2013

Creating an SLI performance profile with load

Christy Marie Seidel
Louisiana State University and Agricultural and Mechanical College, seidel.christy@gmail.com

CREATING AN SLI PERFORMANCE PROFILE WITH LOAD

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Christy M. Seidel
B.S., Louisiana State University, 2007
M.A., Louisiana State University, 2010
December 2013

ACKNOWLEDGMENTS

To say that I was able to complete this body of work on my own would be an overstatement. I would like to take this opportunity to sincerely thank those listed below who have played such a vital role in making this dissertation possible.

To my advisor, Dr. Janet McDonald, I am ever thankful for your unwavering patience, wisdom, and invaluable feedback every step of the way. Without your guidance, I would have never gotten this far. To my dissertation committee—Drs. Brittan Barker, Emily Elliott, Janna Oetting, and Laura Choate, I am appreciative of your thoughtful comments and insight throughout this chapter of my graduate career. To all of my colleagues and mentors in the Psychology and Communication Sciences and Disorders departments, thank you for contributing to my growth and development during my transformation into a researcher. To my parents, Joseph and Lea Seidel, and cousin, Jesse Lind, I am thankful for your continued faith in me. To Kelly Tate, I am grateful for your friendship and humor. Finally, to Misty Smith, I offer my warmest thanks for always providing a constant stream of understanding, encouragement, inspiration, and support. As cliché as it sounds, I thank you for being my rock throughout this crazy journey and standing by my side to the end.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The purpose of this body of work was to gain a clearer understanding of the potential cognitive factors that may contribute to Specific Language Impairment (SLI). This study attempted to simulate a performance profile of SLI in typically developing children within a grammaticality judgment task, featuring structures historically difficult (third person singular -S and auxiliary BE) and easy (plural -S and progressive -ING) for individuals with SLI. Cognitive load was manipulated through the length of the sentences to be judged, and individual differences in phonological short term memory (PSTM) and working memory were measured (WM). For a successful simulation to occur, problematic structures should display lower performance than easier structures, particularly for longer sentences, even after taking into consideration individual differences in cognitive abilities. A successful simulation was not achieved as lengthening failed to polarize performance between the historically easier structures and historically difficult structures in the systematic way predicted, even after accounting for differences in working memory ability.

1. INTRODUCTION

Specific Language Impairment (SLI) is a term used to classify individuals who display typical nonverbal intelligence, yet struggle in areas of language (Leonard, 1998). It is estimated that approximately 7% of the kindergarten population meets the criteria for a diagnosis of SLI (Tomblin, Records, Buckwalter, Zhang, & Smith, 1997). Although attention has recently been given to SLI (Archibald & Gathercole, 2006; Briscoe & Rankin, 2009; Graf Estes, Evans, & Else-Quest, 2007), gaps in the literature still exist which need to be addressed. For example, it is still unclear which factors, be they environmental or physiological, drive the impairment. Furthermore, most of our current knowledge of SLI comes from various measures of spoken language production, and our understanding could be broadened by implementing different experimental tools. For these reasons, continued research on SLI is vital, first to expand our understanding of the impairment, and then to help guide clinicians in diagnosing and treating SLI both earlier and more successfully. The current study attempts to simulate a performance profile of SLI in typically developing children to uncover how external factors and personal limitations in working memory may contribute to the weaknesses shown in SLI. By doing so, this work represents an attempt to better understand the possible cognitive factors that may significantly influence SLI.

1.1 Specific Language Impairment Overview

Individuals with SLI display difficulty in many