati, Farshi, Ashayeri, & Aghdasi, 2008) or applying higher order dependencies in the sequences (as reviewed above: Howard & Howard, 1997; Howard et al., 2007), observe a weaker learning performance in the elderly. To summarize, information complexity charging attention or memory resources can affect learning, as well as the extent to which explicit processes can influence learning.

Although as the literature reviewed above shows, there have been studies comparing SL in children and adults, or even in children in different age groups, developmental studies of learning rarely focus on comprehensive sampling of all ages, and thus they are limited in the scope of their conclusions. In fact, there is only one study testing the development of implicit sequence learning systematically across the lifespan. Janacsek et al. (2012) was the first study aiming to examine differences in mechanisms behind complex

SL, and, more specifically in IL on the ASRT task, across the life span, between 4 and 85 years. Janacsek et al. found that although all age groups showed evidence of sequence learning, there is a significant change in SL abilities in early adolescence. When measured on raw RT differences, sequence learning was significantly more effective in the 4–12-year-old groups than any older group; then it decreased and stayed stable from 14 to 59 years, and became even less efficient in 60–85-year-olds. In the analysis on accuracy measures, only groups older than 6 showed learning, in groups between 7 and 44, the magnitude of learning was the same, and it was smaller (but present) again in the two older groups. When RTs were z-transformed, the developmental trend of sequence learning followed an inverted U-curve increasing from 4 to 12 years, staying at the same level until the mid-thirties, and then declining into old age. This difference between trends shown by raw versus z-transformed RTs suggests that raw RTs reflect developmental differences in absolute RTs (children start with the task with large RTs and have a larger range to reduce RTs in both motor and sequence learning). In spite of this possibility, the authors argue that SL is most effective before adolescence.

This study examines differences in SL throughout the lifespan with the aim of comparing different models of SL against the results. We tested 480 individuals on three different SL tasks, the Serial Reaction Time task testing the learning of motor sequences, the AGL task relying on the extraction of abstract regularities from auditory sequences, and the WP task examining non-sequential probabilistic categorization, to see whether SL is (a) stable with development, as the invariance hypothesis suggests; (b) increases till adulthood and decreases into older age, as the majority of results show; or (c) has a dedicated period before adolescence and decreases with age from childhood through adulthood into older age, as Janacsek et al. (2012) argue. With the help of three different tasks, and by sampling participants in narrow age groups across a wide age range we hope to extend the previous findings on the development of SL and to get a better picture of age-related differences in effectiveness and to test the hypothesis presented in Janacsek et al. (2012) that childhood and early school years are the prime time for SL by tracking age-related changes in performance in three different paradigms of SL. The three-task design also allows us to see whether age-related differences follow similar patterns in different aspects of complex SL.

Table 1
Demographic data of participants by age group

Age group (years)	7 9	9 11	11 14	14 18	18 25	25 35	35 45	45 55	55 65	65+	Total
Age group codes	U09	U11	U14	U18	U25	U35	U45	U55	U65	U99	
Mean age	7.9	9.8	11.9	15.5	20.6	29.4	40.8	49.8	60.0	72.1	31.8
N	64	63	63	57	37	37	28	45	43	43	480
Female	32	29	26	31	23	20	19	25	29	29	263
Male	32	34	37	26	14	17	9	20	14	14	217

Note. Age group codes are assigned to make presentation of results easier, “U” stands for “under” (but above previous age group).

2. Method

2.1. Participants

Altogether 480 participants were tested. The age of participants varied between 7 and 87 years of age. All participants above 18, and parents of participants under 18 provided a written informed consent in accordance with the principles set out in the Declaration of Helsinki and the stipulations of the local Institutional Review Board. Children or adults with known neurological or cognitive deficits were not included in the study. Participants were clustered into 10 age groups. The number of participants in each age group with gender distributions and means and standard deviations of ages are provided in Table 1. All participants of all age groups were exposed to the same procedure, which included three different tasks.

2.1.1. Design and stimuli

All participants completed three different tasks, the SRT task, the AGL task, and the WP task. All three tasks were presented on a 640 × 480 display, on a computer using E-prime 1.2 (Psychology Software Tools Inc., Pittsburgh, PA, USA).

2.1.1.1. The SRT task: The SRT task was an adaptation of Nissen and Bullemer (1987). Participants saw four circles (diameter approximately 55 pixels). Out of the four circles one was always black, while the other three were white with a black contour. The circles were arranged horizontally in the vertical center line with equal distances between the circles.

Participants were asked to press the button that corresponds to the location of the black circle. The response buttons were Y, C, B, and M (on Hungarian keyboards, these buttons are arranged horizontally in the bottom line of the keyboard with one button between the neighboring response keys). The target item was on screen until one of the response keys was pressed. If the answer was incorrect, a short, 560 ms tone indicated the presence of the error. The response–stimulus interval was set to 250 ms.

The task consisted of 12 blocks. Each block included 60 stimulus presentations. There was a 12 element-long repeating sequence (121423413243; numbers represent the position of the black circle) in Blocks 1–11. In Block 12 stimuli appeared in pseudorandom order, with the constraint that the target stimulus cannot appear in the same location twice in a row.

Participants were not informed about the goal of the task. They were only told to be as quick and as accurate in response as possible. Participants were asked to respond with two hands and to keep their fingers on the response keys during the whole session.

2.1.1.2. The AGL task: Stimuli of the AGL task were adapted from the P language of Saffran (2002). The task consisted of a training phase and the test phase. During the training phase, 58 sentences were presented auditorily to the participants in a preset pseu-

Table 2
Words in the artificial language by category (with Hungarian transcription)

Category				
A	bif	hep	mib	rud
C	kav	lam	neb	szig
D	lor	gal		
F	dup	dók	rász	vot
G	tez	péf		

dorandom order. The pseudorandomized set of 58 sentences were repeated twice. Sentences in the training phase followed the rules of the grammar in Eq. (1), with the vocabulary for categories A to G provided in Table 2; S stands for Sentence, and P stands for different types of phrases that are building blocks of sentences. Both children and adults were offered to draw or watch randomly appearing photos. Although systematic data collection was made, adults mostly chose watching pictures, while children were more likely to choose drawing. Participants were told to draw or watch pictures and to keep the headphones over their ears in which strange strings of words are going to be played.

After the training phase, participants were told that they had heard sentences of a strange language. They were also told that their following task would be to listen to two sentences and decide which one was similar to the language previously heard. Pairs were preset, but the order of the pairs was random. Across the 24 test pairs, the legal–illegal order was counterbalanced. Participants could respond by clicking number 1 (for the first sentence of the pair) or number 2 (for the second sentence) appearing on the screen (younger children got help from the experimenter in clicking on the chosen number). Each ungrammatical sentence included violations of one of the four rules listed in Eq. (1), illustrated by examples of grammatical–ungrammatical sentence pairs like *hep gal lam pe̋f dők* [A-D-C-G-F] - * *hep pe̋f lam gal dők* [A-G-C-D-F]; or *rud gal neb dup* [A-D-C-F] -: * *rud gal dup* [A-D-F]). Each rule was associated with six legal–illegal sentence pairs, also following Saffran (2002).

$$\begin{aligned} S &\rightarrow AP + BP + (CP) \\ AP &\rightarrow A + (D) \\ BP &\rightarrow CP + F \\ CP &\rightarrow C + (G) \end{aligned} \tag{1}$$

2.1.1.3. The WP task: The WP task was an adaptation of Knowlton, Squire, and Gluck (1994), adapted by Kemény and Lukács (2010) for testing children. One, two or three out of four different cues appeared at the same time in the WP task. Cue 1 was a square, Cue2 was a triangle, Cue3 was a pentagon, and Cue4 was a rhombus. The size of the cues was adjusted to fit into a 125 × 125 pixel square. Cues always appeared 144 pixels from the top. If one cue appeared at a time, the cue appeared in the horizontal center line. If more than one cue appeared at the same time, the cues were arranged horizontally. In the case of two simultaneous cues, the gap between the two cues was in the horizontal center line, while in the case of the combination of three cues, the second cue was aligned to the horizontal center line. The order of cues in combinations always appeared in the way that Cue1 if present was the leftmost cue, Cue2 was always spatially before Cue3 and Cue4, but after Cue1, and Cue3 always preceded Cue4.

After each prediction a feedback was displayed. The feedback was a 83 pixels wide and 86 pixels high icon of the outcome; that is, the icon of the sun or a cloud with rain. The icon appeared in the horizontal center line, 343 pixels from the top.

One, two, or three cues could appear at the same time in the WP task. Participants are asked to decide whether it would be sunshine or rain. If they predict sunshine, they have to press ENTER, while if the prediction is rain, they have to press SPACE. Participants completed 4 blocks with 50 predictions in each block. Items followed a preset pseudorandom order. The only constraint in the randomization was that two consecutive items could not employ the same stimuli. That is, the combination of cues 1 and 2 could not appear twice in a row, while this combination could be followed by, for example, the combination of cues 1, 2, and 3.

Unknown to the participants, each cue had a preset predictive value. There were two cues predicting sunshine in the majority of the cases, while two cues were likely to predict rain. There was a strong and a weak cue for both outcomes. Cue1 predicted sunshine in 85.7% of all its appearances, Cue2 predicted sun in 70% of its appearances, Cue3 was associated with sun in only 30% of the cases, while Cue4 predicted sun in 14.3% (note that the cues were associated with rain in all other cases). Table 3 summarizes the design.

Table 3
Types and occurrences of cues or cue-combinations per blocks of 50 trials

Cues	Frequency	*p* (SUN)
A	8	0.875
B	4	0.75
C	4	0.25
D	8	0.125
AB	8	0.875
AC	1	1
BC	2	0.5
BD	1	0
CD	8	0.125
ABC	2	1
ABD	1	1
ACD	1	0
BCD	2	0

Note. The first column (Cues) shows which cues are present in a given combination: A is cue1, B is cue2, C is cue3, D is cue4. Frequency is the number of appearances within a block of 50 trials. The third column provides the probability that the given cue or combination leads to sunshine.

3. Results

3.1. SRT

In the case of the SRT task, our primary focus was on age-related differences in sequence learning, which is characterized by the increase in RT from the last sequence block (Block 11) to the random block (Block 12) without the confounding the practice effect on motor responses, but we also include analyses on overall practice effects (comparisons of RTs from the first and last sequence blocks, Block 1 and Block 11). As accuracy averages were close to ceiling in all age groups (they varied between 92% and 98%), they are not analyzed. Only items with correct responses were considered in data analysis, and participants with an overall accuracy below 80% were excluded from the analysis ($N = 14$). RTs below and above two standard deviations were considered as outliers and were excluded. The exclusion criterion was applied for each participant. Analyses were conducted on mean RTs. Since we seek age-based differences in performance, we used LSD post hoc tests for pairwise comparisons. Normally the use of LSD post hoc tests is not encouraged due to the high possibility of Type I errors. However, if we are interested in clustering different groups as similar, reducing the alpha level would elevate the possibility of Type II errors, leading to the possibly false conclusion that the specific groups do not differ from each other. For this reason we did not correct the alpha-levels; this way, it is more likely that similar performances of groups will not differ from each other significantly. Analysis of raw data is included to show groupwise differences in performance itself. However, groups with different baseline RTs are difficult to compare. For this reason we included analyses of normalized data. We first present the analysis on raw RTs, then proceed to normalized RT data.

3.1.1. Raw RTs

To look for age-related differences in sequence learning, a 2 × 10 repeated-measures ANOVA was conducted with Block (Sequence vs. Random) as a within-subject variable and age group (U09 vs. U11 vs. U14 vs. U18 vs. 025 vs. U35 vs. U45 vs. U55 vs. U65 vs. U99; U stands for "under"; participants of the U11 group are above 9 and below 11 years of age) as between-subject variable. Results showed that the mean RTs of the Sequence block were significantly lower than those of the Random block, as evidenced by a significant main effect of Block, $F(1, 456) = 81.367$, $p < 0.001$, $\eta_p^2 = .151$. There was a significant main effect of age group too, $F(9, 456) = 42.866$, $p < 0.001$, $\eta_p^2 = .458$. The Block × Age group interaction was not significant, $F(9, 456) = 1.138$, $p = 0.334$, $\eta_p^2 = .022$, showing that the sequence-learning effect is present and its magnitude does not differ by age. Due to the lack of Block × Age group interaction, sequence-specific knowledge by age group is not analyzed further. Also, simple RT-baseline differences are not reported in more detail.

To test whether practice effect appears and varies between age groups, we conducted a 2 × 10 repeated-measures ANOVA with Block (Block1 vs. Block 11) as within-subject variable, and age group as between-subject variable. The ANOVA revealed that Block11 RTs were significantly lower, shown by a significant main effect of Block, $F(1, 456) = 552.471$, $p < 0.001$, $\eta_p^2 = .548$. There was a significant main effect of age group, $F(9, 456) = 12.348$, $p < 0.001$, $\eta_p^2 = .196$. The Block × Age group interaction was also significant, $F(9, 456) = 10.691$, $p < 0.001$, $\eta_p^2 = .174$. To test whether the practice effect appears in each age group, we conducted a separate repeated-measures ANOVA with Block (Block1 vs. Block11) as a within-subject variable. Block11 RTs were significantly lower for all age groups (all ps < .001).

With a One-sample *T*-test, we also tested whether the amount of learning was significantly different from chance in each age group. Using a difference score of 0 as chance level revealed the following results. Performance in the youngest group did not differ from chance, whereas we found significant differences for all other age groups ($p < 0.05$ for U99, $p < 0.01$ for U45, and $p < 0.001$ for all other age groups).

3.1.2. *Z-transformed RTs*

As explained above, there may be differences in the baseline RTs between the groups, which could mask significant group differences in RT change. Due to these differences in baseline RTs, we decided to apply z-transformations: each individual mean and standard deviation was computed separately, then the response latency for each item was z-transformed by subtracting the given participant's overall mean from the raw RT, and dividing the value with the participant's standard deviation (Christ, White, Mandernach, & Keys, 2001).[3] The Z-transformation was applied to all participants individually, and RTs of all 12 blocks were included.

We computed a single measure of sequence learning by subtracting mean Block 11 Z-transformed RT from mean Block 12 Z-transformed RT for each participant; see

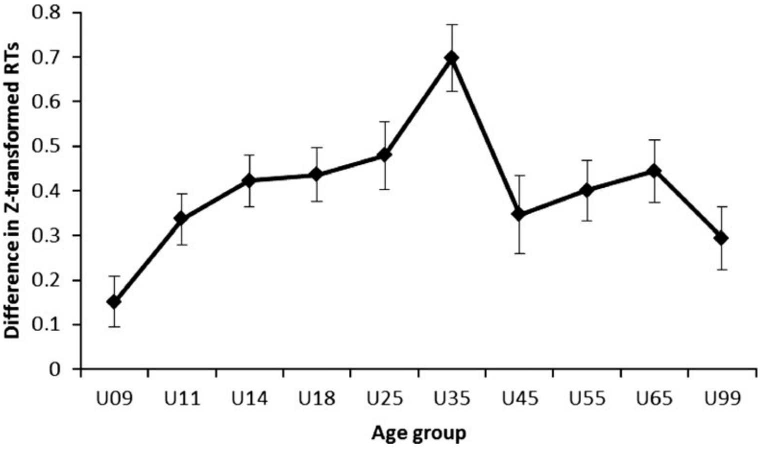


Fig. 1. Block 11–12 differences in Z-transformed reaction time (RT) data by age groups. Error bars indicate *SEM*.

Table 4
Pairwise comparisons between age groups in Block 12-Block 11 Z-transformed RT data of the SRT task

U09	—									
U11	*	—								
U14	***		—							
U18	***			—						
U25	***				—					
U35	***	***	**	**	*	—				
U45						**	—			
U55	**					**		—		
U65	**					*			—	
U99						***				—
Z-RT	U09	U11	U14	U18	U25	U35	U45	U55	U65	U99

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
RT, reaction time; SRT, serial reaction time task.

Fig. 1. A univariate ANOVA was employed on the RT difference as dependent variable and age group as between-subject variable. The ANOVA revealed that the different age groups showed a different rate of sequence-specific RT increase, as confirmed by a significant main effect of age group, $F(9, 456) = 4.131$, $p < 0.001$, $\eta_p^2 = .075$. LSD post hoc tests revealed that the U35 group showed the highest sequence-specific RT increase: It differed from all other groups (all *p*s < .05). The U09 group showed the lowest sequence-specific RT increase, differing from all groups ($p < 0.05$, except U45, where $p = 0.072$, and U99 where $p = 0.134$). The groups in between: U11, U14, U18, U25, U45, U55, U65, and U99 showed no significant pairwise differences, all *p*s > .126, except for the U25 and U99 difference, which approached significance ($p = 0.072$). All post hoc comparisons are provided in Table 4.

With a one-sample *T*-test, we also tested whether the amount of learning was significantly different from chance in each age group. Using a difference score of 0 as chance level showed significant differences for all age groups ($p < 0.01$ for U9 and $p < 0.001$ for all other age groups).

Table 5
Pairwise comparisons between age groups in performance on the artificial grammar learning (AGL) task

U09	—									
U11	*	—								
U14			—							
U18		*		—						
U25	*				—					
U35						—				
U45	**			*			—			
U55								—		
U65		*					*		—	
U99		*					*			—
AGL	U09	U11	U14	U18	U25	U35	U45	U55	U65	U99

Note. *$p < 0.05$, **$p < 0.01$.

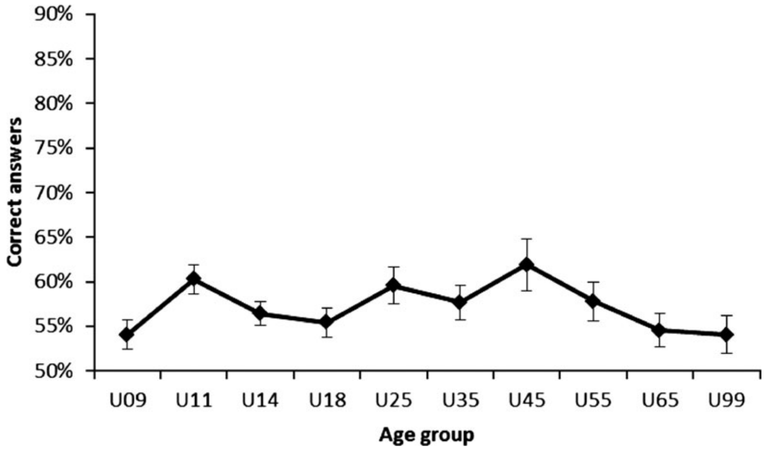


Fig. 2. Percentage of correct choices between legal–illegal test pairs on the artificial grammar learning task by age group. Error bars indicate *SEM*.

3.2. AGL

In the AGL task, performance was measured by the percentage of correct answers in the test phase. A one-way ANOVA was conducted with age group as between-subject variable. As in the analysis of the SRT data, for post hoc comparisons we used LSD tests. The higher possibility of false positives might enhance spotting similarities.

There was a significant main effect of age group, $F(9, 470) = 2.007$, $p < 0.05$, $\eta_p^2 = .037$. LSD post hoc tests showed that the highest performance was observed in the U45 and U11 age groups, and lowest performances were seen in the U09, U18, U65, and U99 groups. Performances differed significantly between groups with lower and groups with higher performance, all ps < .05. The U25 group showed a similar performance to the high performance groups, but only differed from the U09 group ($p < 0.05$). The U14, U35, and U55 groups did not significantly differ from any other groups. Pairwise comparisons are shown in Table 5; see also Fig. 2.

As in other studies with AGL (e.g., Conway & Christiansen, 2006), each group's performance was compared to a target value of 50% (chance level) using a one-sample *T*-test. The *T*-tests revealed that the performance of the U99 group did not differ significantly from chance level, $t(42) = 1.917$, $p = 0.062$. Performance was significantly higher than chance level in all other age groups ($p < 0.05$ for U09 and U65, $p < 0.001$ for all other groups).

3.3. WP

In the case of the WP task, we tested age effects on categorization performance measured as percentage of correct responses on Block 4. A univariate ANOVA was employed with age group as between-subject variable, which showed a significant main effect on categorization performance, $F(9, 470) = 13.905$, $p < 0.001$, $\eta_p^2 = .210$. LSD post hoc

Table 6
Pairwise comparisons between age groups in performance on the weather prediction task

U09	—									
U11	**	—								
U14			—							
U18	***		*	—						
U25	***	***	***	***	—					
U35	***	***	***	**		—				
U45	***	**	***		**		—			
U55	***	**	***		***			—		
U65	***	**	***		***				—	
U99	**				***	***	*		*	—
PCL	U09	U11	U14	U18	U25	U35	U45	U55	U65	U99

Note. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

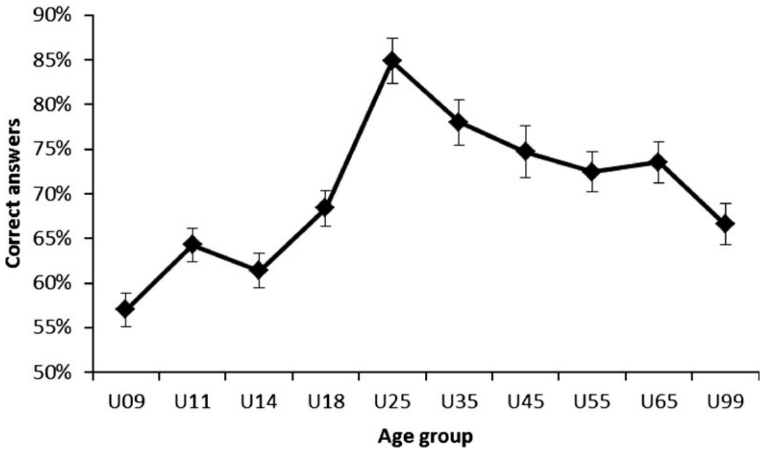


Fig. 3. Block 4 categorization performance in the weather prediction task. Error bars indicate *SEM*.

comparison of the groups revealed that adults under 25 performed significantly higher than all other groups (all *p*s < .001, except comparing with U45, where $p < 0.01$). Categorization performance in the U35, U45, U55, and U65 was lower than in the U25 group, as described above, and did not differ from each other significantly (all *p*s > .05). Performance was also similar in the U11, U18, and U99 group (no significant differences were observed across the groups, all *p*s > .05), and each of them shows lower than U35, U45, U55, and U65 (all *p*s < .01). The U14 group performed yet a little lower (differing significantly from the U18 group, $p < 0.05$). The lowest performance was observed for the U09 group, differing from all groups, except U14 ($p = 0.106$, otherwise *p*s < .01 for the U11 and U99 groups, and *p*s < .001 for all other groups. See Table 6 and Fig. 3.

As for the AGL task, the different groups' performance was compared to a target value of 50% (chance level) using One-sample *T*-tests, to test whether the different age groups' performance is above chance at all. All groups differed significantly from chance level ($p < 0.05$ for the U09 group, and $p < 0.001$ for all other group).

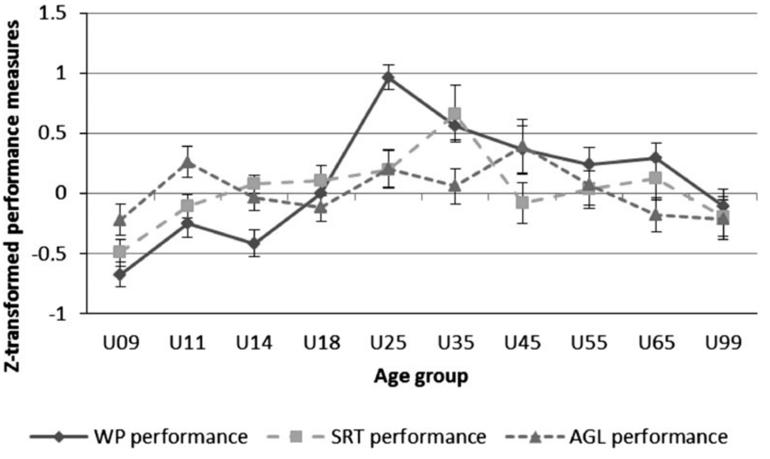


Fig. 4. Age-related differences in normalized performance for three different forms of skill learning.

Fig. 4 shows Z-transformed results on the three tasks to facilitate comparison of differences in performance on different forms of SL.

3.4. Associations between different forms of skill learning

To test relationships between different forms of SL, participants' performance on the three tasks was correlated. Pearson's bivariate correlations revealed that Block4 performance on the WP task has a significant, but negligible positive correlation with the Z-transformed sequence learning measure of the SRT task, $r = .168$, $N = 480$, $p < 0.001$. Block 4 WP performance also significantly, but very weakly correlated with AGL performance, $r = .093$, $N = 480$, $p < 0.05$. The correlation between performance on the SRT task and the AGL task was not significant, $p = 0.705$.

Partial correlations controlling for age revealed the same pattern and strengths of associations: Block 4 performance on the WP task was very weakly, but significantly positively correlated with the Z-transformed sequence learning measure of the SRT task, $r = .168$, $N = 480$, $p < 0.001$. Block 4 WP performance also correlated with AGL performance, $r = .102$, $N = 480$, $p < 0.05$. The correlation between the SRT task and the AGL task was not significant, $p = 0.715$.

4. Summary and discussion

Our aim in this paper was to examine age-related differences in three forms of SL in three different paradigms between 7 and 87 years, by collecting data from the same pool of participants from an SRT, an AGL, and a WP task. Using the same method facilitates direct comparisons across age groups, and this way offers a unique opportunity to extend previous findings on the development of SL. Additionally, in contrast to previous studies comparing children and adults, or young adults and the elderly, the large sample size allows us to group participants into 10 relatively narrow age ranges, improving understanding of the developmental trajectory. Tracking differences on the three tasks also gives us the possibility to see whether age-related differences follow the same trend on different tasks which test different aspects of complex SL. By comparing age-related differences on these three tasks, we can test whether learning of motor sequences (SRT), the extraction of abstract regularities from auditory sequences (AGL), and learning on a non-sequential categorization task (WP) show the same developmental trajectory and argue for the same model of the development of SL. Observing the same pattern in three different tasks provides a stronger case for a developmental model; differences in developmental trajectories raise questions about the unitary nature of the SL process.

Our theoretical interest in plotting developmental differences in different forms of SL was to see whether SL is an ability that (a) is most effective in childhood, and then gradually declines, or (b) is uniformly available at all ages as a robust learning process or (c) improves into adulthood and then declines with aging. The findings from three different tasks uniformly argue for a model where SL gets better with age, with a peak between 18 and 35, no or milder decline until 65, and a larger drop in performance above 65. This pattern shows that learning is more effective in adolescence and adulthood than in childhood: The forms of SL tested by these three tasks (learning of deterministic motor sequences—SRT, the extraction of abstract regularities from auditory sequences—AGL, and learning on a non-sequential probabilistic categorization task) get better with age, and are more effective in adulthood than in childhood. The results also support previous observations that some forms of SL are not as effective at old age as in young adulthood. These findings are in contrast with the conclusions of Janacsek et al. (2012) from a similarly large sample and wide age range, who argue for a dedicated period of SL in childhood ending around 14 years of age, and decrease in effectiveness in adults and the elderly. Although conclusions of the two studies differ, results, where comparable, in fact show similar trends. As discussed above, they used an ASRT task and took sequence learning measures from raw RTs as the basis for their model. At the same time, they also analyzed Z-transformed RTs controlling for age-related differences in baseline RTs, which yielded a similar developmental trend as the one we found for all three tasks: effectiveness of sequence learning increased from 4 to 12 years, stayed the same until the mid-thirties, and then declined into old age.

Looking for relationships between these different forms of SL, we haven't found any meaningful associations between performances on the three tasks. The lack of strong associations may be due to differences in the nature of the three tasks and the measures collected in the tasks: The SRT task is an online motor-based sequence-learning task, the AGL task is an offline similarity judgment task, while the WP task is a non-sequential forced choice paradigm.

Although the overall developmental pattern of differences in performance was similar across the three tasks, there were also differences in the specific timing of these changes, and also between the sizes of age group differences. These differences, together with diverging results of previous studies, may partly be due to task complexity differences. As Howard and Howard (1997, 2012) and Howard et al. (2007) argued, task complexity can be an important factor. Also, while Thomas et al. (2004, reviewed in the introduction) found better sequence learning on the SRT task in adults than children, in an earlier study the authors (Thomas & Nelson, 2001) did not find differences in sequence learning magnitude between 4, 7, and 10-year-old children (although there were differences in accuracy, and also in the rate of children who showed sequence-specific learning in each age group—fewer children in younger groups: 72% in 4-year-olds, 87% in 7-year-olds, and every 10-year-old). They explain this contrast with differences in the complexity of the sequences used in the two studies and make the important point that task or sequence complexity is an important factor in tapping into age-related differences in IL, and it can affect efficiency of sequence learning in adults as well. Thomas and colleagues also cite infant results supporting the effect of complexity on learning from Clohessy, Posner, and Rothbart (2001) from sequence learning. Infants younger than 18 months were only able to learn sequences with no ambiguous elements (like 1-2-3), while infants over 18 months could also learn a sequence with an ambiguous element (1-2-1-3).

As in our study we did not vary task complexity, we have no easy or direct way of comparing the three tasks in this respect. For the AGL and WP tasks, one indication of complexity differences is the difference in accuracy peaks: at the level of age groups, for AGL this is at around 62%, while for the WP, highest performance is 75%. The complexity of the WP allows for more variation in performance across ages, while the AGL may be too complex even for those in the most effective period of learning. Another difference can be pointed out between the SRT and AGL tasks: The SRT relies on learning a fixed sequence of specific elements, while the AGL requires abstraction. These are only speculations on the potential complexity differences. Future studies should directly address the issue of complexity by manipulating it within the same task. This way we would get a better picture of how too easy and too difficult tasks can mask age differences in the effectiveness of learning.

On a related issue, we also have to point out that in the case of the three tasks we used (and often in other studies where age-related differences are found in performance on SL), it is difficult to tell whether these reflect developmental shifts of IL abilities. Implicit tasks vary greatly, and the nature and complexity of the material to be learned can charge working memory, attention, executive functions, and explicit processes differently; these effects are difficult to rule out, and capacities behind each are known to change with age. There is evidence that working memory deficits are associated with diminished efficiency of learning on the SRT task (Gomez Beldarrain, Grafman, Pascual-Leone, & Garcia-Monco, 1999). Several studies (e.g., Dienes, 2008; Frensch & Rünger, 2003; Kihlstrom, Dorfman, & Park, 2007) argue for the contribution of both implicit and explicit processes in these traditional SL paradigms. The WP task has been found to rely on explicit processes by a number of experimental psychological studies (Kemény & Lukács, 2013; Lagnado, Newell, Kahan, & Shanks, 2006; Newell, Lagnado, & Shanks, 2007; Price, 2009). In the case of sequence learning (both AGL and SRT tasks), previous studies found a high rate of explicit knowledge using free-recall and cued recall, as well as recognition (e.g., Jackson & Jackson, 1995). From a theoretical perspective, these tasks are expected to selectively rely on implicit processes; however, it is almost impossible to find process-pure learning tasks (e.g., Destrebecqz & Cleeremans, 2001). In sum, newer studies of IL suggest the cooperation of implicit and explicit processes in these paradigms (and there are studies that suggest a unitary explicit memory system, e.g., Lagnado et al., 2006). Further research has to focus on factoring out contributions from developmental changes in working memory or explicit learning to age-related difference in IL performance. One study by Vinter and Perruchet (1999) attempts do so; they applied a "neutral parameter procedure" to minimize the effect of explicit factors on learning by focusing participants' attention on aspects of the task other than the target of learning. This way, the authors found that learning was equally effective in children between 4 and 10 years, supporting Reber's (1993) claim of age independence.

The contribution of working memory and explicit learning is unclear for the tasks we tested. Performance on the SRT, AGL, and WP tasks probably relies on a combination of different memory systems, but so does complex SL in real-life circumstances, and on that point, our findings argue for improved effectiveness of learning into adulthood, and a decline in acquiring such skills in old age. For IL per se, our results are also compatible with invariance models modulated to different extents by changes in working memory and explicit effects; further studies including working memory measures of participants should resolve this question.

Our findings are in contrast with the third model of SL, which proposes the existence of critical periods for acquiring skills (see e.g., Lenneberg, 1967; Johnson & Newport, 1989 for language, Janacsek et al., 2012 for SL), suggesting the existence of early age windows in which experience, stimuli, or practice relevant to the skill in question is more effective than at older ages. It is certainly true that to become a world champion in chess or football, or a great pianist or painter, it is best to start working on one's gifts early. There are factors that might make childhood and adolescence especially effective for acquiring certain skills: physical changes in the muscles and joints, the amount of practice and time available for learning besides other activities, the type and timing of learning one is able to engage in (these also vary with age), and the focus on discovering the gifts a child has. In the light of the results presented in the paper, together with the majority of studies on developmental differences in SL, it seems that age-related changes in the learning mechanism itself are not among those factors: SL shows its peak efficiency in young adulthood and operates at the same level until at least the late fifties. This paper is the first to provide direct evidence for that trajectory in three different domains of SL systematically sampling a large age range in narrow age groups.

Acknowledgments

The paper benefitted a lot from discussions with Vinciane Gaillard, Rebecca Gómez, Darlene V. Howard, James H. Howard, Karolina Janacsek, Beat Meier, Dezső Németh, and Kathleen M. Thomas. We are grateful for all their helpful comments. We would also like to thank Kornél Németh, Gábor Szvoboda, Kata Fazekas, and psychology students of the Péter Pázmány Catholic University, who were instrumental in collecting the data. 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. During the period of research and writing the paper, Ágnes Lukács was a grantee of the János Bolyai Research Scholarship of the Hungarian Academy of Science. This work is connected to the scientific program of the “Talent care and cultivation in the scientific workshops of BME (Budapest University of Technology and Economics, Hungary)” project and is supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.2/B-10/1-2010-2009).

Notes

1. Studies sometimes consider implicit and procedural learning as equivalent, similarly to explicit and declarative learning (e.g., Price, 2009). Others suggest that the two functions do not necessarily map onto each other (e.g., Berry & Dienes, 1993). In the sequence learning literature, implicit and statistical learning are also considered as the same phenomenon with two approaches (Perruchet & Pacton, 2006). Because trying to separate the processes involved in statistical/implicit/procedural learning tasks is beyond the scope of the paper, we will not differentiate between these concepts, admitting that the theoretical questions these concepts raise is a very important one.
2. The literature on age-related changes in SL and IL often uses the terms *improve* and *decline* to describe differences in performance between different age groups based on a cross-sectional design. We are aware of the limitations of drawing such inferences from cross-sectional studies like ours (and the majority of studies on SL). A more neutral terminology would perhaps be a more accurate description of the findings; on the other hand, it would yield a very complicated description of previous and current findings. For ease of exposition, we use terms like *improve* and *decline* throughout the paper and ask the reader to keep in mind that we are talking about differences between age groups.
3. Since there are no expectations for baseline differences between different age groups in the AGL and WP tasks, we did not apply Z-transformations for these two tasks. We expect participants to have no rule or probability knowledge on the AGL and WP tasks prior to the experiment, setting the theoretical baseline at chance level, that is, 50% before the training.

References

Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to adulthood: Item frequencies, relations, and cognitive flexibility. *Developmental Psychobiology*, *54*(6), 664–673.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442–481.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science*, *2*, 83–89.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, UK: Lawrence Erlbaum.

Christ, S. E., White, D. A., Mandernach, T., & Keys, B. A. (2001). Inhibitory control across the life span. *Developmental Neuropsychology*, *203*, 653–669.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406–416.

Clohessy, A. B., Posner, M. I., & Rothbart, M. K. (2001). Development of the functional visual field. *Acta Psychologica*, *106*, 51–68.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *2104466*, 207–210.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities—Pitting abstract against stimulus-specific representations. *Psychological Science*, *1710*, 905–912.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*, 343–350.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research*, *168*, 49–64.

Fletcher, J., Maybery, M. T., & Bennett, S. (2000). Implicit learning differences: A question of developmental level? *Journal of Experimental Psychology: Learning, Memory, Cognition*, *261*, 246–252.

Frensch, P. A., & Miner, C. S. (1994). Effects of presentation rate and of individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition*, *22*, 95–110.

Frensch, P. A., & Rünger, D. (2003). Implicit learning. *Current Directions in Psychological Science*, *12*, 13–18.

Gaillard, V., Destrebecqz, A., & Cleeremans, A. (2009). Effects of age and practice in sequence learning: A graded account of aging, learning, and control. *European Journal of Cognitive Psychology*, *21*, 255–282.

Gomez Beldarrain, M., Grafman, J., Pascual-Leone, A., & Garcia-Monco, C. (1999). Procedural learning is impaired in patients with prefrontal lesions. *Neurology*, *529*, 1853–1860.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *702*, 109–135.

Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*, 183–206.

Howard, D. V., & Howard Jr., J. H. (1989). Age differences in learning serial patterns: Direct versus indirect measures. *Psychology and Aging*, *4*, 357–364.

Howard, D. V., & Howard Jr., J. H. (1992). Adult age differences in the rate of learning serial patterns: Evidence from direct and indirect tests. *Psychology and Aging*, *7*, 232–241.

Howard Jr., J. H., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology of Aging*, *12*, 634–656.

Howard, D. V., & Howard Jr., J. R. (2012). Dissociable forms of implicit learning in aging. In M. Naveh-Benjamin & N. Ohta (Eds.), *Perspectives on human memory and aging* (pp. 125–151). New York: Psychology Press.

Howard Jr., J. H., Howard, D. V., Dennis, N. A., & Yankovith, H. (2007). Event timing and age deficits in higher-order sequence learning. *Aging, Neuropsychology, and Cognition*, *14*, 647–668.

Huang-Pollock, C., Maddox, T., & Karalunas, S. L. (2011). Development of implicit and explicit category learning. *Journal of Experimental Child Psychology*, *109*, 321–335.

Jackson, G. M., & Jackson, S. R. (1995). Do measures of explicit learning actually measure what is being learnt in the serial reaction time task? A critique of current methods. *Psyche*, *2*. Available at: http://www.theassc.org/files/assc/2352.pdf. Accessed July 27, 2012.

Janacsek, K., Fiser, J., & Németh, D. (2012). The best time to acquire new skills: Age related changes in implicit learning across human life span. *Developmental Science*, *154*, 496–505.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–100.

Kemény, F., & Lukács, Á. (2010). Impaired procedural learning in language impairment: Results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology*, *323*, 249–258.

Kemény, F., & Lukács, Á. (2013). Self-insight in probabilistic category learning. *Journal of General Psychology*, *140*(1), 57–81.

Kihlstrom, J. F., Dorfman, J., & Park, L. (2007). Implicit and explicit memory and learning. In M. Velmans, & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 525–539). Malden, MA: Blackwell.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic category learning in amnesia. *Learning & Memory*, *1*, 106–120.

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Language Sciences*, *3*, 59.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General*, *135*, 162–183.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Maddox, W. T., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby, F. G. (2010). Category label and response location shifts in category learning. *Psychological Research*, *74*, 219–236.

Marcus, G. (2012). *Guitar zero: The new musician and the science of learning*. New York: Penguin.

Maybery, M. T., Taylor, M., & O'Brien-Malone, A. (1995). Implicit learning: Sensitive to age but not IQ. *Australian Journal of Psychology*, *471*, 8–17.

Meulemans, T., Van Der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, *69*, 199–221.

Nejati, V., Farshi, M. T., Ashayeri, H., & Aghdasi, M. T. (2008). Dual task interference in implicit sequence learning by young and old adults. *International Journal of Geriatric Psychiatry*, *23*, 801–804.

Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2007). Challenging the role of implicit processes in probabilistic category learning. *Psychonomic Bulletin & Review*, *14*, 505–511.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning—Evidence from performance-measures. *Cognitive Psychology*, *19*(1), 1–32.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233–238.

Price, A. L. (2009). Distinguishing the contributions of implicit and explicit processes to performance of the weather prediction task. *Memory & Cognition*, *372*, 210–222.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Reber, A. S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition*, *1*, 93–113.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Roter, A. (1985). Implicit processing: A developmental study. Unpublished Doctoral Dissertation, City University of New York.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*, 172–196.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by adults and infants. *Cognition*, *70*, 27–52.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening and learning out of the corner of your ear. *Psychological Science*, *8*, 101–105.

Salthouse, T. A., McGuthry, K. E., & Hambrick, D. Z. (1999). A framework for analyzing and interpreting differential aging patterns: Application to three measures of implicit learning. *Aging, Neuropsychology, and Cognition*, *61*, 1–18.

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. S. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience*, *16*, 1339–1351.

Thomas, K. M., & Nelson, C. A. (2001). Serial reaction time learning in preschool- and school-age children. *Journal of Experimental Child Psychology*, *79*, 364–387.

Vaidya, C. J., Huger, M., Howard, D. V., & Howard Jr., J. H. (2007). Developmental differences in implicit learning of spatial context. *Neuropsychology*, *21*, 497–506.

Vinter, A., Pacton, S., Witt, A., & Perruchet, P. (2010). Implicit learning, development and education. In J.-P. Didier & E. Bigand (Eds.), *Rethinking physical and rehabilitation medicine* (pp. 111–127). Paris: Springer Verlag.

Vinter, A., & Perruchet, P. (1999). Isolating unconscious influences: The neutral parameter procedure. *Quarterly Journal of Experimental Psychology*, *52A*, 857–876.

Weiermann, B., & Meier, B. (2012). Incidental sequence learning across the lifespan. *Cognition*, *123*(3), 380–391.

Study 10.

Lukács Á., Ladányi, E., Fazekas, K. & Kemény F. (2016) Executive Functions and the Contribution of Short-Term Memory Span in Children With Specific Language Impairment. *Neuropsychology*, 30(3):296-303. . doi: 10.1037/neu0000232

Thesis 8. SLI difficulties in executive functions are mainly present on verbal versions of EF tasks, and are eliminated by controlling for verbal short term memory span. (10)

Neuropsychology
2016, Vol. 30, No. 3, 296–303

0894-4105/16/$12.00 http://dx.doi.org/10.1037/neu0000232

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, Trueswell, & 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-

This article was published Online First August 17, 2015.

Ágnes Lukács, Department of Cognitive Science, Budapest University of Technology and Economics; Enikő Ladányi, Department of Cognitive Science, Budapest University of Technology and Economics, and Research Institute for Linguistics, Hungarian Academy of Sciences; Kata Fazekas and Ferenc Kemény, Department of Cognitive Science, Budapest University of Technology and Economics.

We are grateful to Lucy Henry for providing us with the odd-one-out task. The paper benefited a lot from discussions with members of the Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment group (COST Action IS0804). 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.

Correspondence concerning this article should be addressed to Ágnes Lukács, BME Department of Cognitive Science, Egry J. u, 1, H-1111 Budapest, Hungary. E-mail: alukacs@cogsci.bme.hu

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, & Schechtman, 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.

Updating

Updating as an EF is defined as the monitoring of the contents of working memory, and the replacement of old information with new content. Children with SLI (between 7 and 12 years) were poorer than controls on the one-back condition in a dot-pattern recognition N-back task, and both groups performed at chance on the two-back task (Im-Bolter et al., 2006).

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., Dollaghan, 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 *SD*s 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 Mondatutánmondási Teszt; Kas & Lukács, in preparation), and a nonword repetition test (Racsmány, Lukács, Németh, & Pléh, 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

	TD		SLI				
	Mean	*SD*	Mean	*SD*	*F*	Sig	η_p^2
Age (years)	**7.81**	1.81	**7.84**	1.81	0.004	0.949	0.000
Raven IQ (standard score)	**107.06**	13.39	**103.10**	12.01	1.508	0.224	0.025
Nonword repetition (span)	**5.90**	1.35	**2.97**	1.58	61.799	<.001	0.507
Peabody Picture Vocabulary Test (raw scores)	**110.45**	24.7	**86.39**	21.83	16.518	<.001	0.216
Test for Reception of Grammar blocks (raw scores)	**16.77**	2.33	**11.90**	3.13	48.164	<.001	0.445
Sentence repetition (raw scores)	**35.35**	6.35	**17.90**	9.31	74.384	<.001	0.554

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

final word of each sentence in the set in the correct order. Our Hungarian version of the task was modeled after the listening recall task from Working Memory Test Battery for Children (Pickering & Gathercole, 2001).

The odd-one-out task is a nonverbal analogue of the listening recall task developed by Henry (2001). Participants are presented with sets of horizontal arrays of three nonsense shapes and are asked to spot the odd one out in each array. At the end of the set, they have to recall the location of the odd one out by pointing to the left, middle, or right box in an array of three empty boxes in the case of each trial of the set in the correct order. The span version of the odd-one-out task has a reliability of .80 (Henry, 2001).

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, Eschman, & Zuccolotto, 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 appears 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 × 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 × Group was also significant, $F(1, 60) = 26.199$, $p < .001$, $\eta_p^2 = 0.304$.

To further analyze the Type × 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 × 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 × 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
Means and SDs for the Measures of the Different Tasks for the TD and SLI Groups, Together With Results of One-Way ANOVAs as Group Comparisons

	TD		SLI				
	Mean	*SD*	Mean	*SD*	*F*	Sig	η_p^2
Digit span	4.58	0.92	3.39	0.67	34.055	<.001	0.362
Corsi span	4.29	1.01	4.26	0.86	0.018	0.892	0.000
Listening recall	2.23	0.92	1.29	0.86	17.024	<.001	0.221
Odd-one-out	2.97	1.30	2.52	1.09	2.188	0.144	0.035
Accuracy verbal Stroop	−3.00	6.76	−5.61	13.08	0.976	0.327	0.016
Accuracy nonverbal Stroop	−8.55	15.37	−11.29	16.18	0.468	0.497	0.008
RT verbal Stroop	324	290	306	359	0.048	0.827	0.001
RT nonverbal Stroop	397	371	317	337	0.795	0.376	0.013
Verbal one-back hit	8.84	1.83	8.35	2.39	0.803	0.374	0.013
Nonverbal one-back hit	9.23	1.12	8.90	1.70	0.780	0.381	0.013
Verbal one-back false alarm	0.48	1.15	2.16	4.75	3.655	0.061	0.057
Nonverbal one-back false alarm	3.81	7.91	3.81	8.78	0.000	1.000	0.000
Verbal two-back hit	5.48	2.63	3.97	2.70	5.010	0.029	0.077
Nonverbal two-back hit	4.65	2.67	4.58	2.87	0.008	0.927	0.000
Verbal two-back false alarm	1.39	1.99	2.87	5.99	1.714	0.195	0.028
Nonverbal two-back false alarm	2.90	3.00	7.32	11.67	4.172	0.045	0.065
Verbal fluency correct	31.55	12.26	22.45	11.16	9.329	0.003	0.135
Nonverbal fluency correct	8.84	5.01	10.10	6.45	0.735	0.395	0.012
Verbal fluency errors	1.48	2.79	2.19	2.63	1.063	0.307	0.017
Nonverbal fluency errors	7.48	5.53	6.06	5.46	1.033	0.314	0.017

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 $\times$ Group interaction) or RT measures (median RT for congruent items subtracted from the median RT for incongruent items in milliseconds; all p values > 0.436).

N-Back Tasks

In the case of one-back condition, no effects were significant for hit rate: all p values >0.111. Both groups of children produced a significantly higher number of false alarms in the nonverbal task, as revealed by a main effect of type, $F(1, 60) = 8.059$, $p = .006$, $\eta_p^2 = 0.118$. Neither the main effect of group ($p = .542$), nor the Type $\times$ 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 $\times$ Group interaction).

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

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

Further post hoc ANOVAs showed that there was a significant difference between the groups on the verbal, $F(1, 60) = 9.329$, $p = .003$, $\eta_p^2 = 0.135$, 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, $F(1, 60) = 46.770$, $p < .001$, $\eta_p^2 = 0.438$, with a higher number of errors in the nonverbal task. Neither the group main effect ($p = .671$), nor the Type $\times$ Group interaction ($p = .145$) was significant. No effects were significant in the case of repetition errors: all p values >0.273.

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 (Dispaldro et al., 2013; Finneran 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

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8,* 71–82. http://dx.doi.org/10.1076/chin.8.2.71.8724

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bulletin & Review, 14,* 919–924. http://dx.doi.org/10.3758/BF03194122

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49,* 5–28. http://dx.doi.org/10.1080/713755608

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

Bishop, D. V. M. (2012). *TROG—Test for Reception of Grammar handbook*. Budapest, Hungary: OS Hungary Tesztfejlesztő Kft.

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders, 49,* 381–415. http://dx.doi.org/10.1111/1460-6984.12101

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology. 2: Response inhibition. *Autism: The International Journal of Research and Practise, 9,* 29–43. http://dx.doi.org/10.1177/1362361305049028

Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research, 63,* 279–288. http://dx.doi.org/10.1007/s004269900006

Cambridge Cognition. (2006). *The Cambridge Neuropsychological Test Automated Battery (CANTAB): Test–retest reliability characteristics*. Cambridge, England: Author.

Cheuk, D. K. L., Wong, V., & Leung, G. M. (2005). Multilingual home environment and specific language impairment: A case-control study in Chinese children. *Paediatric and Perinatal Epidemiology, 19,* 303–314.

Csányi, F. I. (1974). *Peabody Szókincs-Teszt*. Budapest, Hungary: Bárczi Gusztáv Gyógypedagógiai Fõiskola.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS) examiner's manual*. San Antonio, TX: The Psychological Corporation.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378,* 279–281. http://dx.doi.org/10.1038/378279a0

Dibbets, P., Bakker, K., & Jolles, J. (2006). Functional MRI of task switching in children with Specific Language Impairment (SLI). *Neurocase, 12,* 71–79. http://dx.doi.org/10.1080/13554790500507032

Dispaldro, M., Leonard, L. B., Corradi, N., Ruffino, M., Bronte, T., & Facoetti, A. (2013). Visual attentional engagement deficits in children with specific language impairment and their role in real-time language processing. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 49,* 2126–2139. http://dx.doi.org/10.1016/j.cortex.2012.09.012

Dollaghan, C. A. (2007). *The handbook of evidence-based practice in communication*. Baltimore, MD: Brookes.

Engel de Abreu, P. M. J., Gathercole, S. E., & Martin, R. (2011). Disentangling the relationship between working memory and language: The roles of short-term storage and cognitive control. *Learning and Individual Differences, 21,* 569–574. http://dx.doi.org/10.1016/j.lindif.2011.06.002

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–159). New York, NY: Academic Press.

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 52,* 915–929. http://dx.doi.org/10.1044/1092-4388(2009/07-0053)

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17,* 172–179. http://dx.doi.org/10.1111/j.1467-9280.2006.01681.x

Gioia, G. A., Espy, K. A., & Isquith, P. K. (2003). Behavior Rating Inventory of Executive Function–PreSchool version. Lutz, FL: Psychological Assessment Resources, Inc.

Guy, S. C., Isquith, P. K., & Gioia, G. A. (2004). *Behavior. Rating inventory of executive function-self-report version*. Lutz, FL: Psychological Assessment Resources.

Hackman, D. A., Gallop, R., Evans, G. W., & Farah, M. J. (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science*. Advance online publication.

Henry, L. A. (2001). How does the severity of a learning disability affect working memory performance? *Memory, 9,* 233–247. http://dx.doi.org/10.1080/09658210042000085

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry, 53,* 37–45. http://dx.doi.org/10.1111/j.1469-7610.2011.02430.x

Hughes, D. M., Turkstra, L. S., & Wulfeck, B. B. (2009). Parent and self-ratings of executive function in adolescents with specific language impairment. *International Journal of Language & Communication Disorders, 44,* 901–916.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44,* 2017–2036. http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.010

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77,* 1822–1841. http://dx.doi.org/10.1111/j.1467-8624.2006.00976.x

Kail, R., & Leonard, L. B. (1986). Word-finding abilities in language-impaired children. *Asha Monographs, 25,* 1–39.

Kas, B., & Lukács, Á. (in preparation). *Magyar Mondatutánmondási Teszt*.

Leonard, L. (1998/2014). *Children with specific language impairment*. Cambridge, MA: The MIT Press.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10,* 447–454. http://dx.doi.org/10.1016/j.tics.2006.08.007

Lukács, Á., Győri, M., & Rózsa, S. (2012). A TROG pszichometriai jellemzőinek magyar vizsgálata, a normák kialakítása. In D. V. M. Bishop (Ed.), *TROG—Test for Reception of Grammar handbook* (pp. 47–86). Budapest, Hungary: OS Hungary Tesztfejlesztő Kft.

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Disorders, 43,* 181–200. http://dx.doi.org/10.1080/16066350701340719

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46,* 1138–1153. http://dx.doi.org/10.1044/1092-4388(2003/089)

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202. http://dx.doi.org/10.1146/annurev.neuro.24.1.167

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. http://dx.doi.org/10.1006/cogp.1999.0734

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology, 19,* 78–94. http://dx.doi.org/10.1044/1058-0360(2009/09-0028)

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation. Advances in research and theory* (pp. 1–18). New York, NY: Plenum Press. http://dx.doi.org/10.1007/978-1-4757-0629-1_1

Noterdaeme, M., Amorosa, H., Mildenberger, K., Sitter, S., & Minow, F. (2001). Evaluation of attention problems in children with autism and children with a specific language disorder. *European Child & Adolescent Psychiatry, 10,* 58–66. http://dx.doi.org/10.1007/s007870170048

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective & Behavioral Neuroscience, 5,* 263–281. http://dx.doi.org/10.3758/CABN.5.3.263

Novick, J. M., Trueswell, J. C., & Thomspon-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass, 4,* 906–924. http://dx.doi.org/10.1111/j.1749-818X.2010.00244.x

Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children*. London, England: Psychological Corporation.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle, 60,* 479–506. http://dx.doi.org/10.1556/MPSzle.60.2005.4.3

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

Reilly, S., Tomblin, B., Law, J., Mckean, C., Mensah, F., Morgan, A., . . . Wake, M. (2014). SLI: A convenient label for whom? *International Journal of Language & Communication Disorders, 49,* 415–433.

Robinson, R. J. (1991). Causes and associations of severe and persistent specific speech and language disorders in children. *Developmental Medicine and Child Neurology, 33,* 943–962.

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

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283,* 1657–1661. http://dx.doi.org/10.1126/science.283.5408.1657

Spaulding, T. J. (2008). *Attentional control in preschool children with specific language impairment*. Doctoral dissertation, University of Arizona.

Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53,* 725–738. http://dx.doi.org/10.1044/1092-4388(2009/09-0041)

Spaulding, T. J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research, 51,* 16–34. http://dx.doi.org/10.1044/1092-4388(2008/002)

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1275–1278. http://dx.doi.org/10.1044/jslhr.4205.1275

Weckerly, J., Wulfeck, B., & Reilly, J. (2001). Verbal fluency deficits in children with specific language impairment: Slow rapid naming or slow to name? *Child Neuropsychology, 7,* 142–152. http://dx.doi.org/10.1076/chin.7.3.142.8741

Weismer, S. E., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1249–1260. http://dx.doi.org/10.1044/jslhr.4205.1249

Weyandt, L. L., & Willis, W. G. (1994). Executive functions in school-aged children: Potential efficacy of tasks in discriminating clinical groups. *Developmental Neuropsychology, 10,* 27–38. http://dx.doi.org/10.1080/87565649409540564

Wittke, K., Spaulding, T. J., & Schechtman, C. J. (2013). Specific language impairment and executive functioning: Parent and teacher ratings of behavior. *American Journal of Speech-Language Pathology, 22,* 161–172. http://dx.doi.org/10.1044/1058-0360(2012/11-0052)

Received August 27, 2014
Revision received May 27, 2015
Accepted June 22, 2015 ■

Study 11.

Ladányi, E. & Lukács Á. (2016). Lexical Conflict Resolution in Children with Specific Language Impairment. *Journal of Communication Disorders*, 61, 119-130. doi: 10.1016/j.jcomdis.2016.04.004

Thesis 9. Lexical inhibition is effective in SLI (11).

Journal of Communication Disorders 61 (2016) 119–130

Contents lists available at ScienceDirect

Journal of Communication Disorders

Lexical conflict resolution in children with specific language impairment

Enikő Ladányi[a,b,*], Ágnes Lukács[b]

[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.

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).

http://dx.doi.org/10.1016/j.jcomdis.2016.04.004

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, Trueswell, & Thompson-Schill, 2009; Novick, Trueswell, & Thompson-Schill, 2005; Novick, Trueswell, & Thomspon-Schill, 2010), lexical ambiguity resolution (Bedny, Hulbert, & Thompson-Schill, 2007), the assessment of common ground (i.e. the set of shared beliefs by the interlocutors; Brown-Schmidt, 2009), and, most importantly from the point of view of our research question, in word production as well (Kan & Thompson-Schill, 2004; Schnur, Schwartz, Brecher, & Hodson, 2006; Schnur et al., 2009). When a word is retrieved several other words are activated and for successful production of the target word, the activation level of its lexical representation has to be higher than that of the competing words. When the difference between the activation level of the target word and the other words is not big enough, word retrieval difficulties may appear.

The blocked cyclic naming paradigm is a widely used paradigm to investigate word production under competition (e.g. Belke, Meyer, & Damian, 2005; Damian, Vigliocco, & Levelt, 2001; Kroll & Stewart, 1994; Schnur et al., 2006, 2009). In this task 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 . . .*). Sets of items are repeated in succession multiple times in various orders. For example, six fruits are presented after each other and then they are repeated several times in different orders. A robust finding is that adults name pictures in the homogeneous blocks significantly slower than pictures in the mixed blocks. This effect is usually not present in the first cycle (or first homogeneous cycles are named even faster than first mixed cycles), in the second cycle there is a large drop in reaction times

dc_1365_16

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 propose 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.

	TD		SLI		F	Sig	η_p^2
	Mean	SD	Mean	SD			
Age	8;9	1;4	8;10	1;4	0.16	0.901	0.001
Raven IQ (standardized scores)	108.54	9.71	100.54	5.83	6.473	0.018	0.212
Nonword repetition (raw scores)	6.23	1.01	3	1.47	42.506	0.000	0.639
PPVT (raw scores)	124.5	13.75	97	20.86	15.792	0.001	0.397
TROG blocks (raw scores)	18.38	1.45	13.15	2.67	38.533	0.000	0.616
Sentence repetition (raw scores)	37.38	3.23	19.76	8.21	51.862	0.000	0.684

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, Győri, & 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 ($F(1,34) = 0.001$, $p = 0.97$, $\eta_p^2 = 0.000$; 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

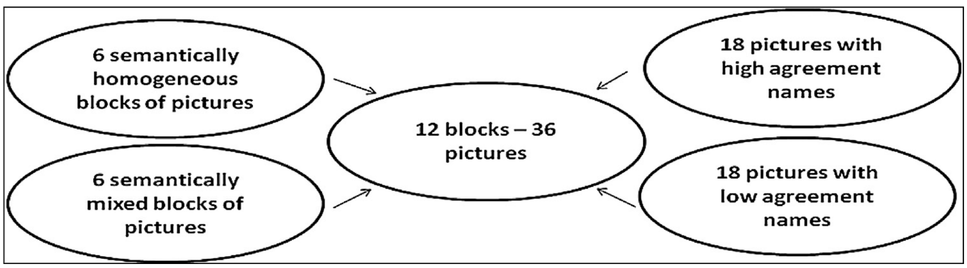


Fig. 1. Pictures were selected from six semantic categories with six pictures from each category, therefore 36 pictures were used in the task. These pictures were presented in two blocks: once together with members of the same category and once with members of other categories. Therefore altogether 12 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: *szófa* 'sofa', *kanapé* 'couch', *dívány* '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 × 2 × 2 × 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 ($F(1, 24) = 13.845$, $p < 0.001$, $\eta_p^2 = 0.366$): low agreement pictures took longer to name than high agreement pictures. All other main effects were nonsignificant (Homogeneity: $F(1,24) = 1.162$, $p = 0.242$ $\eta_p^2 = 0.046$; Cycle: $F(1,24) = 0.28$, $p = 0.867$ $\eta_p^2 = 0.001$; Group: $F(1,24) = 0.105$, $p = 0.794$ $\eta_p^2 = 0.004$). A significant interaction appeared between Homogeneity, Agreement and Group ($F(1, 24) = 8.841$, $p = 0.007$, $\eta_p^2 = 0.269$), Homogeneity and Cycle ($F(1,24) = 27.079$, $p < 0.001$, $\eta_p^2 = 0.530$), and Agreement, Cycle and Group ($F(1,24) = 5.092$, $p = 0.033$, $\eta_p^2 = 0.175$). All other interactions were nonsignificant (Homogeneity x Group: $F(1, 24) = 0.364$, $p = 0.552$, $\eta_p^2 = 0.015$; Agreement x Group: $F(1,24) = 0.157$, $p = 0.695$, $\eta_p^2 = 0.007$, Cycle x Group: $F(1,24) = 1.112$, $p = 0.302$, $\eta_p^2 = 0.044$; Homogeneity x Agreement: $F(1, 24) = 0.906$, $p = 0.351$, $\eta_p^2 = 0.036$; Homogeneity x Cycle x Group: $F(1, 24) = 1.038$, $p = 0.318$, $\eta_p^2 = 0.041$; Agreement x Cycle: $F(1, 24) = 3.349$, $p = 0.080$, $\eta_p^2 = 0.122$).

For breaking down the Homogeneity x Agreement x Group interaction, we analyzed the effect of Homogeneity and Agreement in 2 × 2 repeated measures ANOVAs separately for the two groups. In the SLI group the main effect of Agreement was significant ($F(1,12) = 5.076$, $p = 0.044$, $\eta_p^2 = 0.297$): 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.

	High name agreement		Low name agreement	
	SLI	TD	SLI	TD
Homogeneous cycles				
1.	1095 (165)	1087 (146)	1186 (166)	1163 (192)
2.	1089 (179)	1183 (211)	1288 (223)	1197 (173)
3.	1274 (209)	1165 (249)	1370 (193)	1309 (261)
4.	1249 (231)	1170 (182)	1222 (203)	1302 (237)
Mixed cycles				
1.	1167 (191)	1126 (176)	1329 (198)	1208 (280)
2.	1114 (174)	1062 (156)	1278 (174)	1241 (332)
3.	1130 (192)	1163 (263)	1153 (190)	1111 (235)
4.	1083 (286)	1094 (241)	1131 (267)	1144 (279)

agreement pictures. The main effect of Homogeneity was nonsignificant ($F(1,12) = 0.160$, *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 ($F(1, 12) = 1.804$, *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 ($F(1, 12) = 1.069$, *n.s.*), neither was the Homogeneity x Agreement interaction ($F(1, 12) = 2.937$, *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 ($F(1, 12) = 0.346$, *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: $F(1, 12) = 0.153$, *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 ($F(1, 23) = 0.866$, *n.s.*) and the Agreement x Cycle interaction ($F(1,12) = 0.57$, *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

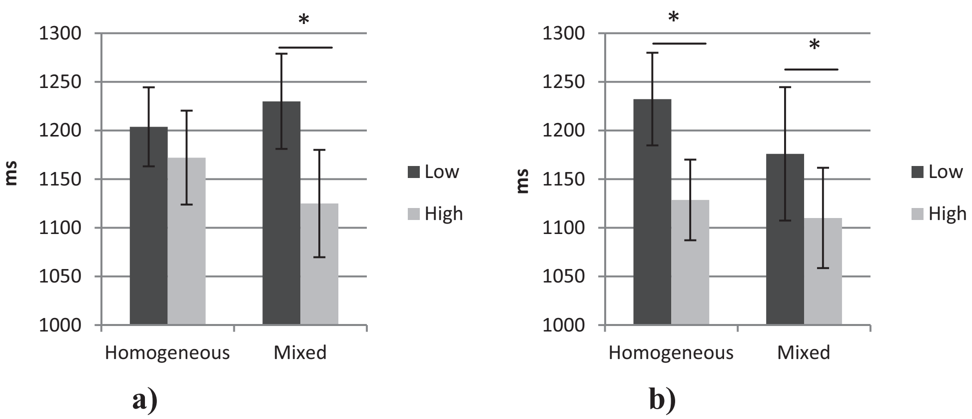


Fig. 2. Mean reaction times by Homogeneity and Agreement in the a) SLI and b) TD groups. Differences marked by * are significant at p < 0.05.

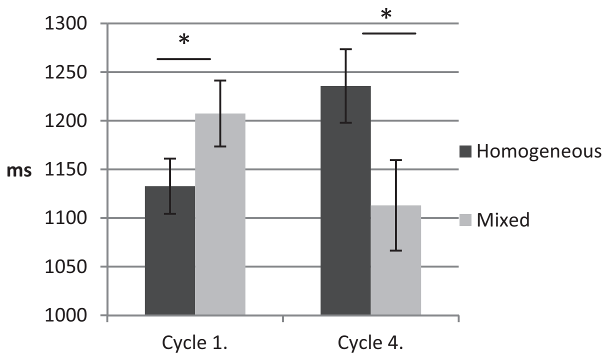


Fig. 3. Mean reaction times by Homogeneity and Cycle. Differences marked by * are significant at $p < 0.05$.

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 party 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

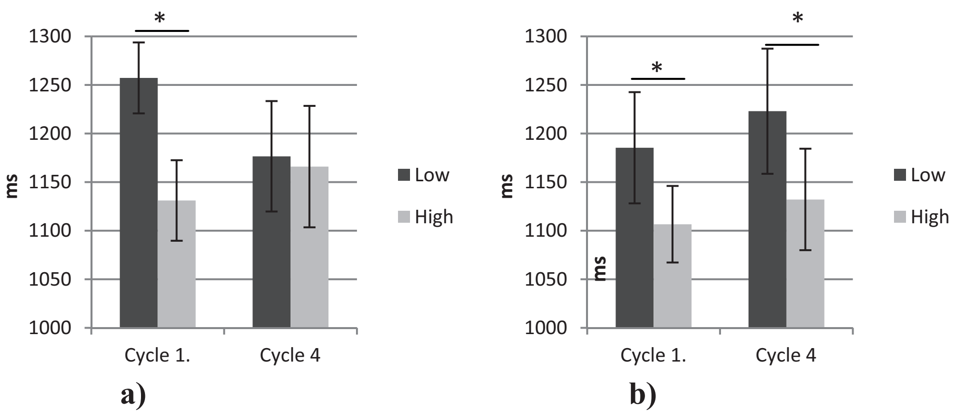


Fig. 4. Mean reaction times by Cycle and Agreement in the a) SLI and b) TD groups. Differences marked by * are significant at $p < 0.05$.

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 (Oppenheim 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 apects 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 *squab* 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

Anderson, J. (1965). Initiatory delay in congenital aphasoid conditions. *Cerebral Palsy Journal*, *26*, 9–12.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*, 344–380.

Bedny, M., Hulbert, J. C., & Thompson-Schill, S. L. (2007). Understanding words in context: the role of Broca's area in word comprehension. *Brain Research*, *1146*, 101–114.

Belke, E. (2013). Long-lasting semantic context effects on object naming are necessarily conceptually mediated. *Journal of Memory and Language*, *69*, 228–256.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in semantic processing as assessed in a blocking paradigm. *Quarterly Journal of Experimental Psychology*, *58*, 667–692.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology*, *66*, 2135–2160.

Biegler, K. A., Crowther, J. E., & Martin, R. C. (2008). Consequences of an inhibition deficit for word production and comprehension: evidence from the semantic blocking paradigm. *Cognitive Neuropsychology*, *25*(4), 493–527.

Bishop, D. (1997). *Uncommon understanding: development and disorders of language comprehension in children*. Psychology Press: Hove, UK.

Bishop, D. V. M. (1983). *Test for reception of grammar*. Manchester, UK: Medical Research Council.
Brown-Schmidt, S. (2009). The role of executive function in perspective-taking during on-line language comprehension. *Psychonomic Bulletin and Review*, *16*, 893–900.
Ceci, S. (1983). Automatic and purposive semantic processing characteristics of normal and language/learning-disabled children. *Developmental Psychology*, *19*, 427–439.
Crosbie, S. L., Howard, D., & Dodd, B. (2004). Auditory lexical decisions in children with specific language impairment. *British Journal of Developmental Psychology*, *22*, 103–121.
Crowther, J., & Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: the role of cognitive control and learning. *Frontiers in Human Neuroscience*, *8*, 1–20.
Csányi, F. I. (1974). *Peabody Szókincs-Teszt [Peabody receptive vocabulary test]*. Budapest: Bárczi Gusztáv Gyógypedagógiai Főiskola.
Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*, B77–B86.
Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition and task switching. *Neuropsychologia*, *44*, 2037–2078.
Dunn, M., & Dunn, L. M. (1981). *Peabody picture vocabulary Test—Revised*. Circle Pines, MN: AGS.
Edwards, J., & Lahey, M. (1996). Auditory lexical decisions of children with specific language impairment. *Journal of Speech and Hearing Research*, *39*, 1263–1273.
Ellis Weismer, S. (2007). Typical talkers, late talkers, and children with specific language impairment: a language endowment spectrum? In R. Paul (Ed.), *Language disorders and development from a developmental perspective: essays in honor of Robin S. Chapman* (pp. 83–101).Mahwah, NJ: Lawrence Erlbaum Associates.
Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in preschool-age children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *52*, 915–929.
Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I., & Trón, V. (2004). Creating open language resources for Hungarian. *Proceedings of language resources and evaluation conference (LREC04). LREC*203–221.
Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W. V., Hoffman, A., et al. (2011). Examining continuity of early expressive vocabulary development: the Generation R study. *Journal of Speech, Language, and Hearing Research*, *54*, 854–869.
Henry, L. A., Messer, D., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry*, *53*(1), 37–45.
January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of stroop and syntactic ambiguity resolution in broca's area: implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience*, *21*(12), 2434–2444.
Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech and Hearing Research*, *37*, 418–421.
Kail, R., & Leonard, L. (1986). Word-finding abilities in language-impaired children. *ASHA Monographs*25.
Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming Cognitive. *Affective & Behavioral Neuroscience*, *4*(1), 43–57.
Kas, B., & Lukács, Á. Magyar Mondatutánmondási Teszt [Hungarian sentence repetition test]. Manuscript in preparation
Katz, W., Curtiss, S., & Tallal, P. (1992). Rapid automatized naming and gesture by normal and language-impaired children. *Brain and Language*, *43*, 623–641.
Kawamoto, A. H., Liu, Q., Mura, K., & Sanchez, A. (2008). Articulatory preparation in the delayed naming task. *Journal of Memory and Language*, *58*(2), 347–365.
Kornai, A., Halácsy, P., Nagy, V., Trón, V., & Varga, D. (2006). Web-based frequency dictionaries for medium density languages. A. Kilgarriff, & M. Baroni (Eds.), *Proceedings of the 2nd international workshop on web as corpus*1–9.
Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
Lahey, M., & Edwards, J. (1996). Why do children with specific language impairment name pictures more slowly than their peers? *Journal of Speech and Hearing Research*, *39*, 1081–1098.
Lahey, M., & Edwards, J. (1999). Naming errors of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 195–205.
Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press 2014.
Leonard, L., Nippold, M., Kail, R., & Hale, C. (1983). Picture naming in language-impaired children. *Journal of Speech and Hearing Research*, *26*, 609–615.
Leonard, L. B., & Deevy, P. (2004). Lexical deficits in specific language impairment. In L. Verhoeven, & H. von Balkom (Eds.), *Classification of Developmental Language Disorders: theoretical issues and clinical implications* (pp. 209–233).Mahwah, NJ: Lawrence Erlbaum Associates.
Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research*, *50*, 408–428.
Lukács, Ladányi, E., Fazekas, K., & Kemény, F. (2015). Executive functions and the contribution of short-term memory span in children with specific language impairment. *Neuropsychology* [Advance online publication. http://dx.doi.org/10.1037/neu0000232].
Lukács, A., Győri, M., & Rózsa, S. (2012). A TROG pszichometriai jellemzőinek magyar vizsgálata, a normák kialakítása [The psychometric analysis of Hungarian data from the TROG]. In D. V. M. Bishop (Ed.), *TROG—Test for reception of grammar. handbook* (pp. 47–86).Budapest, Hungary: OS Hungary.
Lum, J., & Bavin, E. L. (2007). A study of grammaticality judgments by children with and without SLI. *Journal of Speech, Language and Hearing Research*, *50*, 1618–1630.
Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2008). Lexical Representations in Children with SLI: Evidence from a Frequency-Manipulated Gating Task. *Journal of Speech, Language, and Hearing Research*, *51*, 381–393.
Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2010). Explaining lexical-semantic deficits in specific language impairment: the role of Phonological similarity, phonological working memory, and lexical competition. *Journal of Speech, Language, and Hearing Research*, *53*, 1742–1756.
McGregor, K. K., & Leonard, L. (1995). Intervention for word-finding deficits in children. In M. Fey, J. Windsor, & S. Warren (Eds.), *Language intervention: Preschool through the elementary years* (pp. 85–105).Baltimore: Paul H. Brookes.
McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *45*, 998–1014.
McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: implications for SLI. *Cogntive Psychology*, *60*(1), 1–39.
Miller, C., Kail, R., Leonard, L., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 416–433.
Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
Montgomery, J., & Leonard, L. (1998). Real-time inflectional processing by children with specific language impairment: effects of phonetic substance. *Journal of Speech, Language, and Hearing Research*, *41*, 1432–1443.
Montgomery, J., Scudder, R., & Moore, C. (1990). Language-impaired children's real-time comprehension of spoken language. *Applied Psycholinguistics*, *11*, 273–290.
Navarrete, E., Del Prato, P., & Mahon, B. Z. (2012). Factors determing semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology*, *3*, 38.
Norbury, C. F. (2005). Barking up the wrong tree?: Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology*, *90*, 142–171.

Noterdaeme, M., Amorosa, H., Mildenberger, K., Sitter, S., & Minow, F. (2001). Evaluation of attention in children with autism and children with a specific language disorder. *European Journal of Child Psychology*, *10*, 58–66.
Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, *26*(6), 527–567.
Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: reexamining the role of broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 263–281.
Novick, J. M., Trueswell, J. C., & Thomspon-Schill, S. L. (2010). Broca's area and language processing: evidence for the cognitive control connection. *Language and Linguistics Compass*, *4*(10), 906–924.
Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*, 227–252.
Poll, G. H., & Miller, C. A. (2013). Late talking, typical talking: and weak language skills at middle childhood. *Learning and Individual Differences*, *26*, 177–184.
Racsmány, M., Lukács, Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai [Hungarian methods for studying verbal working memory]. *Magyar Pszichológiai Szemle*, *60*, 479–505.
Raven, J., Court, J., & Raven, J. (1987). *Raven's progressive matrices and raven's colored matrices*. London: H. K. Lewis.
Rescorla, L. (2011). Late talkers: do good predictors of outcome exist? *Developmental Disabilities Research Reviews, Special Issue on Developmental Delay*, *17*, 141–150.
Rice, M., Buhr, J., & Nemeth, M. (1990). Fast mapping word-learning abilities of language delayed preschoolers. *Journal of Speech and Hearing Disorders*, *55*, 33–42.
Rice, M., Buhr, J., & Oetting, J. (1992). Specific-language-impaired children's quick incidental learning of words: the effect of a pause. *Journal of Speech and Hearing Research*, *35*, 1040–1048.
Rice, M., Oetting, J., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research*, *37*, 106–122.
Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked cyclic naming: evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227.
Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, *106*(1), 322–327.
Seiger-Gardner, L., & Brooks, P. (2008). Phonological processing of onset and rhyme-related distractors in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *51*, 1263–1281.
Sheng, L., & McGregor, K. K. (2010). Lexical-semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *53*, 146–159.
Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *53*(3), 725–738.
SPSS Inc. Released 2009. PASW Statistics for Windows, Version 18.0. Chicago: SPSS Inc.
Stark, R., & Montgomery, J. (1995). Sentence processing in language-impaired children under conditions of filtering and time compression. *Applied Psycholinguistics*, *16*, 137–154.
Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 1275–1278.
Trauner, D., Wulfeck, B., Tallal, P., & Hesselink, J. (1995). *Neurologic and MRI profiles of language impaired children (Technical Report CND-9513)*. Center for Research in Language: University of California at San Diego.
Watkins, R., Kelly, D., Harbers, H., & Hollis, W. (1995). Measuring children's lexical diversity: differentiating typical and impaired language learners. *Journal of Speech and Hearing Research*, *38*, 1349–1355.
Wiegel-Crump, C. A., & Dennis, M. (1986). Development of word-finding. *Brain and Language*, *27*, 1–23.
Wiig, E., Semel, E., & Nystrom, L. (1982). Comparison of rapid naming abilities in language-learning-disabled and academically achieving eight-year-olds. *Language, Speech, and Hearing Services in Schools*, *13*, 11–23.
Windsor, J., & Hwang, M. (1999a). Testing the generalized slowing hypothesis in specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 1205–1218.
Windsor, J., & Hwang, M. (1999b). Children's auditory lexical decisions: a limited processing capacity account of language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 990–1002.
Windsor, J., & Hwang, M. (1999c). Testing the generalized slowing hypothesis in specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 1205–1218.

Study 12.

Lukács Á., Kemény F., Lum, J.A.G.& Ullman, M.T. (2017). Learning and overnight retention in declarative memory in specific language impairment. *PLoS ONE* 12(1): e0169474.
doi: 10.1371/journal.pone.0169474

Thesis 7. In concert with the Procedural Deficit hypothesis, procedural learning is vulnerable in SLI, while processes of declarative learning and retention are relatively intact (3, 8, 12)

PLOS | ONE

RESEARCH ARTICLE

Learning and Overnight Retention in Declarative Memory in Specific Language Impairment

Ágnes Lukács[1]*, Ferenc Kemény[1,2], Jarrad A. G. Lum[3], Michael T. Ullman[4]*

1 Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary, **2** Institute of Psychology, University of Graz, Graz, Austria, **3** Cognitive Neuroscience Unit, School of Psychology, Deakin University, Melbourne, VIC, Australia, **4** Brain and Language Lab, Department of Neuroscience, Georgetown University, Washington, DC, United States of America

* alukacs@cogsci.bme.hu (ÁL); michael@georgetown.edu (MU)

OPEN ACCESS

Citation: Lukács Á, Kemény F, Lum JAG, Ullman MT (2017) Learning and Overnight Retention in Declarative Memory in Specific Language Impairment. PLoS ONE 12(1): e0169474. doi:10.1371/journal.pone.0169474

Editor: Etsuro Ito, Waseda University, JAPAN

Received: September 8, 2016

Accepted: December 17, 2016

Published: January 3, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Hungarian National Science Foundation (OTKA) under OTKA K 83619 'Nonlinguistic abilities in Specific Language Impairment' to Ágnes Lukács, and by the National Institutes of Health under RO1 HD049347 and the Mabel H. Flory Trust to Michael Ullman. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

We examined learning and retention in nonverbal and verbal declarative memory in Hungarian children with ($n = 21$) and without ($n = 21$) SLI. Recognition memory was tested both 10 minutes and one day after encoding. On nonverbal items, only the children with SLI improved overnight, with no resulting group differences in performance. In the verbal domain, the children with SLI consistently showed worse performance than the typically-developing children, but the two groups showed similar overnight changes. The findings suggest the possibility of spared or even enhanced declarative memory consolidation in SLI.

Introduction

An increasing awareness of the importance of memory systems in language has inspired researchers to examine these systems in Specific Language Impairment (SLI). Much of the memory research in SLI has focused on working memory, suggesting impairments in this domain, in particular of verbal working memory [1–4]. More recently however, research has begun to examine long-term memory systems, in particular procedural and declarative memory. In this spirit, the Procedural Deficit Hypothesis (PDH) posits that the pattern of language and other deficits in SLI can be largely explained by abnormalities of brain structures underlying procedural memory, in particular frontal and basal ganglia structures [5,6]. Crucially, the PDH also proposes that declarative memory generally remains relatively spared or even enhanced in the disorder, and that it plays important compensatory roles for grammatical and other impairments [6,7]. Although an increasing number of studies of SLI have focused on procedural memory [8], declarative memory has received much less attention.

Declarative memory

The declarative memory system is rooted in the hippocampus and other medial temporal lobe structures [9–12]. The system underlies the learning, consolidation (the stabilization of memories

Competing Interests: The authors have declared that no competing interests exist.

after their initial acquisition), storage, and use of multiple types of knowledge, including of facts (semantic knowledge), events (episodic knowledge), and words (lexical knowledge). Learning in declarative memory can be very fast, and can occur after as little as a single presentation of the stimulus. The system underlies not only explicit knowledge, that is, available to conscious awareness, but also implicit knowledge [10,13,14]. The functional neuroanatomy of the system has been quite well studied. Briefly, the hippocampus and other medial temporal lobe structures are critical for learning and consolidating new knowledge that depends on this system, though ultimately the long-term storage of this knowledge depends largely on neocortical regions, particularly in the temporal lobes [9–12,15]. Note that the declarative memory systems refers here to the entire neurocognitive system involved in the learning, consolidation, representation, and use of the relevant knowledge, not just to those portions underlying learning and consolidating new knowledge, which is how some researchers refer to the system.

Declarative memory in SLI

Procedural memory has been increasingly well studied in SLI, with numerous studies now suggesting deficits in this domain, including in consolidation [2,16–20]. For a recent meta-analysis of procedural memory in SLI, which reveals clear impairments in this domain, see Lum, Conti-Ramsden, Morgan and Ullman [8]. However, we still know very little about the status of declarative memory in the disorder.

In the nonverbal domain, studies testing declarative memory have generally revealed largely intact performance in SLI. No differences have been reported between SLI and typically developing (TD) groups in a range of tasks probing a variety of types of nonverbal stimuli, including abstract visual and spatial information, faces, and complex nonverbal sounds [2,21–25]. When such tasks have used easily verbalizable items (e.g., pictures of everyday events), SLI deficits have been found in some [2,23] but not other [21,24] studies. Such impairments may be due to the association of these items with language [26]. Type of task also seems to have an influence. For example, Kuppuraj et al. [25] found intact nonverbal declarative memory performance in SLI with incidental encoding and later recognition (similar to the approach employed here), while performance lagged behind controls with intentional encoding and later recall.

In the verbal domain, studies have generally–but not always [21,27,28]–found impairments in tasks probing declarative memory [2,21–24,27–31]. However, most of these studies used list learning tasks, which rely heavily on working memory (due to the repetition of items during the learning phase). Since verbal working memory is often impaired in SLI [1–4,6], any deficits at these tasks could be due to problems with working memory rather of declarative memory itself [2,6]. Indeed, verbal declarative memory deficits observed in a range of tasks in Lum et al. [2] were reduced or eliminated after covarying out working memory (and were completely absent after language abilities were controlled for). More recently, Lum, Ullman, and Conti-Ramsden [32] found that in a list learning task only those children with SLI who had working memory deficits showed impairments at declarative memory; those without working memory problems showed normal performance at the task. Together, the data suggest that declarative memory deficits in SLI may be due largely, if not entirely, to accompanying working memory deficits, as well as to their underlying language impairments.

However, almost all studies of declarative memory in SLI have focused on the initial stages of learning, in which the acquired knowledge is typically tested after a short delay of minutes following encoding (enough time to reduce the likelihood of maintaining the information in working memory). But subsequent *retention* of that knowledge is also critical, since the goal of learning is generally to retain information beyond the range of a few minutes. This is especially the case, of course, with language. To our knowledge, however, there is only one study that specifically examined longer-term retention, i.e., one day or more after initial learning [33]. In this study, word learning was tested in a heterogeneous group of young adults with different forms of language impairment (with previous diagnoses not only of "language impairment", but also of dyslexia, learning disability, dysnomia, or a combination of these) and was compared to a TD group of the same age. The study was designed to separately examine learning word forms, word referents, and links between form and referent. Retrieval of items was tested immediately after training; after 12 hours (either involving sleep or not); 24 hours after training; and one week later. The main result relevant to retention was that after delays involving sleep the language-impaired group did not differ from the TD group at remembering referents or form-referent links, but did show impairments at retrieving word forms. The authors concluded that the "consolidation of declarative memory is a relative strength for young adults with LI [language impairment]", though primarily when the items were not purely verbal in nature, perhaps since learning novel word forms may also depend on procedural memory [6]. However, weaknesses of the study involving participant criteria suggest caution in generalizing the results. Moreover, a dearth of direct group comparisons in the paper makes it difficult to interpret the results. Finally, purely nonverbal information was not tested, since even the nonverbal referents were learned initially in the context of the novel words that referred to them. Together with the fact that this is the only study to date to examine retention in SLI suggests that further studies are needed.

The present study

The aim of the present study was to examine both learning and retention in children with SLI, as compared to TD children, in both nonverbal and verbal domains. The study was designed to minimize the influence of functions that can affect performance on declarative memory tasks, and are often impaired in SLI, namely working memory and free recall–both of which depend on frontal/basal ganglia circuits, which appear to be abnormal in SLI [6,34]. We therefore tested declarative memory with a recognition memory task, following incidental encoding. To test initial learning we probed recognition memory 10 minutes after encoding. Both nonverbal items (pictures of real and novel objects) and verbal items (auditorily presented real and novel words) were examined. To test for retention we probed recognition memory of the same items 24 hours later.

Consistent with the PDH and previous studies, we predicted largely intact recognition memory for nonverbal items in SLI at initial learning (after the short delay of 10 minutes). In contrast, we expected the possibility of impairments at recognition memory for verbal items at this time point. Given the dearth of previous evidence regarding retention in declarative memory in SLI, we had no clear predictions for recognition memory following the overnight delay.

Materials and Methods

The Budapest University of Technology and Economics Behavioral and Biomedical Institutional Review Board reviewed and approved the study (IRB #: IRB00004964 Project Title: Nonlinguistic abilities in Specific Language Impairment). All children were tested with the informed written consent of their parents (by asking them individually to sign a detailed consent form), in accordance with the principles set out in the Declaration of Helsinki, and approved by the Budapest University of Technology and Economics Behavioral and Biomedical Institutional Review Board. In approving this Research Project, the Review Board followed the requirements of the Common Rule and the Helsinki Agreement.

Participants

Children with SLI were recruited from two special schools for children with language impairment, which were in and near Budapest, Hungary. The children had been referred to these schools by speech and language therapists working in clinical practice. Recruitment and screening for this study lasted between 2 and 3 months at each school. No eligible children declined participation. All children in the SLI group met the following inclusion and exclusion criteria, which are commonly used in SLI research [1,35]. They scored at least 1.5 standard deviations (*SDs*) below the mean for their age on at least two of the following four language screening tests: (1) a Hungarian version of the Peabody Picture Vocabulary Test [36,37]; (2) the Hungarian version of the Test for the Reception of Grammar [38,39]; (3) the Hungarian Sentence Repetition Test [40]; and (4) a Hungarian nonword repetition test [41]. Their nonverbal IQ as measured by the Raven Progressive Matrices (RPM) was in the normal range, corresponding to a score above 85 IQ points [42]. Their hearing was also assessed as normal. Typically developing children were recruited from two regular schools, which had no special selection processes for children. All TD children scored normally, that is, above 1.5 *SDs* below the mean for their age, on all four language screening tests. The TD children were matched individually (pair-wise) to children with SLI on chronological age and sex; the two groups were also matched group-wise on nonverbal IQ [42]. None of the TD children were known to have been diagnosed with any neurodevelopmental, psychiatric, or neurological disorder. Similarly, none of the children with SLI had any known current or past neurodevelopmental, psychiatric, or neurological disorders other than SLI; e.g., children with comorbid autism or ADHD were excluded. A total of 42 children were tested: 21 with SLI and 21 who were typically developing (TD). Demographic and screening data for the two groups are shown in Table 1.

Procedure

Declarative memory was tested with a recognition memory task developed by the Brain and Language Lab at Georgetown University, and modified as appropriate for the Hungarian

Table 1. Demographic and screening data for the two groups.

	SLI	TD	Group Comparisons
n	21	21	
Sex	15M, 6F	15M, 6F	
Age (years)	8.89 (1.06)	8.85 (1.03)	$F(1, 40) = 0.02$, $p = .884$
Vocabulary (PPVT; raw scores)	98.00 (19.60)	124.43 (12.50)	$F(1, 40) = 27.14$, $p < .001$
Grammar (TROG; blocks raw score)	13.52 (2.02)	18.05 (1.47)	$F(1, 40) = 69.21$, $p < .001$
Sentence Repetition (raw scores)	20.24 (8.33)	37.67 (2.80)	$F(1, 40) = 82.60$, $p < .001$
Nonword Repetition (span)	3.29 (1.27)	6.48 (0.98)	$F(1, 40) = 82.98$, $p < .001$
Nonverbal IQ (RPM; standard score)	103.90 (9.82)	107.38 (11.19)	$F(1, 40) = 1.14$, $p = .291$

Note. Means (and standard deviations) are shown for each variable. Results from one-way ANOVAs are shown for group differences. SLI: specific language impairment; TD: typically developing; M: male; F: female. Vocabulary scores are computed from the PPVT (Peabody Picture Vocabulary Test), grammar scores from the TROG (Test for the Reception of Grammar), and nonverbal IQ scores from the RPM (Raven's Progressive Matrices); see main text.

doi:10.1371/journal.pone.0169474.t001

participants in the present study. The task assesses encoding, initial learning (recognition after a short delay of 10 minutes) and retention (recognition again after a delay of 1 day) in declarative memory. Separate subtasks examine nonverbal and verbal learning, using two sets of stimuli: the nonverbal subtask visually presents pictures of real and novel objects, while the verbal subtask auditorily presents words and nonwords.

Each of the two subtasks consists of three phases. First, in the Encoding phase, participants are presented with 32 real and 32 novel items; that is, pictures of 32 real and 32 made-up objects in the nonverbal subtask, or 32 real and 32 made-up words in the verbal subtask. Participants are asked to make a real/novel decision on each item, that is, an object decision in the nonverbal subtask, and a lexical decision in the verbal subtask; see below for further details. This incidental encoding task is followed by an initial Recognition phase after a 10-minute delay, and a Retention phase after a 24-hour delay. These two phases are virtually identical. Both phases present all 64 target items that were seen or heard in the encoding phase (old items), together with 64 foils (new items), for a total of 128 items. Half the old items and half the new items are real (objects or words) and half are novel (novel objects or nonwords). The foils (new items) are entirely new in both the Recognition and Retention phases; that is, the foils in the Retention phase are foils that were not presented previously in the Recognition phase.

Stimuli were presented on a Lenovo z61m PC laptop running Windows 7, using E-Prime 1.2 [43]. A display resolution of 1024 x 768 pixels was used. The objects in the nonverbal subtask were presented as 640 x 480 pixel pictures. Children sat approximately 40–60 centimeters from the screen. Testing took place in a quiet room in the children's school. For the verbal subtask, the stimuli were presented via headphones to further decrease noise. Responses were made on the left and right buttons located just below the touchpad on the laptop (see below for Instructions).

Items in the nonverbal subtask were presented with the following presentation and timing parameters in all three phases (Encoding, Recognition, Retention). A 1000 millisecond (msec) preparation period with a fixation cross at the center of the screen signaled the imminent presentation of each new item; during the first 200 msec of this preparation period a tone was also presented. After this 1000 msec preparation period, the picture appeared in the center of the screen for 500 msec. If the participant responded during this 500 msec presentation period, the item disappeared. Following the disappearance of the picture (at or before 500 msec), the fixation cross reappeared on the screen. If the participant responded prior to 5000 msec after the appearance of the picture on the screen (i.e., during the allowable response period, including during the initial 500 msec), the next item was when the experimenter pressed a mouse button, at which point the 1000 msec preparation period for the next item began. If the participant did not respond within the 5000 msec response period, then at 5000 msec a 1000 msec time-out period occurred (a 400 msec time-out tone, together with a fixation cross which lasted the full 1000 msec; note that the time-out tone had a different frequency from the tone in the preparation period), after which the 1000 msec presentation period for the next item began.

The presentation and timing parameters for the words/nonwords in the verbal subtask were identical to those for the real/novel objects in the nonverbal task except that the presentation duration of the stimulus (word/nonword) was variable (rather than the consistent 500 msec duration in the nonverbal subtask), lasting the duration of the sound file of the item. As in the nonverbal subtask, the response period was 5000 msec from the onset of the stimulus.

The following instruction procedures were followed for both subtasks. All instructions were given in Hungarian. Each of the three phases (Encoding, Recognition, Retention) began with instructions, which were presented on the screen and also read out and explained by the experimenter. Before all three phases, participants were instructed to place their left and right index fingers on the left and right buttons located just below the touchpad on the laptop, and to make a response by pressing one of these buttons. In the Encoding phase, participants were instructed to decide as quickly and accurately as possible whether the object or word was Real or Made-up, and to press one of two buttons accordingly. In both the Recognition and Retention phases, participants were instructed to indicate by button press whether the item they were presented with was or was not previously presented in the Encoding phase (Yes/No decision on whether they had seen or heard the item earlier). Instructions were followed by 3 practice items in the encoding phase, and 6 practice items in both the Recognition phase (3 old, that is, the three practice items from encoding; and 3 new, that is, that were not presented as practice or experimental items during Encoding) and the Retention phase (3 old, the same practice items as in Encoding and Recognition; and 3 new, which had not been presented during either Encoding or Recognition).

For each subtask two versions (A and B) were created. In one version the left button was associated with Real (Encoding) or Yes, presented earlier (Recognition and Retention), and the left button with Made-up/No, while in the other group, the mapping was the reverse. The two versions were alternatively assigned to consecutive participants within each participant group (SLI and TD). During the task, a reminder appeared at the bottom of the screen indicating the mapping of the buttons.

Participants (within each group and version) were randomly assigned to one of two subtask orders. In one order, participants were first given the Encoding phase of the verbal subtask, followed by the Encoding phase of the nonverbal subtask, then the Recognition phase of the verbal subtask, then the Recognition phase of the nonverbal subtask, and 24 hours later the Retention phases of first the verbal then the nonverbal subtask. In the other order subjects were given first the nonverbal and then the verbal subtask for each phase. Participants were given the Encoding and Recognition phases at varying times, between about 9am and 4pm, and were tested on Retention about 24 hours later. Encoding phases took 5–7 minutes to complete, and Recognition and Retention phases took between 9–11 minutes to complete. Participants took a short self-paced break between the two Encoding phases. The delay between each Encoding phase and its corresponding Recognition phase was about 10 minutes.

Stimuli

The real and made-up object stimuli in the nonverbal task were identical to the stimuli developed in the original version of the task developed in the Brain and Language Lab. These items were validated as real and novel objects by the Hungarian experimenters. The items were black and white line drawings of real and made-up objects. Images of real objects had been taken from various sources, and then modified as necessary. Sources included a number of different clipart galleries (including free websites and purchased collections) and the Snodgrass and Vanderwart [44] set of pictures. Images for made-up objects had been taken from Eals and Silverman [45], Cornelissen et al. [46] and Williams and Tarr [47]; they were then modified as necessary, including to reduce nameability. Low nameability was confirmed in previous pilot studies run in the Brain and Language Lab. All images were resized, touched up, rotated, and/or converted to black-and-white to create the final set of stimuli produced by the Brain and Language Lab. The items were presented in a pseudo-randomized order, with no more than 3 consecutive real or made-up objects.

Stimuli in the verbal subtask comprised auditorily presented real words (concrete nouns) and made-up words (Hungarian adaptations of the English nonwords created by the Brain and Language Lab). The nonwords were matched to the real words in phonological length (range: 1–5 syllables) and syllable structure (e.g., a CVCC word was matched as closely as possible to a CVCC nonword) for all three sets of real/made-up words: the target items, the foils presented during Recognition, and the foils presented during Retention. Additionally, these three sets were matched on the same factors. All items conformed to the rules of Hungarian phonotactics. All words and nonwords were digitally recorded by a male native Hungarian speaker in 16 bit, 44100 Hz, 705 kbps mono–not stereo–wave format (though these were later presented to participants to both ears via the headphones). All sound files were edited to reduce their length to the length of the word or nonword. The length of the sound files ranged between 465 and 1311 msec. See Fig 1 for examples of stimuli in both the nonverbal and verbal subtasks.

Results

Between and within group differences in recognition memory were examined with 2 (Group: SLI vs. TD) X 2 (Real/Novel: Real vs. Novel) X 2 (Delay: 10 minutes/Recognition vs. 24 hours/Retention) mixed ANOVAs run separately on the Nonverbal and Verbal subtasks. In these analyses Group was the between-subject factor, and Real/Novel and Delay were within-subject factors. In order to minimize effects of response bias, accuracy was entered in the analyses as normalized d' (d-prime) scores (d' = z(hit rate)—z(false alarm rate)). For all analyses, we report partial eta-squared (η_p^2) as a measure of effect size. Reaction times were also analyzed, using median RTs (based on raw RTs for correct responses only); however, these analyses are not presented here, since there were no main effects of Group or interactions with Group (all ps > 0.1).

Recognition and retention of nonverbal information

Performance on the Nonverbal subtask is shown in Table 2. Both the SLI and TD groups showed above-chance performance at both Real and Novel items, for both Recognition and Retention (Table 2).

The 2 (Group) X 2 (Real/Novel) X 2 (Delay) ANOVA yielded main effects of Delay ($F(1, 40)$ = 11.774, p = .001, η_p^2 = .227), with better performance at Retention than Recognition over both groups, and of Real/Novel ($F(1, 40)$ = 95.450, $p < .0001$, η_p^2 = .705), with Real objects recognized better than Novel ones. There was no main effect of Group ($F(1, 40)$ = 1.63, p = .209, η_p^2 = .039). However, the two significant main effects were qualified by a significant interaction between Group and Delay ($F(1, 40)$ = 1.478, p = .034, η_p^2 = .107) and an interaction between Group and Real/Novel that approached significance ($F(1, 40)$ = 1.575, p = .057, η_p^2 = .088); see Figs 2 and 3. There were no other interactions (Real/Novel X Delay; $F(1, 40)$ = .433, p = .514, η_p^2 =. 011; Group X Real/Novel X Delay: $F(1, 40)$, = .651, p = .424, η_p^2 = .016).

We first followed up on the Group X Delay interaction, collapsing over both object types (Real and Novel). There was no effect of Delay in the TD group; that is, there was no significant difference in accuracy for the TD children between Recognition and Retention ($F(1, 20)$ = 2.004, p = .172, η_p^2 = .091). However, for the SLI group, accuracy was significantly higher at Retention than Recognition ($F(1, 20)$ = 20.340, $p < .001$, η_p^2 = .504). Additionally, the SLI group was worse than the TD group at Recognition ($F(1, 40)$ = 4.211, p = .047), but not at Retention ($F(1, 40)$ = .133, p = .718). See Fig 2. S1 Fig shows d' performance at Recognition and Retention at the nonverbal subtask for each child with SLI (S1 (A)) and each TD child (S1 (B)).

Next, we followed up on the Group X Real/Novel interaction that approached significance, collapsing over both Delay periods (10 minutes/Recognition and 24 hours/Retention). The effect of Real/Novel was significant in both the TD and SLI groups, with better performance (i.e., collapsed over both Delay periods) on Real than Novel objects in both groups, though this effect was larger for the TD children (TD: $F(1, 20)$ = 79.707, $p < .0001$, η_p^2 = .799; SLI: $F(1, 20)$ = 49.626, $p < .0001$, η_p^2 = .713). Additionally, the effect of Group (again, collapsed over both Delay periods)

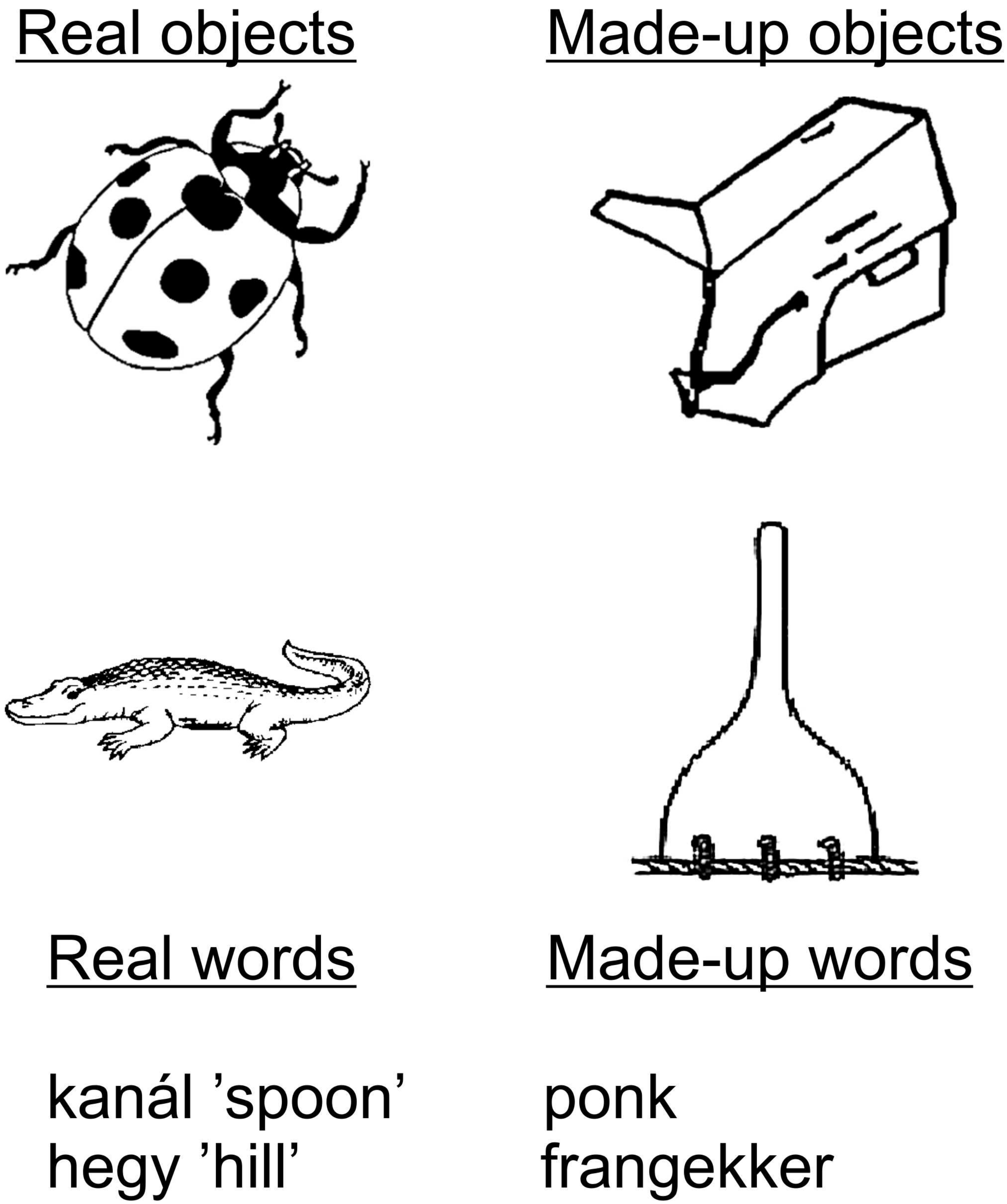


Fig 1. Example stimuli from the nonverbal subtask (real and made-up objects) and the verbal subtask (real and made-up words).

doi:10.1371/journal.pone.0169474.g001

Table 2. Recognition and retention accuracy for nonverbal information.

	SLI	TD
	Recognition (10 minute delay)	
Real		
d'	1.41 (1.12) ***	2.17 (1.05) ***
Hit rate	0.77 (0.20)	0.78 (0.19)
False alarm rate	0.33 (0.32)	0.14 (0.15)
Novel		
d'	0.65 (0.75) **	0.92 (0.78) ***
Hit Rate	0.38 (0.23)	0.38 (0.16)
False Alarm Rate	0.21 (0.25)	0.16 (0.17)
	Retention (24 hour delay)	
Real		
d'	1.90 (1.21) ***	2.18 (0.95) ***
Hit rate	0.74 (0.23)	0.75 (0.22)
False alarm rate	0.20 (0.24)	0.11 (0.10)
Novel		
d'	1.12 (0.76) ***	1.12 (0.94) ***
Hit Rate	0.45 (0.23)	0.38 (0.21)
False Alarm Rate	0.15 (0.21)	0.10 (0.11)

Note. Accuracy in the Nonverbal subtask, showing means (and standard deviations) of *d'*, as well as of hit rates and false alarm rates. SLI: specific language impairment; TD: typically developing. Asterisks indicate performance greater than chance (mean *d'*s significantly greater than zero, one-sample t-tests, df = 20):
***: p < .001
**: p < .01
*: p < .05.

doi:10.1371/journal.pone.0169474.t002

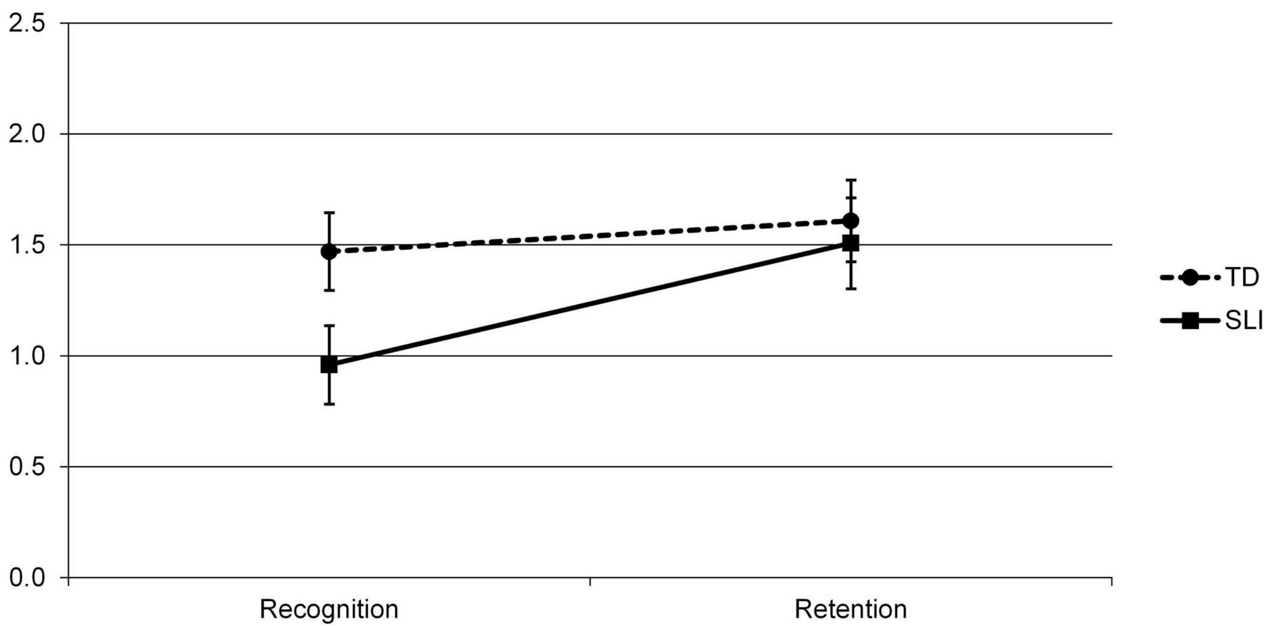


Fig 2. Nonverbal subtask performance by Group (SLI vs. TD) and Delay (10 minutes/Recognition vs. 24 hours/Retention), showing mean *d'* scores and standard errors.

doi:10.1371/journal.pone.0169474.g002

showed a trend for Real objects, with somewhat better performance by the TD than the SLI group ($F(1, 40) = 2.947$, $p = .094$, $\eta_p^2 = .069$), while it there was no group difference for Novel objects, $F(1, 40) = .256$, $p = .616$, $\eta_p^2 = .006$. See Fig 3.

To understand, at a more fine-grained level, the increase in performance between Recognition and Retention found for the SLI group but not the TD group, we tested whether this pattern held separately for Real and Novel items. Note that although the three-way interaction between Group, Real/Novel and Delay was not significant, such higher-level interactions are difficult to obtain, leaving open the possibility that the observed patterns for the SLI and TD groups could differ between Real and Novel items. However, consistent with the lack of a three-way interaction, the SLI group showed an increase in performance between Recognition and Retention for both types of items (Real: $F(1, 20) = 7.745$, $p = .011$, $\eta_p^2 = .279$; Novel: $F(1, 20) = 11.65$, $p = .003$, $\eta_p^2 = .368$), whereas the TD group did not show an increase for either (Real: $F(1, 20) = 0.20$, $p = .889$, $\eta_p^2 = .001$; Novel: $F(1, 20) = 1.412$, $p = .249$, $\eta_p^2 = .066$). Interestingly, however, the two groups did not differ for either Real or Novel items at either Delay period, with the notable exception of Real items at Recognition, where the TD group showed superior performance (Recognition: Real: $F(1, 40) = 5.150$, $p = .029$, $\eta_p^2 = .114$; Novel: $F(1, 40) = 1.349$, $p = .252$, $\eta_p^2 = .033$. Retention: Real: $F(1, 40) = .724$, $p = .400$, $\eta_p^2 = .018$; Novel: $F(1, 40) = .000$, $p = .998$, $\eta_p^2 = .000$).

Recognition and retention of verbal information

Performance on the Verbal subtask is shown in Table 3. Both the SLI and TD groups showed above-chance performance at both Real and Novel items, for both Recognition and Retention, with the exception of the SLI group at Novel items at Recognition (Table 3).

The 2 (Group) X 2 (Real/Novel) X 2 (Delay) ANOVA yielded main effects of Group ($F(1, 40) = 21.86$, $p = .0001$, $\eta_p^2 = .353$), with better overall performance by the TD children than children with SLI, and Real/Novel ($F(1, 40) = 24.173$, $p < .0001$, $\eta_p^2 = .377$), with Real words recognized

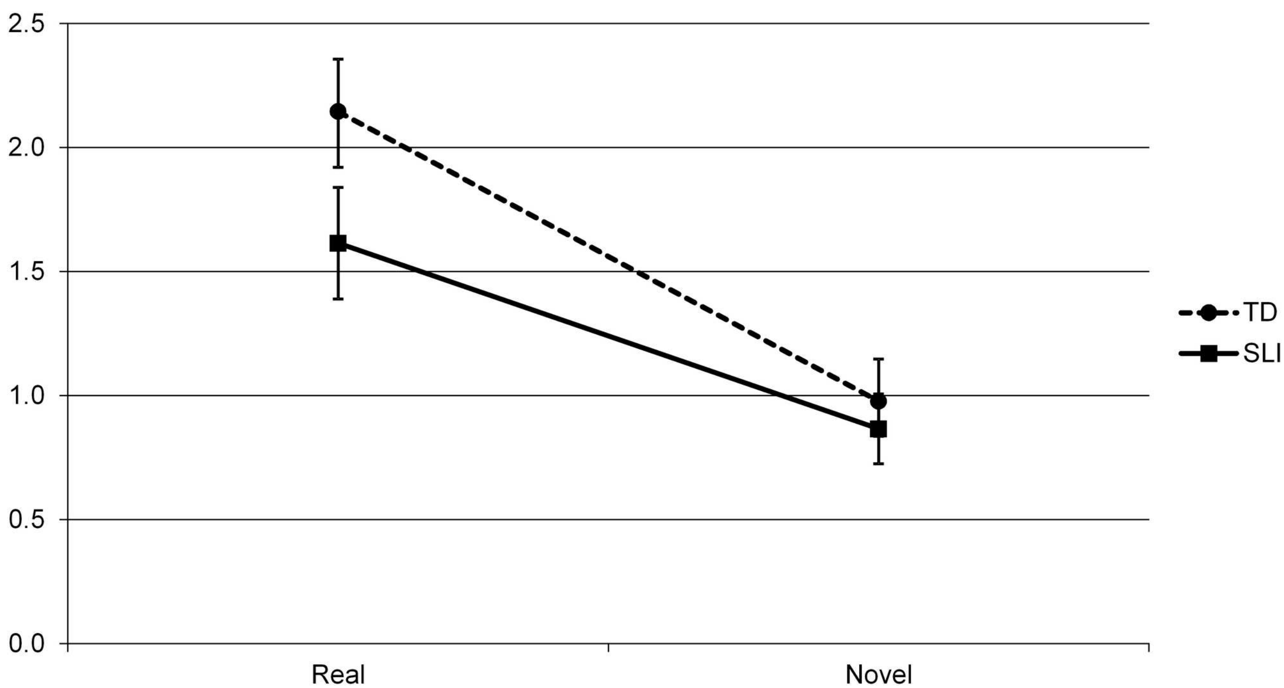


Fig 3. Nonverbal subtask performance by Group (SLI vs. TD) and Real/Novel, showing mean *d'* scores and standard errors.

doi:10.1371/journal.pone.0169474.g003

Table 3. Recognition and retention accuracy for verbal information.

	SLI	TD
	Recognition (10 minute delay)	
Real		
d'	0.83 (0.83) ***	1.51 (0.49) ***
Hit Rate	0.71 (0.17)	0.74 (0.09)
False Alarm Rate	0.42 (0.31)	0.22 (0.13)
Novel		
d'	0.05 (0.61)	0.91 (0.62) ***
Hit Rate	0.28 (0.17)	0.42 (0.20)
False Alarm Rate	0.27 (0.23)	0.16 (0.14)
	Retention (24 hour delay)	
Real		
d'	0.65 (0.65) ***	1.14 (0.69) ***
Hit Rate	0.55 (0.27)	0.64 (0.21)
False Alarm Rate	0.35 (0.27)	0.27 (0.21)
Novel		
d'	0.32 (0.43) **	1.09 (0.67) ***
Hit Rate	0.33 (0.17)	0.52 (0.23)
False Alarm Rate	0.05 (0.03)	0.06 (0.03)

Note. Accuracy in the verbal subtask, showing means (and standard deviations) of *d'*, as well as of hit rates and false alarm rates. SLI: specific language impairment; TD: typically developing. Asterisks indicate performance greater than chance (mean *d'*s significantly greater than zero, one-sample t-tests, df = 20)

***: p < .001

**: p < .01

*: p < .05.

doi:10.1371/journal.pone.0169474.t003

better than Novel ones. There was no main effect of Delay ($F(1, 40) = .148$, $p = .703$, $\eta_p^2 = .004$). However, the main effect of Real/Novel was qualified by a significant Real/Novel X Delay interaction ($F(1, 40) = 20.359$, $p < .0001$, $\eta_p^2 = .337$, see Fig 4 below). No other interactions were significant (Group X Real/Novel: $F(1, 40) = 1.700$, $p = .200$, $\eta_p^2 = .041$; Group X Delay: $F(1, 40) = .952$, $p = .335$, $\eta_p^2 = .023$; Group X Real/Novel X Delay: $F(1, 40) = .221$, $p = .641$, $\eta_p^2 = .005$).

To investigate the Real/Novel X Delay interaction, we examined the effect of Delay separately for Novel and Real words, collapsing across the two participant groups. For both types of words there were significant effects of Delay, but in opposite directions for the two types of words. Whereas performance on Novel words increased between Recognition and Retention ($F(1, 41) = 5.487$, $p = .024$, $\eta_p^2 = .118$), performance on Real words decreased during this interval ($F(1, 41) = 12.099$, $p = .001$, $\eta_p^2 = .228$). Additionally, the effect of Real/Novel was significant at Recognition ($F(1, 41) = 38.935$, $p < .0001$, $\eta_p^2 = .487$), but only approached significance at Retention ($F(1, 41) = 3.480$, $p = .069$, $\eta_p^2 = .078$)–though in both cases performance on Real words was higher than on Novel words. See Fig 4.

Comparison of performance on the nonverbal and verbal subtasks

To compare overall performance on nonverbal and verbal items directly, we ran a 2 (Group: SLI vs. TD) X 2 (Modality: Nonverbal vs. Verbal) X 2 (Delay: 10 minutes/Recognition vs. 24 hours/Retention) mixed ANOVA on normalized *d'* scores (for simplicity, averaged over Real and Novel items). This analysis yielded main effects of Modality ($F(1,40) = 43.721$, $p < 0.001$, $\eta_p^2 = 0.522$), with better performance on the nonverbal than verbal subtask, of Delay ($F(1,40) = 7.688$, $p = 0.008$,

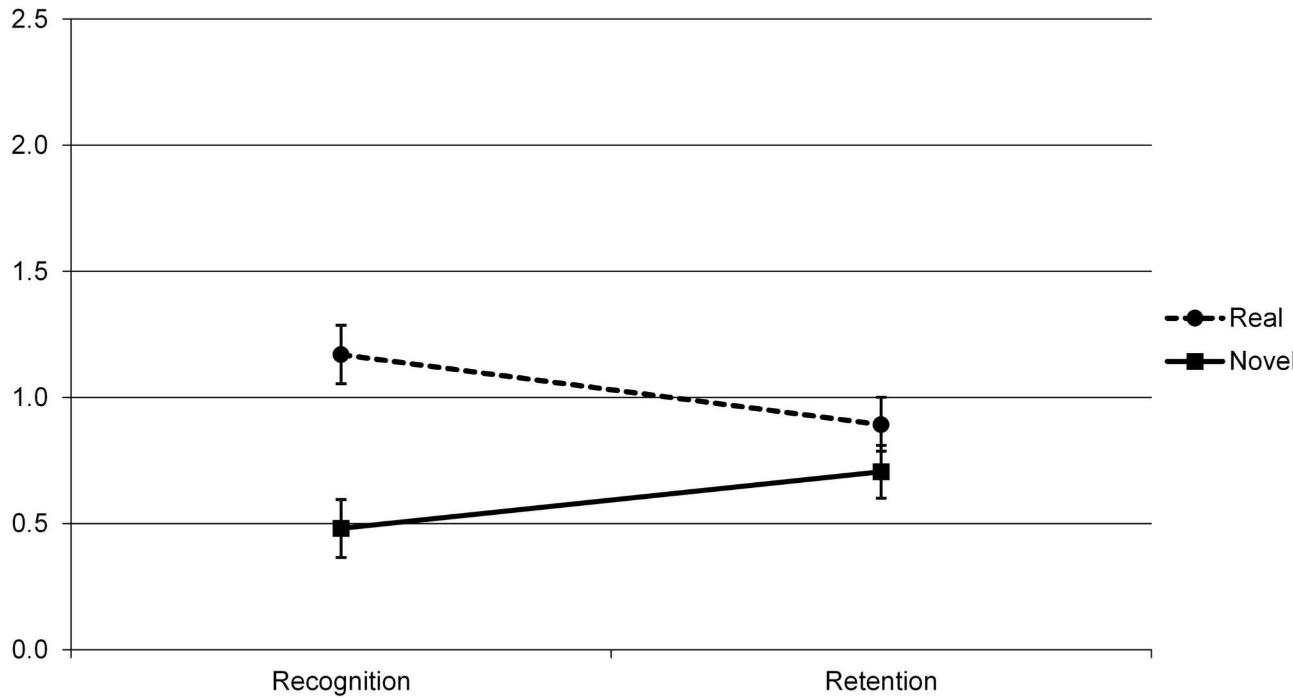


Fig 4. Verbal subtask performance by Delay (10 minutes/Recognition vs. 24 hours/Retention) and Real/Novel, showing mean *d'* scores and standard errors.

doi:10.1371/journal.pone.0169474.g004

$\eta_p^2 = 0.161$), with better performance at Retention than Recognition, and of Group ($F(1,40) = 7.038$, $p = 0.011$, $\eta_p^2 = 0.150$), with the TD group showing better performance overall than the children with SLI. However, these main effects were qualified by an interaction between Group and Modality that approached significance ($F(1,40) = 3.822$, $p = 0.058$, $\eta_p^2 = 0.087$), and significant interactions both between Group and Delay ($F(1,40) = 5.426$, $p = 0.025$, $\eta_p^2 = 0.119$) and Modality and Delay ($F(1,40) = 15.189$, $p < 0.001$, $\eta_p^2 = 0.275$). The three-way interaction was not significant ($F(1,40) = 2.708$, $p = 0.108$, $\eta_p^2 = 0.063$).

Following up on these interactions, we found, first of all, that Modality had a significant effect in both groups, with better performance on the nonverbal than verbal subtask, though this effect was larger in the SLI group (SLI: $F(1,20) = 34.636$, $p < 0.001$, $\eta_p^2 = 0.634$; TD: $F(1,20) = 11.531$, $p = 0.003$, $\eta_p^2 = 0.366$). The Group by Delay interaction was explained by better overall performance (over both the Nonverbal and Verbal conditions) at Retention than Recognition by the SLI group but not the TD group (SLI: $F(1,20) = 11.348$, $p = 0.003$, $\eta_p^2 = 0.362$; TD: $F(1,20) = 0.115$, $p = 0.738$, $\eta_p^2 = 0.006$). Finally, follow up analyses for the Modality by Delay interaction revealed a significant effect, over both groups, in the nonverbal subtask, with better performance at Retention than Recognition, but not for in the verbal subtask (Objects: $F(1,20) = 16.947$, $p < 0.001$, $\eta_p^2 = 292$; Words: $F(1,20) = 0.654$, $p = 0.423$, $\eta_p^2 = 0.016$).

Incidental encoding

This paper investigates potential SLI/TD differences in learning and retention in declarative memory. Hence our analyses focus on examining Recognition and Retention differences between the two groups. However, each subtask also includes an incidental encoding phase, in which participants had to judge whether the stimulus they saw or heard was real or novel. Performance at the encoding phases of both the Nonverbal and Verbal subtasks is shown in

Table 4. Analyses were run with normalized *d'* as the dependent measure. Both the SLI and TD groups showed above-chance performance at both subtasks (Table 4).

To examine potential group differences in incidental encoding we performed separate one-way ANOVAs on the Nonverbal and Verbal subtasks. These revealed group differences in both subtasks, with the TD group showing significantly better performance than the SLI group in categorization accuracy (i.e., Real vs. Novel) for both Nonverbal and Verbal stimuli (Nonverbal: $F(1, 40) = 4.789$, $p = .035$, $\eta_p^2 = .107$; Verbal: $F(1, 40) = 28.273$, $p < .001$, $\eta_p^2 = .414$); see Table 4.

These group differences in incidental encoding could potentially affect the group differences reported above for Recognition and Retention. Note that because the encoding was incidental, there is no clear direct relation between success at this phase (distinguishing Real and Novel items) and actually learning the material, which is tested later in Recognition and Retention. Nevertheless, we ran ANCOVAs parallel to the ANOVAs presented above for both the Nonverbal and Verbal subtasks, with encoding (normalized *d'*) included as a covariate.

Importantly, the same pattern of critical results was obtained as for the ANOVAs presented above. The ANCOVA for the Nonverbal task crucially yielded an interaction between Group and Delay ($F(1, 39) = 7.068$, $p = .011$, $\eta_p^2 = .153$; all other effects: $ps > .3$). Likewise, the ANCOVA for the Verbal subtask yielded a similar pattern of results as the ANOVA presented above for this subtask, namely only a borderline significant interaction between Real/Novel and Delay ($F(1, 39) = 3.833$, $p = .057$, $\eta_p^2 = .089$), with no other significant effects ($ps > .1$).

Discussion

The purpose of this study was to examine both learning and retention in declarative memory in SLI, of both nonverbal and verbal information. In order to obtain a clear picture of the status of declarative memory, the study attempted to minimize the influence of functions that can affect measures of learning and retention in this system and are often impaired in SLI, namely free recall and short-term memory. We aimed to achieve this goal by probing declarative memory with a recognition memory task, following incidental encoding. To test for learning we examined recognition memory 10 minutes after encoding, separately for nonverbal items (real and novel objects) and verbal items (real and novel words). To test for retention of this information in the same individuals, we then examined recognition memory of the same items 24 hours later, allowing us to investigate potential differences in overnight retention between

Table 4. Encoding accuracy in the Nonverbal and Verbal subtasks.

	SLI	TD
Nonverbal		
d'	1.99 (1.13) ***	2.65 (0.80) ***
Hit Rate	0.87 (0.14)	0.91 (0.08)
False Alarm Rate	0.27 (0.24)	0.16 (0.13)
Verbal		
d'	2.43 (0.74) ***	3.49 (0.54) ***
Hit Rate	0.87 (0.06)	0.93 (0.05)
False Alarm Rate	0.14 (0.13)	0.03 (0.04)

Note. Accuracy in the Encoding phase, showing means (and standard deviations) of *d'*, as well as of hit rates and false alarm rates. SLI: specific language impairment; TD: typically developing. Asterisks indicate performance greater than chance (mean *d'*s significantly greater than zero, one-sample t-tests, df = 20): ***: p < .001.

doi:10.1371/journal.pone.0169474.t004

the SLI and TD groups. To our knowledge, this is the first study to test incidental learning in declarative memory in SLI, and the second (after McGregor et al., 2013) to test for overnight retention in this system in the disorder.

Analyses revealed the following pattern. In the nonverbal domain we found a Group (SLI vs. TD) by Delay (10 minutes vs. 24 hours) interaction. Analyses revealed that the children with SLI improved significantly in their recognition memory between testing for learning (at the short delay, that is, after 10 minutes) and testing for retention one day later, whereas the TD children showed no change in performance during this period. This pattern held for both real and novel objects. Interestingly, the two groups did not differ significantly in their recognition memory for either real or novel objects at either delay period, except for real objects after the short delay, when the children with SLI performed worse than the TD children.

In the verbal domain, a main effect of group showed that the TD children performed better overall at recognition memory than the children with SLI, that is, over both real and novel words, over both delay periods. Additionally, analyses revealed an improvement of recognition memory for novel words between the short delay and one day later, over both groups. In contrast, recognition memory for real words decreased during this time period over the two groups. Unlike in the nonverbal domain, no interaction between Group and Delay was found.

An analysis directly comparing performance between the nonverbal and verbal subtasks revealed a Group by Delay interaction over both subtasks, due to better overall performance at the long than short delays only in the SLI group, with no three-way Group by Modality by Delay interaction. Additionally, this analysis revealed overall better performance (over both delay periods) at the nonverbal than verbal items in both groups, with this difference being larger in the SLI group.

The results suggest the following patterns regarding declarative memory in SLI, at least when tested with a recognition memory paradigm with incidental encoding, with recognition tested minutes after encoding and then again one day later. In the nonverbal domain, children with SLI appear to have recognition memory deficits only for real objects, and only at a short delay of minutes. Importantly, they do not show recognition memory deficits for real objects one day after learning the items, and do not show impairments for novel objects at either delay period. Moreover, only children with SLI *improve* at remembering items between initial learning and testing one day later, and in fact do so for both real and novel items. In contrast, TD children show no changes in performance during this period. In the verbal domain, children with SLI appear to have recognition memory deficits for both real and novel words, at both short and long delays. However, the *changes* in performance between the short and long delays do not differ between the groups for verbal items.

A key question is how the observed patterns may best be interpreted. First of all, it does not seem likely that these findings can be accounted for by differences between the TD and SLI groups at the incidental encoding task. Success at distinguishing real and novel items in this task does not have any clear relation with actually encoding the material. Moreover, ANCOVAs with performance from the encoding task covaried out yielded similar patterns to those from the ANOVAs without this factor included.

Second, it might be argued that the improvements between the two delay periods observed in the nonverbal task for the SLI but not the TD group could be due to ceiling effects for the latter. On this view, the lack of an increase between the two delay periods for the TD group might simply be explained by the fact that their performance was already very good at the nonverbal task after the short delay, and hence they had less room for improvement. Indeed, the highest performance at the short delay was observed for the TD group, for real objects, with an accuracy score (over hits and correct rejections) of 82%. However, the TD group's accuracy for novel objects was only 61%, yet they showed no improvements for either novel or real objects.

In contrast, the SLI group showed improvements at both real and novel objects, even though their accuracy at real objects was higher at the short delay (72%) than it was for the TD group for novel ones (i.e., 61%). Moreover, even for real objects, accuracy for the TD group at the short delay (i.e., 82%) was not particularly close to ceiling (i.e., 100%). Indeed, the variance of *d*-prime scores for real objects at the short delay did not differ between the TD and SLI groups, also arguing against ceiling effects for the TD group for this condition (see Table 2; Levene's test for equality of variance: $F(1, 40) = .319$; $p = .575$). Together, the data suggest that ceiling effects are unlikely to explain the pattern of improvements at the nonverbal task between the two delays for the SLI but not the TD group.

Third, it might be suggested that the improvements observed between the short and long delays for the SLI but not the TD group could be explained by worse initial learning by the children with SLI. In particular, since the children with SLI showed worse performance than the TD children at real objects at the short delay, it could be argued that they would be positioned to show more additional learning with an additional exposure (i.e., of the target items during the recognition memory task at the short delay)–assuming a classic non-linear (e.g., log-shaped) learning curve [48,49], since the performance of the children with SLI at the short delay is "further left" on the curve. On this view, such additional learning could result in greater improvements in the SLI than TD group between the short and long delay. Alternatively, it might be argued that the children with SLI in particular understood the instructions better the second time, at the long delay. In either case, however, the children with SLI were not significantly worse than the TD children on *novel* objects at the short delay; yet the same pattern was observed on these items as on the real objects, namely, an improvement between the two delays for the children with SLI but not the TD children. Additionally, the performance at the short delay was lower for *both* participant groups on novel objects than even the SLI group on real objects (see Table 2), yet only the SLI group showed improved performance between the delays, moreover on *both* real *and* novel objects. Overall, this suggests that lower performance at the short delay is unlikely to account for the increases between the delays observed for the SLI group but not the TD group.

We suggest instead that the group differences observed in the changes between testing for initial learning and for retention 24 hours later for both real and novel items in the nonverbal task could be due to group differences in *consolidation*. Consolidation, as we have seen above, refers to the stabilization of memories after their initial acquisition. This process, which depends on the medial temporal lobes as well as neocortical regions [11,50,51], and whose molecular mechanisms are quite well studied [52–54], has been examined not only extensively in non-human animals, but also in humans. In humans, consolidation has been observed for both verbal and nonverbal information, over various time periods, ranging from hours to days to weeks [55–58]. Studies have revealed the importance of sleep in consolidation, showing that sleep can help preserve information, often with better retrieval of the learned information after a period involving sleep than after the same period without sleep [56,59–62]. Some of these studies show that sleep can lead to enhanced retrieval not only as compared to conditions without sleep, but even as compared to initial learning [61,62].

Based on the results from the present study, we suggest that the children with SLI may show consolidation strengths in declarative memory, as compared to TD children, at least for nonverbal items over the course of 24 hours with sleep. These strengths seem to hold for different types of nonverbal items, given that increases between the two delays were found for children with SLI but not TD children for both real and novel nonverbal items. Indeed, these strengths seem to lead to normal recognition performance in children with SLI after 24 hours even for items that showed impaired performance at initial learning (i.e., real objects, at the short delay).

The lower recognition memory performance of the SLI than TD group at real (but not novel) objects at the short delay might be explained by the fact that these items are associated with verbal labels, which could impair their processing. This would not be surprising, given the language difficulties found in children with SLI, including with phonology. It is also consistent with the particular impairment observed in this study for the SLI group at the verbal task, including at encoding. More generally, the findings strengthen the view that declarative memory problems in SLI in the verbal domain, in particular with word learning, may be due primarily to language problems rather than to declarative memory per se [2,6].

Although the SLI group was worse than the TD group at verbal items at both the short and long delays, the *change* in performance between the delays did not differ between the groups. This suggests that although consolidation was not enhanced in the SLI group for verbal items, neither was it impaired; rather, the children with SLI showed evidence for normal consolidation in the verbal domain. However, various questions remain about SLI and TD consolidation of verbal items. First, future studies may elucidate why, over both groups, there was a decrease between the two delays for real words, but an increase for novel words. Second, potential group differences between the two groups in consolidation may also be revealed by further research. Although in the current study there was no interaction between group and delay for the verbal material, exploratory analyses on each group, carried out separately for real and novel words as was done for the nonverbal material, suggested an intriguing pattern. For real words, although both groups showed signs of a decrease in performance between the short and long delay, this reached significance only for the TD group (SLI: $F(1, 20) = 3.035$, $p = .097$, $\eta_p^2 = .132$; TD: $F(1, 20) = 9.787$, $p = .005$, $\eta_p^2 = .329$). Moreover, for novel words, although both groups showed increases between the two delays, this effect reached borderline significance for the SLI group ($F(1, 20) = 4.314$, $p = .051$, $\eta_p^2 = .177$) but not for the TD group ($F(1, 20) = 1.591$, $p = .222$, $\eta_p^2 = .074$). These exploratory analyses suggest that TD but not SLI children might show a decrement in performance for real words between the two delays, while only the children with SLI show an improvement at novel words, hinting at the possibility of SLI consolidation strengths in the verbal domain as well. The Group by Delay interaction yielded by the analyses comparing performance on the nonverbal and verbal subtasks, due to better overall performance, only in the SLI group, at the long than short delays over *both* subtasks, with no three-way Group by Modality by Delay interaction, further supports the possibility of SLI consolidation strengths in the verbal domain. Future studies focusing on this issue, with large sample sizes and other tasks, may be useful.

The factors and mechanisms underlying the apparent SLI strengths at consolidation in declarative memory remain to be elucidated. One obvious possibility, though still speculative, is that they may be related in some way to sleep, since previous evidence suggests that sleep, in particular Slow Wave Sleep, is especially important for consolidation in this system [60,63–65]. Indeed, as discussed above, sleep has been found to lead to *improvements* at remembering items as compared to initial learning, as was observed in the present study. Perhaps children with SLI spend more time in Slow Wave Sleep, or have more efficient sleep-related consolidation processes for declarative memory, as compared to TD children. However, further research is required before sleep-related factors can be identified as a source of the observed patterns. Alternatively or in addition, declarative memory consolidation advantages in SLI might be related to the "seesaw" effect, that is, to the enhancement of declarative memory due to impairments of procedural memory [5,12]. On this view, as suggested by the Procedural Deficit Hypothesis (PDH) and the broader Declarative/Procedural model framework upon which the PDH is built (see below), the procedural memory impairments in SLI that appear to lead to their grammatical (and other) deficits may be associated with improvements of declarative memory, due to the seesaw effect. Such an interaction between memory systems might be expected particularly in developmental disorders, given the continuing interactions between systems during development [66,67]. Although the mechanisms of the seesaw effect remain to be elucidated [5,12], and could indeed be related to sleep, the present findings suggest that at least one manifestation of these effects might be related to consolidation, rather than (just) initial learning. Note that no seesaw effect was reported in children with SLI in Kuppuraj et al. [25]; however, this does not counter the possibility of seesaw effects in consolidation in SLI, since in that study retention was examined only one hour after encoding, at which point strong consolidation effects might not be expected.

Implications and future directions

Although this is the first study to suggest possible SLI strengths in declarative memory, and further research and confirmation is clearly needed, the study has various potential implications, and opens up new avenues of research.

First of all, the results support and suggest refinements of the PDH of SLI, specifically regarding the status of declarative memory. In addition to positing abnormalities of brain structures underlying procedural memory, the PDH hypothesizes that in individuals with SLI declarative memory should be largely spared, particularly for nonverbal information, and may even show advantages compared to TD individuals, due to the seesaw effect [5–7]. The findings from the present study are consistent with these predictions, and refine them by revealing not only the normal attainment of nonverbal knowledge in SLI, but, for the first time, apparent retention strengths in declarative memory, which may be related to consolidation.

Additionally, the results suggest future areas of research for the Declarative/Procedural (DP) model, on which the PDH is based. In particular, the possibility of consolidation strengths in declarative memory in SLI, together with evidence suggesting consolidation impairments in procedural memory in SLI [17], suggest that dissociations between lexical/declarative memory and grammatical/procedural processes may extend to consolidation. Future studies should thus further investigate consolidation in the two memory systems and how these might affect language [68].

Apparently normal (or possibly enhanced) consolidation in the verbal domain in SLI might help explain the relative sparing of lexical knowledge in children with the disorder, compared to aspects of grammar [1,6], since learning of lexical but not grammatical knowledge seems to rely critically on declarative memory [5,12]. It could also at least partly explain the observation that lexical abilities appear to gradually improve as children with SLI get older [69,70], in particular because declarative memory improves during childhood [12]. Note that a dependence of lexical memory on declarative memory does not preclude an additional reliance of lexical memory on, or interactions with, various functions impaired in SLI, such as phonology, syntax, working memory, or recall, which would be expected to lead to some level of lexical deficits, perhaps continuing throughout the lifespan [6]. Additionally, note that normal or enhanced consolidation in declarative memory is consistent with and may help explain why this memory system seems to play a compensatory role for grammar in children with SLI (5–7).

The findings reveal, for the first time, the possibility that children with SLI show cognitive *strengths*, as compared to typically developing children. Strengths in various domains and functions have been observed for a variety of disorders, including in declarative memory in both dyslexia and autism [7,71–73]. However, to our knowledge cognitive strengths have never been reported for SLI, in any domain. The findings presented here suggest that children with SLI also show such strengths, perhaps in consolidation in declarative memory. Further studies examining this issue seem warranted.

Possible strengths in declarative memory in SLI are also consistent with the compensation underdiagnosis hypothesis [7]. On this view, SLI may be underdiagnosed partly as a result of compensation by declarative memory, in particular for grammatical/procedural deficits. If indeed aspects of declarative memory are enhanced in SLI, including possibly in the verbal domain (see above), this would facilitate compensation, potentially increasing underdiagnosis of the disorder.

The findings suggest the need for further investigation of declarative memory consolidation in SLI. This could elucidate how broadly the apparent consolidation strengths in this system may hold in SLI–for example, across age groups in both children and adults, types of items (e.g., other types of nonverbal knowledge, such as faces or scenes), tasks (including those that are less episodic in nature, as well those that involve free or cued recall), and delay periods, and how these may interact with sleep and the time of day during which encoding occurs [33]. (Note that comparisons between the present study and McGregor et al. (2013) are difficult, given the difficulty in interpreting results from that study, and the differences between the studies, including incidental vs. intentional training, testing with recognition vs. recall, and the nature of the participants; see Introduction). Of particular interest, it remains to be seen whether children with SLI might show *better* memory than TD children for nonverbal (and perhaps verbal) information after *longer* periods, during which additional consolidation could take place. The apparent procedural memory consolidation deficits in SLI [17] should also be further examined, including the underlying mechanisms. Overall, such investigations of consolidation in SLI may elucidate not only the nature of SLI, but also of consolidation more generally.

The particular patterns of performance of hits and false alarms in the present study also warrant further investigation. This pattern suggests that the SLI consolidation strengths might be due to a larger reduction of false alarms between Recognition and Retention in the SLI than the TD group, rather than an increase in the number of hits, at least in the case of real objects (see Table 2), and perhaps also for novel words (see Table 3). Although research is sparse on hits versus false alarms in retention and consolidation, there is some evidence in the literature that in recognition memory tasks, sleep reduces false recognition (false alarms), while it does not affect correct recognition (hits) [74], consistent with the pattern observed here. Alternatively, it might be argued that participants with SLI may have better understood the requirements of the recognition memory task the second time (i.e., at Retention), perhaps leading to a more consistent rejection of foils, i.e. to a lower number of false alarms. On this account, however, it is unclear why an SLI reduction in false alarms from Recognition to Retention would be found for real but not novel objects (see Table 2), and perhaps novel but not real words (see Table 3). Future studies, with larger numbers of participants, might clarify these issues. Additionally, it remains to be seen why the impaired SLI performance on some conditions (e.g., real objects and real words at the short delay), as compared to TD children, seems to be primarily due to differences in false alarms rather than in hits (see Tables 2 and 3).

The findings of the present study suggest that retention in declarative memory should also be further examined in other disorders. It should be investigated particularly in disorders that may be related to SLI, as evidenced by comorbidities with SLI and similar patterns of deficits and spared functions, including of declarative memory [7,75]. These may include dyslexia, autism, Tourette syndrome, obsessive-compulsive disorder, and ADHD [7]. Indeed, one study found enhanced declarative memory in dyslexia, although the superior performance was observed across both short and long (one day) delays, with no improvement between them [72]. Future studies on consolidation in such disorders seem warranted.

The results of the present study also have methodological implications. In particular, the study suggests that examining the status of declarative memory may be usefully carried out with tasks that minimize the involvement of other functions that interact with, but are not necessary for, the functioning of this system, such as free recall and working memory. This approach seems particularly important in the many developmental and other disorders where these functions are problematic [7]. More generally, examining the status of a neurocognitive function may be best carried out by tasks that minimize other, non-critical, functions, especially if these may be impaired.

Finally, the study of course has limitations, which could be addressed by future research. For example, future studies could attempt to match the TD and SLI groups on performance of real objects at the short delay, in order to test whether group differences in consolidation would still be found, even for real objects. Additionally, whereas in this study the same target items were presented at both the short and long delays, future studies could include different target items in the different test sessions, thereby avoiding potential problems of group learning differences, as discussed above.

Conclusion

In conclusion, the present study reveals normal and perhaps even enhanced consolidation in declarative memory in SLI. To our knowledge this is the first demonstration of apparent cognitive strengths in children with SLI. The findings, should they be supported by further studies, have a range of basic research and clinical implications for SLI as well as for related disorders, and open up new avenues of research.

Supporting Information

S1 Fig. Individual participant performance on the nonverbal task for the SLI (A) and TD (B) groups. The figures show *d'* performance for each individual, collapsed over both Real and Novel items, at both Recognition (10 minutes after encoding) and Retention (24 hours after encoding). SLI: children with specific language impairment; TD: typically-developing children; *d'*: d-prime scores.
(DOCX)

S1 Dataset. Dataset behind analyses in 'Learning and overnight retention in declarative memory in specific language impairment'.
(XLS)

Acknowledgments

The authors are grateful to the children for their participation and to the schools for accommodating this research. We very much appreciate the help of the speech and language therapists in screening, and thank Kata Fazekas, Enikő Ladányi and Borbála Győri for collecting the data.

Author Contributions

Conceptualization: AL MU.

Formal analysis: AL FK JL MU.

Funding acquisition: AL MU.

Investigation: AL FK.

Methodology: AL FK JL MU.

Project administration: AL.

Supervision: AL MU.

Visualization: AL FK MU.

Writing – original draft: AL FK MU.

Writing – review & editing: AL FK MU.

References

1. Leonard LB. Children with Specific Language Impairment. MIT Press; 2014.
2. Lum JAG, Conti-Ramsden G, Page D, Ullman MT. Working, declarative and procedural memory in specific language impairment. Cortex. 2012; 48: 1138–1154. doi: 10.1016/j.cortex.2011.06.001 PMID: 21774923
3. Marton K, Schwartz RG. Working Memory Capacity and Language Processes in Children With Specific Language Impairment. J Speech Lang Hear Res. 2003; 46: 1138. PMID: 14575348
4. Montgomery JW, Magimairaj BM, Finney MC. Working Memory and Specific Language Impairment: An Update on the Relation and Perspectives on Assessment and Treatment. Am J Speech Lang Pathol. 2010; 19: 78. doi: 10.1044/1058-0360(2009/09-0028) PMID: 19948760
5. Ullman MT. Contributions of memory circuits to language: the declarative/procedural model. Cognition. 2004; 92: 231–270. doi: 10.1016/j.cognition.2003.10.008 PMID: 15037131
6. Ullman MT, Pierpont EI. Specific Language Impairment is not Specific to Language: the Procedural Deficit Hypothesis. Cortex. 2005; 41: 399–433. PMID: 15871604
7. Ullman MT, Pullman MY. A compensatory role for declarative memory in neurodevelopmental disorders. Neurosci Biobehav Rev. 2015; 51: 205–222. doi: 10.1016/j.neubiorev.2015.01.008 PMID: 25597655
8. Lum JAG, Conti-Ramsden G, Morgan AT, Ullman MT. Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. Cortex. 2014; 51: 1–10. doi: 10.1016/j.cortex.2013.10.011 PMID: 24315731
9. Eichenbaum H. The Cognitive Neuroscience of Memory: An Introduction. Oxford University Press; 2011.
10. Henke K. A model for memory systems based on processing modes rather than consciousness. Nat Rev Neurosci. 2010; 11: 523–532. doi: 10.1038/nrn2850 PMID: 20531422
11. Squire LR, Wixted JT. The Cognitive Neuroscience of Human Memory Since H.M. Annu Rev Neurosci. 2011; 34: 259–288. doi: 10.1146/annurev-neuro-061010-113720 PMID: 21456960
12. Ullman MT. The declarative/procedural model: A neurobiological model of language learning, knowledge and use. The Neurobiology of Language. Elsevier; 2016.
13. Chun MM. Contextual cueing of visual attention. Trends Cogn Sci. 2000; 4: 170–178. PMID: 10782102
14. Schendan HE, Searl MM, Melrose RJ, Stern CE. An fMRI Study of the Role of the Medial Temporal Lobe in Implicit and Explicit Sequence Learning. Neuron. 2003; 37: 1013–1025. PMID: 12670429
15. Shimamura AP. Hierarchical relational binding in the medial temporal lobe: The strong get stronger. Hippocampus. 2010; 20: 1206–1216. doi: 10.1002/hipo.20856 PMID: 20824723
16. Gabriel A, Stefaniak N, Maillart C, Schmitz X, Meulemans T. Procedural Visual Learning in Children With Specific Language Impairment. Am J Speech-Lang Pathol Am Speech-Lang-Hear Assoc. 2012; 21.
17. Hedenius M, Persson J, Tremblay A, Adi-Japha E, Veríssimo J, Dye CD, et al. Grammar predicts procedural learning and consolidation deficits in children with Specific Language Impairment. Res Dev Disabil. 2011; 32: 2362–2375. doi: 10.1016/j.ridd.2011.07.026 PMID: 21840165
18. Kemény F, Lukács Á. Impaired procedural learning in language impairment: results from probabilistic categorization. J Clin Exp Neuropsychol. 2009; 32: 249–58. doi: 10.1080/13803390902971131 PMID: 19548167
19. Lukács Á, Kemény F. Domain-General Sequence Learning Deficit in Specific Language Impairment. Neuropsychology. 2014; 28.
20. Tomblin JB, Mainela-Arnold E, Zhang X. Procedural Learning in Adolescents With and Without Specific Language Impairment. Lang Learn Dev. 2007; 3: 269–293.
21. Baird G, Dworzynski K, Slonims V, Simonoff E. Memory impairment in children with language impairment. Dev Med Child Neurol. 2010; 52: 535–540. doi: 10.1111/j.1469-8749.2009.03494.x PMID: 19807770
22. Bishop DVM, Hsu H. The declarative system in children with specific language impairment: a comparison of meaningful and meaningless auditory-visual paired associate learning. BMC Psychol. 2015; 3: 3. doi: 10.1186/s40359-015-0062-7 PMID: 25780564

23. Dewey D, Wall K. Praxis and memory deficits in language-impaired children. Dev Neuropsychol. 1997; 13: 507–512.
24. Riccio CA, Cash DL, Cohen MJ. Learning and memory performance of children with Specific Language Impairment (SLI). Appl Neuropsychol. 2007; 14: 255–261. doi: 10.1080/09084280701719203 PMID: 18067421
25. Kuppuraj S, Rao P, Bishop DV. Declarative capacity does not trade-off with procedural capacity in children with specific language impairment. Autism Dev Lang Impair. 2016; 1.
26. Lum JAG, Conti-Ramsden G. Long-term memory: A review and meta-analysis of studies of declarative and procedural memory in specific language impairment. Top Lang Disord. 2013; 33: 282–297. PMID: 24748707
27. Records NL, Tomblin JB, Buckwalter PR. Auditory verbal learning and memory in young adults with specific language impairment. Clin Neuropsychol. 1995; 9: 187–193.
28. Shear PK, Tallal P, Delis DC. Verbal learning and memory in language impaired children. Neuropsychologia. 1992; 30: 451–458. PMID: 1620325
29. Duinmeijer I, de Jong J, Scheper A. Narrative abilities, memory and attention in children with a specific language impairment. Int J Lang Commun Disord. 2012; 47: 542–555. doi: 10.1111/j.1460-6984.2012.00164.x PMID: 22938065
30. Lum JAG, Gelgic C, Conti-Ramsden G. Procedural and declarative memory in children with and without specific language impairment. Int J Lang Commun Disord. 2010; 45: 96–107. doi: 10.3109/13682820902752285 PMID: 19900077
31. Nichols S, Jones W, Roman MJ, Wulfeck B, Delis DC, Reilly J, et al. Mechanisms of verbal memory impairment in four neurodevelopmental disorders. Brain Lang. 2004; 88: 180–189. doi: 10.1016/S0093-934X(03)00097-X PMID: 14965540
32. Lum JAG, Ullman MT, Conti-Ramsden G. Verbal declarative memory impairments in specific language impairment are related to working memory deficits. Brain Lang. 2015; 142: 76–85. doi: 10.1016/j.bandl.2015.01.008 PMID: 25660053
33. McGregor KK, Licandro U, Arenas R, Eden N, Stiles D, Bean A, et al. Why Words Are Hard for Adults With Developmental Language Impairments. J Speech Lang Hear Res. 2013; 56: 1845. doi: 10.1044/1092-4388(2013/12-0233) PMID: 24023376
34. Ullman MT, Pullman MY, Lovelett JT, Pierpont EI, Turkeltaub PE. The neuroanatomy of specific language impairment. Annu Rev Psychol. To appear;
35. Tager-Flusberg H, Cooper J. Present and Future Possibilities for Defining a Phenotype for Specific Language Impairment. J Speech Lang Hear Res. 1999; 42: 1275. PMID: 10515521
36. Csányi FY. Peabody szókincsteszt [Peabody Vocabulary Test]. 1974th ed. Budapest: Bárczi Gusztáv Gyógypedagógiai Főiskola; 1974.
37. Dunn L. The Peabody Picture Vocabulary Test. 1959th ed. Circle Pines, MN: American Guidance Service; 1959.
38. Bishop DVM. Test For Reception of Grammar. 1983rd ed. Manchester, U.K.: Medical Research Council; 1983.
39. Lukács Á, Rózsa S, Győri M. A TROG pszichometriai jellemzőinek magyar vizsgálata, a normák kialakítása. TROG—Test for Reception of Grammar. 2012th ed. Budapest: OS Hungary Tesztfejlesztő Kft.; 2012. pp. 47–86.
40. Kas B, Lukács Á. Magyar Mondatutánmondási Teszt. in preparation. in preparation.
41. Racsmány M, Lukács Á, Németh D, Pléh C. A verbális munkamemória magyar nyelvű vizsgálóeljárásai. Magy Pszichol Szle. 2005; 60: 479–506.
42. Raven J, Court J, Raven J. Raven's Progressive Matrices and Raven's Colored Matrices. [Internet]. 1987th ed. London: H. K. Lewis; 1987. Available: 1987
43. Psychology Software Tools, Inc. [Internet]. 2012. Available: http://www.pstnet.com.
44. Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol [Hum Learn]. 1980; 6: 174–215.
45. Eals M, Silverman I. The Hunter-Gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. Ethol Sociobiol. 1994; 15: 95–105.
46. Cornelissen K, Laine M, Renvall K, Saarinen T, Martin N, Salmelin R. Learning new names for new objects: Cortical effects as measured by magnetoencephalography☆. Brain Lang. 2004; 89: 617–622. doi: 10.1016/j.bandl.2003.12.007 PMID: 15120553
47. Williams P, Tarr MJ. Structural processing and implicit memory for possible and impossible figures. J Exp Psychol Learn Mem Cogn. 1997; 23: 1344–1361. PMID: 9372604

48. Hull CL. Principles of behavior: an introduction to behavior theory. Oxford: Appleton-Century; 1943.
49. Rescorla RA, Wagner AW. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical Conditioning II: Current Research and Theory. New York: Appleton-Century-Crofts; 1972. pp. 64–99.
50. Nader K, Schafe GE, LeDoux JE. The labile nature of consolidation theory. Nat Rev Neurosci. 2000; 1: 216–219. doi: 10.1038/35044580 PMID: 11257912
51. Wixted JT. The Psychology and Neuroscience of Forgetting. Annu Rev Psychol. 2004; 55: 235–269. doi: 10.1146/annurev.psych.55.090902.141555 PMID: 14744216
52. Alberini CM, LeDoux JE. Memory reconsolidation. Curr Biol. 2013; 23: R746–R750. doi: 10.1016/j.cub.2013.06.046 PMID: 24028957
53. Dudai Y. The neurobiology of consolidations, or, how stable is the engram? Annu Rev Psychol. 2004; 55: 51–86. doi: 10.1146/annurev.psych.55.090902.142050 PMID: 14744210
54. Frankland PW, Bontempi B. The organization of recent and remote memories. Nat Rev Neurosci. 2005; 6: 119–130. doi: 10.1038/nrn1607 PMID: 15685217
55. Brown H, Weighall A, Henderson LM, Gaskell GM. Enhanced recognition and recall of new words in 7- and 12-year-olds following a period of offline consolidation. J Exp Child Psychol. 2012; 112: 56–72. doi: 10.1016/j.jecp.2011.11.010 PMID: 22244988
56. Ellenbogen JM, Hulbert JC, Jiang Y, Stickgold R. The Sleeping Brain's Influence on Verbal Memory: Boosting Resistance to Interference. Rogers N, editor. PLoS ONE. 2009; 4: e4117. doi: 10.1371/journal.pone.0004117 PMID: 19127287
57. Henderson LM, Weighall AR, Brown H, Gaskell M. Consolidation of vocabulary is associated with sleep in children: Sleep-associated consolidation of vocabulary. Dev Sci. 2012; 15: 674–687. doi: 10.1111/j.1467-7687.2012.01172.x PMID: 22925515
58. Sheth BR, Varghese R, Truong T. Sleep Shelters Verbal Memory from Different Kinds of Interference. SLEEP. 2012;
59. Barrett TR, Ekstrand BR. Effect of sleep on memory. 3. Controlling for time-of-day effects. J Exp Psychol. 1972; 96: 321–327. PMID: 4345763
60. Diekelmann S, Born J. The memory function of sleep. Nat Rev Neurosci. 2010;
61. Payne JD, Tucker MA, Ellenbogen JM, Wamsley EJ, Walker MP, Schacter DL, et al. Memory for semantically related and unrelated declarative information: the benefit of sleep, the cost of wake. PloS One. 2012; 7: e33079. doi: 10.1371/journal.pone.0033079 PMID: 22457736
62. Tucker MA, Fishbein W. Enhancement of declarative memory performance following a daytime nap is contingent on strength of initial task acquisition. Sleep. 2008; 31: 197–203. PMID: 18274266
63. Holz J, Piosczyk H, Feige B, Spiegelhalder K, Baglioni C, Riemann D, et al. EEG sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation: EEG sigma and SWA and memory consolidation. J Sleep Res. 2012; 21: 612–619. doi: 10.1111/j.1365-2869.2012.01017.x PMID: 22591117
64. Marshall L, Born J. The contribution of sleep to hippocampus-dependent memory consolidation. Trends Cogn Sci. 2007; 11: 442–450. doi: 10.1016/j.tics.2007.09.001 PMID: 17905642
65. Mednick SC, Cai DJ, Shuman T, Anagnostaras S, Wixted JT. An opportunistic theory of cellular and systems consolidation. Trends Neurosci. 2011; 34: 504–514. doi: 10.1016/j.tins.2011.06.003 PMID: 21742389
66. Karmiloff-Smith A. Development itself is the key to understanding developmental disorders. Trends Cogn Sci. 1998; 2: 389–398. PMID: 21227254
67. Thomas M, Karmiloff-Smith A. Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. Behav Brain Sci. 2002; 25: 727-750-787.
68. Morgan-Short K, Finger I, Grey S, Ullman MT. Second Language Processing Shows Increased Native-Like Neural Responses after Months of No Exposure. Stamatakis EA, editor. PLoS ONE. 2012; 7: e32974. doi: 10.1371/journal.pone.0032974 PMID: 22470434
69. Rice ML. Language growth and genetics of specific language impairment*. Int J Speech Lang Pathol. 2013; 15: 223–233. doi: 10.3109/17549507.2013.783113 PMID: 23614332
70. Rice ML, Hoffman L. Predicting Vocabulary Growth in Children With and Without Specific Language Impairment: A Longitudinal Study From 2;6 to 21 Years of Age. J Speech Lang Hear Res. 2015; 58: 345. doi: 10.1044/2015_JSLHR-L-14-0150 PMID: 25611623
71. Boucher J, Mayes A, Bigham S. Memory in autistic spectrum disorder. Psychol Bull. 2012; 138: 458–496. doi: 10.1037/a0026869 PMID: 22409507

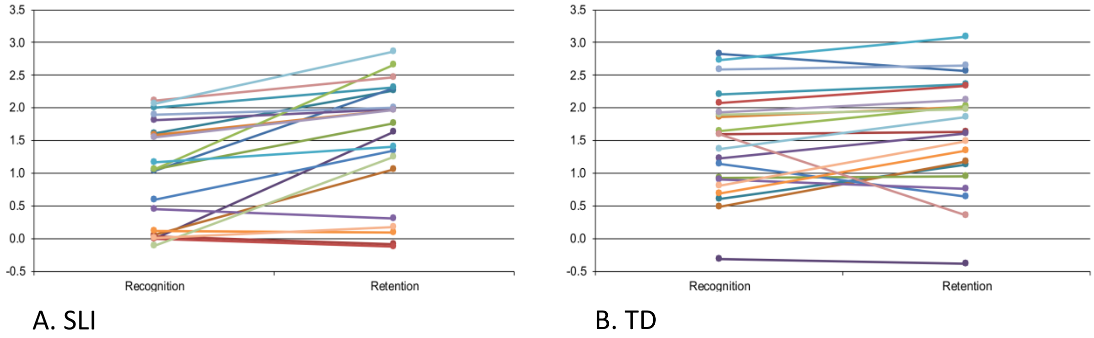
72. Hedenius M, Ullman MT, Alm P, Jennische M, Persson J. Enhanced Recognition Memory after Incidental Encoding in Children with Developmental Dyslexia. Chao L, editor. PLoS ONE. 2013; 8: e63998. doi: 10.1371/journal.pone.0063998 PMID: 23717524

73. Walenski M, Mostofsky SH, Gidley-Larson JC, Ullman MT. Brief Report: Enhanced Picture Naming in Autism. J Autism Dev Disord. 2008; 38: 1395–1399. doi: 10.1007/s10803-007-0513-y PMID: 18163206

74. Fenn KM, Gallo DA, Margoliash D, Roediger HL, Nusbaum HC. Reduced false memory after sleep. Learn Mem Cold Spring Harb N. 2009; 16: 509–513.

75. Krishnan S, Watkins KE, Bishop DVM. Neurobiological Basis of Language Learning Difficulties. Trends Cogn Sci. 2016; 20: 701–714. doi: 10.1016/j.tics.2016.06.012 PMID: 27422443

dc_1365_16



S1 Fig. Individual participant performance on the nonverbal task for the SLI (A) and TD (B) groups. The figures show *d'* performance for each individual, collapsed over both Real and Novel items, at both Recognition (10 minutes after encoding) and Retention (24 hours after encoding). SLI: children with specific language impairment; TD: typically-developing children; *d'*: d-prime scores.

Summary and conclusions

Before proceeding to the general discussion of the results, I present the results by thesis points of the dissertation.

Results by theses of the dissertation

Thesis 1. Alongside the grammatical deficit, there is also evidence of lexical impairments in SLI, arguing against the selective impairment of grammar (1-2, 4-6.)[1]

In all our studies on language abilities of children with SLI, we compared performance of the SLI group to a group of typically developing children matched individually in vocabulary size. If the lexicon was intact, as predicted by grammar specific accounts, vocabulary controls would have been chronological age controls as well. As shown by our results, children in the VC groups were significantly younger than children with SLI (around one year younger in pre-school groups, and 2-3 years younger at school-age), suggesting that vocabulary size of children with SLI matched those of much younger TD children. The greater the age difference, the greater the lexical delay. SLI performance on grammatical tasks was closely associated with vocabulary size in several studies. Lexical deficits are also reflected in the finding that within the area of case marking we observed a special difficulty in the use of suffixes in their lexically specified, nonspatial and nontransparent meanings.

Thesis 2. Agreement deficits in SLI are better explained by processing difficulties than by a selective grammatical impairment targeting agreement. (1, 4)

In an elicited production task disguised as a sentence repetition task with masked inflections, performance of the SLI group lagged significantly behind VC children, but the group difference disappeared when performance on a nonword repetition task was included as a covariate. Detailed analysis showed that the group difference was especially large for the second person plural. Processing factors (inflection frequency) were significant in modelling performance in both the SLI and the VC groups, and the effect was stronger in the SLI group. Although performance levels were different in the two groups, SLI and VC children generally showed the same performance profile across the inflection types. The two groups of children were also similar in their pattern of errors. Inflections produced in place of the correct

[1] Numbers in parentheses show the numbers assigned to the publications (in the 'Publications supporting the thesis points of the dissertation' list) associated with the thesis points.

inflection usually differed from the correct form on a single dimension (e.g., tense or definiteness), though no single dimension was consistently problematic.

There were no significant group differences in the accuracy of grammaticality judgments with either types of errors. Both groups recognized well-formed sentences significantly more accurately than the ill-formed ones, and all children were better in recognizing agreement errors than morphophonological ones. We have found no evidence of a special difficulty with any of the agreement features. Tense errors were more difficult for both groups, probably due to a distance effect. Although there was no difference between performance levels of the SLI and the VC groups, we found differences in the specific effects of performance factors on grammaticality judgments. Performance in the SLI group was best modelled by nonword repetition span, while in the VC group, this factor did not seem to influence performance. At the same time, TROG scores (reflecting comprehension of sentences of different structural complexity) showed a weak association in VC, which was not present in the SLI group.

Weaker performance on the sentence repetition task requiring morpheme restoration suggests a production problem. Results from the two studies show that grammatical competence is by far not the sole determinant of performance in the production and comprehension of agreement. Accuracy for both elicited production and for grammaticality judgments depended on individual capacity measures (especially nonword repetition span measuring verbal short term memory) and processing difficulty of the items (low frequency, complex phonotactics, length), and the influence of such factors was especially strong in language impairment. Grammar-specific accounts that assume constraints on checking or problems specific to agreement do not provide an explanation for the observed pattern of findings. The findings indicate that models assuming processing limitations on the part of children with SLI like the Morphological Richness Account are more compatible with the pattern of verb inflection use seen in Hungarian-speaking children with SLI.

Thesis 3. Difficulties with multiple suffixation and with morphophonologically irregular forms suggest lexical and processing problems instead of a grammar-specific deficit in SLI (2)

The younger group of children with SLI was less accurate than the younger VC children when two suffixes (marking plural and accusative case) were required, at least when noun stem classes were regular. Older children with SLI had more difficulties with accusative forms of irregular items than TD peers, but with scores not counting morphophonological overgeneralizations among errors group differences disappeared. All groups showed significant overgeneralization of irregular stem forms with correct morphosyntactic selection of suffixes. The general patterns across suffixes and regular and irregular items, just like errors patterns, were similar in all groups. However, there were strong word frequency effects in the SLI, but not in the VC groups, suggesting group differences in learning strategies: the SLI groups seem to rely more (although not exclusively) on rote learning of individual inflected forms, and are less effective in pattern extraction and analogical extension. These findings also argue against the selective impairment of grammar because 1) difficulties were observed with lexically more complex items and with forms where children had to rely on multiple operations 2) overgeneralization errors were markedly present in the SLI group as well. These results suggest that morphological knowledge in SLI is strongly intertwined with limitations in lexical knowledge.

Thesis 4. Difficulties in aspect marking in production but not in comprehension in past tense forms suggest a processing problem instead of a selective impairment of aspect marking (5)

In our study on the comprehension and production of both imperfective and perfective verb forms in past tense contexts, children with SLI did not differ in their comprehension performance from either the AC or the VC group, but children with SLI were less accurate than both comparison groups in producing both imperfective (*A szamár itta a tejet*) and perfective (*A szarvas megitta a kólát*) forms. Based on these results, it appears that children with SLI have difficulties selecting the appropriate aspectual marking in past tense contexts. Comparable recognition of the correct form in the comprehension task in the two groups together with the nature of production errors suggest a processing deficit instead of a specific grammatical deficit in the marking of aspect. Children made mistakes in both directions: they used an imperfect form in a perfect context and perfect forms in imperfect contexts; i.e. they did not use a ’default’ form. Errors were mostly “near-miss” errors of aspect only and errors of tense only; errors of both aspect and tense were rare. The act of retrieving the appropriate form for production was relatively more difficult for the children with SLI than for the TD children. Processing problems are also supported by the fact that the presence of adverbs facilitated children's accuracy.

Thesis 5. Problems with case marking in SLI suggest lexical and processing deficits instead of a selective case marking impairment within grammar (6)

Analysing narrative samples showed that the number of different nouns, the number of case marked nouns and of different case markers was significantly lower for children with SLI than for typically developing children. The number of case marking errors was very low in both groups, showing less diversity but correct spontaneous use of morphology in SLI.

In an elicited production task with masked inflections of case markers in their spatial and nonspatial meanings, overall performance of the children with SLI was significantly below that of TD children in both age groups, and SLI children found suffixes with nonspatial meaning especially difficult (although the overall pattern of results was very similar: both SLI and TD children scored higher with spatial than with nonspatial meanings). These findings of lexical (frequency) effects and problems with semantically opaque marking are suggestive of processing problems and argue against a selective difficulty with grammar and a specific impairment in case marking in Hungarian language impairment.

Thesis 6. Cognitive impairments in SLI do not selectively target language; deficits also occur in skill learning outside the language domain, most prominently for sequentially organized stimuli (3, 8.)

In our study testing 3 different forms of skill learning in SLI, a significantly smaller proportion of children showed any evidence of learning in the SLI than in the TD group for the two sequential skill learning tasks (13/29 learners on the SRT, 7/29 learners on the AGL task in the SLI group; 101/159 learners on the SRT, 83/159 learners on the AGL task in the TD group). The proportion of learners on the nonsequential PCL task was the same in the SLI and TD groups. The amount of learning for those children with SLI who did show evidence of learning was overall comparable to the level of learning in TD children on all tasks (with great individual variation). In one of our earlier studies, however, children with SLI showed no evidence of learning and performed at chance on the PCL task as well.

Taken together, these findings suggests that implicit sequence learning tends to be impaired in SLI, but not for all children. We have not found any significant associations between skill learning indices and measures of vocabulary and grammar. Since these language measures were results from our screening tests, further research is needed to clarify the relationship of deficits in implicit learning and more sophisticated and relevant language measures. These results point to deficits in fundamental and domain general learning skills outside language and in some cases, even beyond sequential learning, arguing against the language specificity of the impairment within SLI.

Thesis 7. In concert with the Procedural Deficit hypothesis, procedural learning is vulnerable in SLI, while processes of declarative learning and retention are relatively intact (3, 8, 12).

Results on deficits in procedural learning are summarized under Thesis 6. Declarative memory was tested by examining recognition memory after incidental encoding with both verbal and non-verbal stimuli 10 minutes after encoding, and also 1 day later. Results suggest that in the visual domain, declarative memory is a strength in SLI: on nonverbal items, only the children with SLI improved overnight, with no resulting group differences in performance at the long delay. In the verbal domain, the children with SLI consistently showed worse performance than the typically-developing children, but the two groups showed similar overnight changes. The results did not appear to be explained by group differences in encoding or by ceiling effects among the TD children. The findings strengthen the claim that declarative memory is relatively intact, and can play an important compensatory role in SLI.

Thesis 8. SLI difficulties in executive functions are mainly present on verbal versions of EF tasks, and are eliminated by controlling for verbal short term memory span. (10).

Children with SLI (n=31) were tested on verbal and nonverbal versions of simple and complex span, fluency, N-back and Stroop tasks. Their performance was compared to TD children matched on age and nonverbal IQ. The SLI group showed difficulties in verbal versions of simple and complex span (Digit span and Listening span task) and fluency but not in inhibition (Stroop tasks) relative to TD age-matched children. Performance of the two groups was comparable on nonverbal tasks (simple spatial span measured by Corsi blocks, a nonverbal oddball task and nonverbal stroop and n-back tasks). Including simple verbal span (Digit span) as a covariate eliminated group differences on verbal tasks, which shows that the observed deficits in EF are secondary to the well-documented reduced capacity of verbal short-term memory span in SLI.

Thesis 9. Lexical inhibition is effective in SLI (11).

Children in both the SLI and TD groups took longer to name pictures in high conflict conditions (in semantically homogeneous as opposed to mixed blocks, and for pictures with low as opposed to high naming agreement) than in low conflict conditions. 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.

Thesis 10. Age-related changes in different forms of skill learning with potential roles in language acquisition argue against the existence of a critical period for these learning mechanisms (9).

Age-related changes in three different forms of skill learning between 7 and 80 years using 3 paradigms argue against the existence of a critical period: The three tasks were the following: 1) the Serial Reaction Time Task (learning of motor sequences), 2) Artificial Grammar Learning (the extraction of regularities from auditory sequences) 3) and Probabilistic Category Learning in the Weather prediction task (a non-sequential categorization task). Age-related changes on all three tasks show an inverted U-shaped curve (with different time windows): learning gets more effective during childhood and adolescence, peaking in adulthood, and becoming less efficient with old age. This pattern of results is in concert with earlier findings and models of age-related improvement and then decline in skill learning. It is also compatible with age-invariance models if age-related changes are in fact due to task differences in working memory and explicit load; this question is open for further research. If such domain-general skill learning abilities support language acquisition, then these results, supporting recent findings from second language learning challenge the hypothesis of critical period of language learning, at least concerning the learning mechanisms themselves.

Thesis 11. In aphasia, the acquired language impairment is not specific to language: it is often accompanied by the impairment of nonverbal executive functions (7).

Our results show evidence of impairments in updating working memory representations and inhibition of prepotent responses in aphasia. We found deficits in EF in both individuals with transcortical motor aphasia (TMA), and in conduction aphasia. Individuals with TMA showed impaired inhibition as indexed by the Stop-signal and the nonverbal Stroop tasks, as well as a deficit of updating of working memory representations as indexed by the auditory *n*-back task. Participants with conduction aphasia had difficulties in only one of the tasks measuring inhibition, but no clear evidence for impairment of updating of working memory representations was found. Although the results show different patterns of EF deficits in the two groups with aphasia, the findings clearly demonstrate that EF deficits are not specific to participants with TMA. These results show that the impairment of language is often accompanied by deficits in executive functions in aphasia, which can have detrimental effects on language itself.

General discussion

The aim of the dissertation was to examine the validity of the argument of selective language impairment in the debate about the domain specificity of language processes: the question of the specificity of linguistic impairment was examined in a developmental and an acquired impairment of language by assessing linguistic abilities and non-linguistic cognitive functions. This line of research was complemented by a typical developmental study in which we addressed the hypothesis of a dedicated period of learning that does not necessarily follow from specificity but is often associated with it. The studies presented in the dissertation focused on issues of the specificity of language from different aspects of language and cognition, and support previous claims and results casting doubt on the existence of specialized mechanisms in acquisition and processing.

We examined the problem of language specificity from several aspects. In a detailed examination of language abilities in SLI, we have not found evidence of a selective impairment in any area of grammar (agreement, case marking, aspect marking, regular inflection), moreover, our findings also argue against a selective impairment of grammar as a whole domain. In both school-age and preschool children with SLI we found that performance on measures of both grammar and the lexicon lags behind age-based expectations; verbal working memory proved to be one of the best markers of SLI, and some of our results point to an atypically strong connection in SLI between grammatical performance and verbal working memory. These results suggest processing difficulties that may affect both lexical and grammatical functions, and which may also involve domains outside language. Each of our studies suggested that grammatical competence is far from the only factor that can affect performance on language tasks. Agreement and case marking errors were only observed in elicited production; in spontaneous speech children with SLI avoid the use of complex structures and make very few mistakes. Errors are more prominent in production than in comprehension. We found no overall or selective deficit in any of the grammatical structures targeted by our studies: in several tasks, structures that were difficult in terms of processing (rare, long, non-transparent in their meaning, encoding multiple grammatical functions) were the ones that caused problems in SLI (and in typical development as well, to a lesser extent).

Performance patterns were similar in the SLI and TD groups, which also argues against selective impairments, and supports processing accounts. On grammatical tasks, accuracy measures were dependent on individual processing capacity indices of participants and processing difficulty of items; these associations were especially strong in SLI. Beside similar performance levels, we found evidence that the two groups rely on different learning strategies: the SLI group seemed to rely more (though not exclusively) on the memorization of individual word forms and applies pattern extraction and analogical extension less effectively, as suggested by strong word frequency effects observed in this group. Our results also highlight the observation that the same level and pattern of performance can be supported by different underlying mechanisms in typical and atypical development.

In our studies of cognitive functions outside the linguistic domain, no general impairment was found in executive functions in SLI: the deficit was mostly evident in verbal tasks, where it was a consequence of a more fundamental reduction in verbal short-term memory capacity. Resolution of lexical conflicts also showed a pattern similar to typical development. At the same time, in the acquired disorder of language abilities, aphasia, linguistic impairments were accompanied by deficits in executive functions. These observations, together with previous findings from the literature suggest that language impairment is not necessarily associated with the deficit of executive functions, but deficits in the two groups often co-occur, perhaps depending on the severity of the deficit.

Overall results from our studies on skill learning and declarative memory supported the procedural deficit hypothesis of SLI (PDH, Ullman and Pierpont, 2005): in the skill learning tasks, the SLI group lagged behind typically developing children, primarily in tasks involving the acquisition of sequentially organized information. Age-related changes in various forms of skill learning across a wide age range showed that learning abilities playing an important role in language acquisition do not follow a pattern predicted by a critical period for learning mechanisms and language acquisition. In declarative memory functions, however, we did not find a nonverbal deficit in SLI, and deficits in the verbal domain can probably be explained by the reduced capacity of verbal short-term memory here as well as in EFs.

Taken together, the studies presented in the dissertation argue against the specificity of mechanisms of grammar: language performance is strongly determined by factors that are traditionally regarded as extra-grammatical, such as the processing difficulty of language stimuli and processing capacity of the participants. We have also shown that as an example of selective impairment, neither SLI nor aphasia presents a good argument in support of the specificity of the language: in SLI, we found a deficit in general skill learning abilities outside the language domain, and aphasia was associated with an impairment of executive functions. These results, together with numerous observations in the literature, call into question the specificity of linguistic processes supporting the acquisition, processing and production of language.

Our studies focused on examining specific areas of language and cognition, and because of their focus, they have a limited scope. For this reason, they are unable to address many critical issues related to specificity. For example, we did not critically evaluate the problem of domain-specificity in the neuroanatomical sense (although we mentioned it briefly in the introduction), and did not test whether there are brain regions or neural networks that subserve linguistic functions selectively. The studies do not provide an account of how domain general processes can explain linguistic functions, and they do not have an answer to one of the questions raised in the introduction: are they causal in language problems, or are they associative impairments that tend to co-occur with language impairment because of shared networks or anatomical proximity with language processes. These questions are going to be answered by further research in this direction.

It has been argued that double dissociations present strong arguments for domain-specificity (and, confounded with it, modularity and innateness). Although the studies presented in the dissertation did not examine cases of double dissociations for language, findings from our current and previous research support the above arguments against specificity. Some have proposed that language and cognitive abilities on one hand, and within-language subcomponents of grammar and lexicon on the other show double dissociations in SLI and Williams Syndrome (WS). Based on our previous work on Williams syndrome (e.g. Lukács, 2005; Lukács et al., 2004; Pléh et al., 2003) and the current results on SLI presented in the dissertation, language and cognitive performance patterns in WS and SLI argue against strongly contrasting profiles in these two developmental disorders. Besides lack of evidence for a double dissociation between the lexicon and grammar evidenced by individuals with WS showing impairments in grammar, with the same performance patterns as observed in SLI and TD (e.g. on regular-irregular morphology task; Lukács, 2005; Lukács et al., 2004; Pléh et al., 2003), profiles on more specific language measures are overlapping (although not identical). Differences are present in processing-related measures of verbal short-term memory span in the two groups. The double dissociation is not observed for language versus cognition either: as we have seen, non-linguistic cognitive functions are not intact in SLI, and language is not intact in WS either.

The studies also have general methodological limitations associated with the field of research. SLI research is complicated by participants of different age, patterns of symptoms, severity and aetiology. The etiology of SLI is unknown: genetic and environmental factors both seem to contribute, potentially as multiple risk factors (e.g. Bishop, 2001, 2006; Fattal et al., 2011; Leonard, 1998/2014). Although everybody agrees that SLI includes very heterogeneous groups of children, attempts at creating subcategories have not been successful (see Introduction). The comparison of different theoretical explanations is often made difficult by the contradictory empirical findings frequently coming from methodologically diverse experiments and data analyses.

A further possible limitation comes from aiming to test predictions motivated by theories based on English, and the impossibility of testing all aspects of language and cognition in one project: although our research covered all the main proposals for language specific deficits (see also Ladányi et al., in press, not included in the dissertation), we have not tested all areas of grammar and language, and we cannot exclude that further studies might reveal special difficulties characteristic of Hungarian. Nevertheless, it seems clear from the SLI literature that there is no area of grammar that would be a universal, language-independent marker of specific language impairment. The observation that there are significant cross-linguistic differences and language (or at least language type)-specific symptoms of SLI speaks against the existence of very specific *and* innate mechanisms and representations. While they argue against the specificity of such grammatical representations, cross-linguistic differences in impairments are compatible with language-specific processing mechanisms. Cross-linguistic differences in symptoms can still reflect deficits in similar or even the same cognitive resources (either language-specific or domain general) supporting language. A deficit in these resources (such as problems in working memory both in the storage and in the integration of information across different domains) would result in different symptoms depending on the structure of the language and the areas of grammar where most resources are allocated.

In addition to their theoretical significance, our research on specific language impairment also has practical relevance: it can also be an important step towards systematically profiling language disorders in Hungary. Although we do not have good estimates for Hungary, other reports for Europe show that approximately 7% of children going to school have significant difficulties with speech and/or language (Bercow, 2008, Bishop 2000; Leonard, 2000/2014; Tomblin et al., 1996). By testing a variety of domains and developing tests, this research helps speech and language therapists working with language disorders in comprehensive assessment, diagnosis and effective therapy. Our results on the effective functioning of declarative memory indicating potential mechanisms of compensation can be a basis for efficient training methods. There is a clear need among speech therapists to work with standardized and psychologically grounded methods in assessing language impairments. By describing the linguistic and non-linguistic abilities of SLI, this research hopefully not only supports speech therapists working with language disorder in diagnosis and training, but also contributes to the in-depth study of typical language acquisition in Hungarian.

References

Adams, A.-M. and Gathercole, Susan. E. (2000), Limitations In Working Memory: Implications For Language Development. *International Journal of Language & Communication Disorders, 35*: 95–116.

Alexander, M. P. (1997). Aphasia: clinical and anatomic aspects. *Behavioral neurology and neuropsychology*, 133-150.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71-82.

Baddeley, A. D., Gathercole, S. D. és Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*. 158-173.

Baddeley, A.D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49,* 5-28.

Baird, G., Dworzynski, K., Slonims, V., & Simonoff, E. (2010). Memory impairment in children with language impairment. *Developmental Medicine & Child Neurology*, *52*(6), 535-540.

Bánréti, Z. (2014). Az afázia. In: Pléh, Cs. és Lukács, Á. (2014). (szerk). *Pszicholingvisztika 1-2. Magyar pszicholingvisztikai kézikönyv*. Budapest: Akadémiai. 1167-1241.

Barde, L. H. F. & Thompson-Schill, S. L. (2002). Models of functional organization of lateral prefrontal cortex in verbal working memory: Evidence in favor of the process model. *Journal of Cognitive Neuroscience*, *14*, 1054-1063.

Belton, E., Salmond, C., Watkins, K., Vargha-Khadem, F. & Gadian, D. (2002). Bilateral grey matter abnormalities in a family with a mutation in FOPX2. *Neuroimage, 16*, 10144

Bercow, J. (2008). *The Bercow Report: A review of services for children and young people (0-19) with speech, language and communication needs*. Nottingham: DCSF.

Bialystok, E. & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (ed.), *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T, CATALISE consortium (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children. *PLoS ONE 11*(7): e0158753.

Bishop, D. (2000): How does the brain learn langage? Insights form the study of children with and without language impairement. *Developmental Medicine and Child Neurology, 42,* 133-142.

Bishop, D. (2001): Genetic and environmental risks for specific language impairment in children. *Philos. Transc. Royal Soc., B, 356,* 369-380.

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders*, *49*, 381-415.

Bishop, D. V. M., & Hsu, H. (2015). The declarative system in children with specific language impairment: a comparison of meaningful and meaningless auditory-visual paired associate learning. *BMC Psychology*, *3*(1), 3.

Bishop, D.V., & Adams, C. (1989). Conversational characteristics of children with semantic-pragmatic disorder. II: What features lead to a judgement of inappropriacy? *British Journal of Disorders of Communication*, *24*, 241-263.

Bishop, D.V.M. & Norbury, C.F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology, *Autism, 9*: 7-43.

Bishop, D.V.M. (1997). *Uncommon Understanding: Development and Disorders of Language Comprehension in Children.* Hove, UK: Psychology Press

Bishop, D.V.M. (2006). What causes specific language impairment in children? *Current DirectionsinPsychologicalScience,15,* 217–221.

Bishop, D.V.M. (2012). *TROG - Test for Reception of Grammar. Handbook*. Budapest: OS Hungary Tesztfejlesztő Kft. 47-86.

Blank, I., Kanwisher, N. & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology*, 112(5), 1105-1118.

Bonini MV, Radanovic M. (2015). Cognitive deficits in post-stroke aphasia. *Arquivos de Neuro-Psiquiatria*, 73(10): 840-47.

Bortolini, Umberta, Caselli, Cristina M., Deevy, Patricia and Laurence B. Leonard (2002). Specific language impairment in Italian: the first steps in the search for a clinical marker. International *Journal of Language and Communication Disorders*, 37/2, 77-93.

Burgess, P.W. (2000). Strategy application disorder: the role of the frontal lobes in human multitasking. *Psychological Research, 63,* 279-88.

Caplan, D. (1987). *Neurolinguistics and linguistic aphasiology. An introduction*. Cambridge studies in Speech Science and Communication. Cambridge: Cambridge University Press

Caplan, D. & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(1), 77-126.

Chomsky, N. (1968/1995). *Mondattani szerkezetek/Nyelv és elme*. Budapest: Osiris.

Chomsky, N. (2010). Some simple evo devo theses: how true might they be for language? in Larson, R. K., Déprez, V. & Yamakido, H. (eds.) *The Evolution of Human Language*, Cambridge: Cambridge University Press. 45-62.

Christiansen, M. H., and Chater, N. (2008). Language as shaped by the brain. *Behav. Brain Sci.* 31, 489-558.

Christiansen, M., Conway, C., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes*, *27*, 231-256.

Christiansen, M.H. & Chater, N. (2015). The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology, 6,* 1182.

Clahsen, H. & Hansen, D. (1997). The grammatical agreement deficit in specific language impairment: Evidence from therapy experiments. In Gopnik, M. (ed.) *The Inheritance and Innateness of Grammar*, chapter 7. Oxford University Press, New York.

Clahsen, H. (1991). *Child language and developmental dysphasia.* Amsterdam: John Benjamins.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences, 22*, 991-1060.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences, 2*, 406-416.

Conti-Ramsden, G. & Botting, N. (1999). Classification of Children With Specific Language Impairment: Longitudinal Considerations. *Journal of Speech and Hearing Research, 42*(5), 1195-1204.

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The Extent to Which Psychometric Tests Differentiate Subgroups of Children With SLI. *Journal of Speech and Hearing Research, 40*(4), 765-777.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 24-39.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities - Pitting abstract against stimulus-specific representations. *Psychological Science*, *17*, 905-912.

Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Ann. N.Y. Acad. Sci.* 1145, 113-131.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*, 356-371

Conway, C. M., Karpicke, J. & Pisoni, D. B. (2007). Contribution of implicit sequence learning to spoken language processing: Some preliminary findings with normal-hearing adults. *Journal of Deaf Studies and Deaf Education*, *12*, 317-334.

Conway, C.M., & Christiansen, M.H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences*, *5*, 539-546.

Conway, C.M., Pisoni, D.B., & Kronenberger, W.G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science*, *18*:275-279.

Cornish, H., Dale, R., Kirby, S. & Christiansen, M.H. (2017). Sequence memory constraints give rise to language-like structure through iterated learning. *PLoS ONE* 265-280.

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

D'Esposito M, Postle BR (1999). The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia, 37*:89-101

D'Esposito, M., Detre, J.A., Alsop, D.C., Shin, R.K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive of working memory. *Nature, 378,* 279-281.

Daal, J. V., Verhoeven, L., & Balkom, H. V. (2004). Subtypes of Severe Speech and Language ImpairmentsPsychometric Evidence From 4-Year-Old Children in the Netherlands. *Journal of Speech and Hearing Research, 47*(6), 1411-1423.

Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19*(4), 450–466.

Daneman, M. & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analyses. *Psychonomic Bulletin and Review,* 3, 422-433.

Dankovics, N., & Pléh, C. (2001). Hangrestaurációs jelenségek és alaktani feldolgozás a magyarban: Azt halljuk-e , amit várunk? In C. Pléh & , Á.Lukács (Eds.), *A magyar morfológia pszicholingvisztikája*. Budapest: Osiris. 55-83.

Dewey, D., & Wall, K. (1997). Praxis and memory deficits in language-impaired children. *Developmental Neuropsychology*, *13*(4), 507-512.

Dollaghan, C. (2007). *The Handbook for Evidence-Based Practice in Communication Disorders*. Brookes Publishing.

Dominey, P.F., Hoen, M., Blanc, J.-M., & Lelekov-Boissard, T. (2003). Neurological basis of language and sequential cognition: Evidence from simulation, aphasia, and ERP studies. *Brain and Language, 86*, 207-225.

Dromi, Esther, Leonard, Laurence B., Adam, Galit and Sara Zadunaisky-Ehrlich (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research, 42*, 1414-1431.

Dunn, M. & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test - Revised*. Circle Pines, MN: AGS.

Eisenbeiss, S., Bartke, S., & Clahsen, H. (2005). Structural and lexical case in child German: Evidence from language-impaired and typically-developing children. *Language Acquisition*, *13*(*1)*, 3-32.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996). Rethinking innateness: A connectionist perspective on development. MIT Press.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, Mass.: MIT Press.

Engle, R.W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B.H. Ross (Ed.), *The Psychology of Learning and Motivation (Vol. 44)*. New York: Academic Press. 145-159

Fattal, I., Friedmann, N., & Fattal-Valevski, A. (2011). The crucial role of thiamine in the development of syntax and lexical retrieval: A study of infantile thiamine deficiency. *Brain: A Journal of Neurology, 134,* 1720–1739.

Fedorenko, E. & Thompson-Schill, S.L. (2014). Re-working the language network. *Trends in Cognitive Sciences,* 18(3), 120-126.

Fedorenko, E., M. K. Behr, and N. Kanwisher (2011). Functional Specifi city for High-Level Linguistic Processing in the Human Brain. Proceedings of the National Academy of Science, 108/3 9

Flege, J E, G H. Yeni-Komshian és S Liu (1999), Age Contstraints on Second Language Acquisition. *Journal of Memory and Language* 41, 78-104.

Fodor, J. A. (1983). The Modularity of Mind: An Essay on Faculty Psychology. Cambridge, MA: MIT Press.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78-84.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78-84.

Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends in Cognitive Sciences, 13*(4), 175-181.

Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., & Hewitt, J.K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172-179.

Frost, R., Armstrong, B.C., Siegelman, N. & Christiansen, M.H. (2015). Domain generality vs. modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19,* 117-125.

Gábor, B. & Lukács, Á. (*2012).* Early morphological productivity in Hungarian: evidence from sentence repetition and elicited production. *Journal of Child Language,* 39, 411-442.

Gauger L.M., Lombardino L.J., & Leonard C.M. (1997). Brain morphology in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*: 1272-1284.

Gillam, R. B., Montgomery, J.W., Gillam, S.R., & Evans, J.L. (2017) Working Memory in Child Language Disorders. In: Schwarz, R.G. (ed.) Handbook of Child Language Disorders. 2nd edition. Psychology Press. 213-237.

Gomez R.L. & Gerken L. (1999). Artificial grammar learning by 1-year-olds leads specific and abstract knowledge. *Cognition*, *70*, 109-135.

Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178-186.

Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*(2), 183-206.

Goodglass, H. (1993). *Understanding Aphasia.* Academic Press.

Goodglass, H., Kaplan, E., & Barressi, B. (2001).*The assessment of aphasia and related disorders*. Lippincott Williams & Wilkins.

Gopnik, M. & Crago, M. B. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39*, 1-50.

Greenfield, P. M. (1991). Language, tools, and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, 14(4):531--551.

Grodzinsky, Y. & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences, 12.12,* 474-480.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behav. Brain Sci. 23*, 1–21.

Gupta, P. & Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review*, *109*, 401-448.

Gupta, P. & Dell, G. S. (1999). The emergence of language from serial order and procedural memory.In B. MacWhinney (Ed.), *The Emergence of Language* , 28th Carnegie Mellon Symposium on Cognition. Mahwah, NJ, Lawrence Erlbaum.

Hauser, M. D., N. Chomsky, and W. T. Fitch. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, *298/5598*, 1569-1579.

Hernandez, A. ; Li, P. & Macwhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9* (5):220-225.

Hiscock, M. & Kinsbourne (1995). M. Phylogeny and ontogeny of cerebral lateralization. In: Davidson, R.J., & Hugdahl, K., (szerk.) *Brain Asymmetry*. Cambridge, MA: MIT Press. 535-578.

Hsu, N, Jaeggi, SM, & Novick, JM (2017). A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language,* 166, 63-77.

Huizinga, M., Dolan, C., & van der Molen, M. (2006). Age related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017-2036.

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

Jackendoff, R. (2002). *Foundations of Language: Brain, meaning, grammar, evolution.* Oxford and New York: Oxford University Press.

Janacsek, K., Fiser, J., & Ne´meth, D. (2012). The best time to acquire new skills: Age related changes in implicit learning across human life span. *Developmental Science*, 154, 496-505.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-100.

Johnston, J., & Kamhi, A. (1984). Syntactic and Semantic Aspects of the Utterances of Language-Impaired Children: The Same Can Be Less. *Merrill-Palmer Quarterly, 30*(1), 65-85.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Just, M. A. & Carpenter, P. A. (1996). The Capacity Theory of Comprehension: New Frontiers of Evidence and Arguments. *Psychological Review, 103*(4), 773-780.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension [Review]. *Trends in Cognitive Sciences*, *6*, 350-356.

Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective & Behavioral Neuroscience, 4*(1), 43-57.

Kan, I.P., & Thompson-Schill, S.L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. Cognitive, Affective and Behavioral Neuroscience, 4(1), 43-57.

Kapa, L. L., & Plante, E. (2015). Executive Function in SLI: Recent Advances and Future Directions. *Current Developmental Disorders Reports*, *2*(3), 245–252.

Karmiloff-Smith, A. (1996). Túl a modularitáson: A kognitív tudomány fejlődéselméleti megközelítése. In: Pléh Csaba (szerk.) *Kognitív tudomány.* Budapest, Osiris.

Kas B., & Lukács, Á. (előkészületben). *Magyar Mondatutánmondási Teszt*.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax*. Developmental Psychology, 48,* 171 - 184.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences, 97*(22): 11850-11857.

Ladányi, E. Kas, B. & Lukács, Á. (in press). The role of cognitive control in anaphor resolution in children with specific language impairment. *Applied Psycholinguistics*

Lai, C.S.L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated is a severe speech and language disorder. *Nature, 413*, 519-523.

Lee, B. & Pyun, S.B. (2014). Characteristics of cognitive impairment in patients with post-stroke aphasia. *Ann Rehabil Med., 38*(6): 759-65.

Lenneberg, E. (1974). A nyelv biológiai szempontból. In: Pap M. (szerk.): *A nyelv keletkezése.* Bp: Kossuth, 310-330.

Lenneberg, E.H. (1967). *Biological foundations of language*. New York: Wiley.

Leonard, L. (1998/2014). *Chilren with Specific Language Impairment*. Cambridge, MA: The MIT Press.

Leonard, L., & Deevy,P. (2010). Tense and aspect in sentence interpretation by children with specific language impairment. *Journal of Child Language, 37*, 397–418.

Leonard, L., Deevy, P., Miller, C., Charest, M., Kurtz, R., & Rauf, L. (2003). The use of grammatical morphemes reflecting aspect and modality by children with specific language impairment. *Journal of Child Language*, 31, 231-246.

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

Leonard, L., Sabbadini, L., Leonard, J., & Volterra, V. (1987). Specific language impairment in children: A crosslinguistic study. *Brain and Language*, *32*, 233-252.

Leonard, L.B., Bortolini, U., Caselli, M.C., McGregor, K.K., and Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, 2, 151-179.

Lewis, R. L., Vasisth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in cognitive sciences*, *10*(10), 447-454.

Lukács, Á. (2005). *Language Abilities in Williams Syndrome*. Budapest: Akadémiai.

Lukács, Á., Győri, M., & Rózsa, S. (2012). A TROG pszichometriai jellemzőinek magyar vizsgálata, a normák kialakítása. In.: Bishop, D.V.M. (2012). *TROG - Test for Reception of Grammar.* Handbook. OS Hungary Tesztfejlesztő Kft. 47-86.

Lukács, Á., Kas, B. és Pléh, Cs. (2014). A specifikus nyelvfejlődési zavar. In: Pléh, Cs. és Lukács, Á. (2014). (szerk). *Pszicholingvisztika 1-2. Magyar pszicholingvisztikai kézikönyv*. Budapest: Akadémiai. 1265-1325.

Lukács, Á., Leonard, L. B. & Kas, B. & Pléh, Cs. (2009). The Use of Tense and Agreement by Hungarian-Speaking Children with Language Impairment. *Journal of Speech, Language and Hearing Research*. 52/1, 1-20.

Lukács, Á., Kas, B. & Leonard, L. B. (2013). Case marking in Hungarian children with specific language impairment. *First Language*, 33/4. 331-353.

Lukács, Á., Pléh, Cs. and M. Racsmány (2004). Language in Hungarian children with Williams syndrome. In: Susanne Bartke and Julia Siegmüller (eds.). *Williams Syndrome across Languages*. Amsterdam: John Benjamins. 187-220.

Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, *48*(9), 1138-1154. http://doi.org/10.1016/j.cortex.2011.06.001

MacWhinney, B.(1999). (ed.) *The emergence of Language*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language,* 21(2), 339-366.

Marchman, V.A., Wulfeck, B. & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: a study of the English past tense. *Journal of Speech, Language, and Hearing Research, 42*: 206-19

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant Rule Learning Facilitated by Speech. *Psychological Science*, *18*(5), 387-391. https://doi.org/10.1111/j.1467-9280.2007.01910.x

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77-80.

Marshall, C. R., & van der Lely, H. K. J. (2007). Derivational morphology in children with grammatical-specific language impairment. *Clinical Linguistics & Phonetics* , *21*, 71-91.

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

Milham, M.P., Banich, M.T., & Barad, V. (2003). Competition for priority in processing increases prefrontal cortex's involvement in top-down control: An event-related fMRI study of the Stroop task. *Cognitive Brain Research,* 17, 212-222.

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

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100.

Miyake, A., Just, M. A. & Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language, 33*(2), 175–202.

Montgomery, J. W. (2003). Working memory and comprehension in children with specific language impairment: What we know so far. *Journal of Communication Disorders, 36*, 221-231.

Müller, R.A. (1996). Innateness, autonomy, universality? Neurological approaches to language. *Behavior & Brain Sciences, 19*, 611-675.

Németh D. (2006). *A nyelvi folyamatok és az emlékezeti rendszerek kapcsolata*. Budapest: Akadémiai Kiadó.

Norbury, C. F., Bishop, D., & Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *Journal of Speech, Language, and Hearing Research, 44*, 165 - 178.

Norman, D.A., & Shallice, T. (1986). Attention to action: willed and automatic control of behaviour. In Davidson, R.J., Schwartz, G.E., & Shapiro, D. (Eds), *Consciousness and Self-Regulation. Advances in Research and Theory* (pp.1-18). New York: Plenum.

Novick, J. M., Trueswell, J. C., & Thomspon-Schill, S. L. (2010). Broca's Area and Language Processing: Evidence for the Cognitive Control Connection. *Language and Linguistics Compass, 4*(10), 906-924.

Novick, J. M., Trueswell, J. C., &Thompson-Schill, S. L. (2005). Cognitive Control and Parsing: Reexamining the Role of Broca's Area in Sentence Comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 263-281.

Novick, JM, Kan, IP, Trueswell, JC, & Thompson-Schill, SL (2009). A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, 26(6), 527-567.

Oetting, J., & Horohov, J. (1997). Past-tense marking by children with and without specific language impairment. Journal of Speech, Language, and Hearing Research, 40, 62 - 74.

Peña, M., Werker, J. F., & Dehaene-Lambertz, G. (2012). Earlier speech exposure does not accelerate speech acquisition. *Journal of Neuroscience, 32*(33): 11159-11163.

Pinker, S. (1991): Rules of language. *Science, 253*, 530-535.

Pinker, S. (1999). A nyelvi ösztön. Budapest: Typotex.

Pléh, C. (2000). A modularitás és a pragmatika. *Erdélyi Pszichológiai Szemle, 1*(4), 9-30.

Pléh, C., & Lukács, Á. (2002). A magyar morfológia és a szabály viták. In F. Altrichter & K. Nyiri & C. Pléh & E. S. Vizi (Eds.), *Agy és tudat*. Budapest: BIP.

Pléh, Cs. (1998). *A mondatmegértés a magyar nyelvben.* Budapest: Osiris

Pléh, Cs. (2013). A lélek es a nyelv. Budapest: Akademiai Kiado.

Pléh, Cs., Á. Lukács & M. Racsmány (2003). Morphological patterns in Hungarian children with Williams syndrome and the rule debates. *Brain and Language*, *86*, 377–83.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle*, *60*, 479-505.

Rapin, I. & Allen, D.A. (1983). Developmental language disorders: nosological considerations. In: Kirk, U. (ed.) *Neuropsychology of language, reading and spelling*. New York: Academic Press.

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

Riccio, C. A., Cash, D. L., & Cohen, M. J. (2007). Learning and memory performance of children with Specific Language Impairment (SLI). *Applied Neuropsychology*, *14*(4), 255-261.

Rice, M., & Wexler, K. (1996). Towards tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research,* 39, 1239 - 1257.

Rice, M., Wexler, K., & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*, 850-863.

Rice, M., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 41*, 1412 - 1431.

Saffran, J. R. (2002). Constraints on Statistical Language Learning. *Journal of Memory and Language*, *47*, 172-196.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27-52.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*, 669-680.

Santi, A. and Y. Grodzinsky. "Working Memory and Syntax Interact in Broca's Area." Neuroimage, v"37 ll (2007).

Saygin, A.P., Dick, F, Wilson, S., Dronkers, N., Bates, E. (2003a). Neural resources for processing language and environmental sounds: Evidence from aphasia. Brain, 126.

Saygin, A.P., Dronkers, N., Wilson, S., Ludy, C. & Bates, E. (2003b). Pantomime interpretation and reading comprehension in patients with aphasia. 10th Annual Meeting of the Cognitive Neuroscience Society

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences, 106*(1), 322-327.

Schnur, T.T., Schwartz M.F., Brecher A., & Hodgson C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. Journal of Memory and Language, 54, 199-227.

Shafer, V.L., Yu, Y.H., & Garrido-Nag, K. (2012). Neural mismatch indices of vowel discrimination in monolingually and bilingually exposed infants: Does attention matter? *Neuroscience Letters, 526*: 10–14.

Squire, L.R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology , 44*, 453-495.

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

Tallal, P., & Piercy, M. (1973). Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality. *Neuropsychologia, 11,* 389-398.

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., & Merzenich, M.M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271,* 81-84.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.

Tomblin, J.B., Records, N.L. és Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research 39,* 1284-1294.

Trauner, D., Wulfeck, B., Tallal, P. és Hesselink, J. (1995). Neurologic and MRI profiles of language impaired children. Technical report CND-9513, Center for Research in Language, University of California at San Diego

Turi, Zs., Németh, D., & Hoffmann, I. (2014). Nyelv és emlékezet. In Pléh, Cs. és Lukács, Á. (szerk.) Pszicholingvisztika: magyar pszicholingvisztikai kézikönyv. 2. köt. Akadémiai Kiadó, Budapest. 743-776.

Ullman, M. T. (2001). A neurocognitive perspective on language: the declarative/procedural model, *Nature*, 717-726.

Ullman, M. T. (2001). A neurocognitive perspective on language: the declarative/procedural model, *Nature*, 717–726.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, *92*(1-2), 231-270. http://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T. (2016). The declarative/procedural model: A neurobiological model of language learning, knowledge and use. In *The Neurobiology of Language.* Elsevier.

Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment is not Specific to Language: the Procedural Deficit Hypothesis. *Cortex*, *41*(3), 399-433. http://doi.org/10.1016/S0010-9452(08)70276-4

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, *51*, 205-222. http://doi.org/10.1016/j.neubiorev.2015.01.008

van der Lely, H. & Stollwerck, L. (1997). Binding theory and specifically language impaired children. *Cognition, 62*, 245-29.

van der Lely, H. (2005). Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in Cognitive Sciences, 9*, 53-59.

van der Lely, H. és Stollwerck, K. (1997). A grammatical specific language impairment in children. *Brain and Language, 52*, 484-504.

van der Lely, H. K. J. (1994). Canonical linking rules: Forward versus reverse linking in normally developing and specifically language impaired children. *Cognition, 51*, 29-72.

van der Lely, H. K. J. (1997a). Narrative discourse in Grammatical Specific Language Impaired children: A modular language deficit? *Journal of Child Language, 24*, 221 - 256.

Vargha-Khadem F, Watkins K, Alcock K, Fletcher P, Passingham R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences, USA, 92*, 930 - 933.

Vargha-Khadem F, Watkins KE, Price CJ, Ashburner J, Alcock KJ, Connelly A, Frackowiak RS, Friston KJ, Pembrey ME, Mishkin M, Gadian DG, Passingham RE. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences, USA, 95*, 12695-12700.

Warren, R. M. (1970). Perceptual restorations of missing speech sounds. *Science*, *167*, 392-393.
Waters, G. S. & Caplan D. (1996). The capacity theory of sentence comprehension: critique of Just and Carpenter (1992). *Psychological Review*, 103, No.4., 761-772.
Watkins, R.V., Kelly, D.J., Harbers, H.M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38*:1349-1355.
Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: new directions. *Annual review of psychology, 66*: 173-196.
Whitehurst, G. J., Fischel, J. E., Arnold, D. S., & Lonigan, C. J. (1992). Evaluating outcomes with children with expressive language delay. In S. Warren & J. Reichle (Eds.), *Causes and effects in communication and language intervention*. Baltimore, MD: Paul H. Brooks Publishing. 277-313.
Windsor, J. (2017) Processing Speed, Attention, and Perception in Child Language Disorders. In: Schwarz, R.G. (ed.) Handbook of Child Language Disorders. 2nd edition. Psychology Press. 481-499.
Yu, C.; Smith, L. B. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science*, *18*(5): 414–420.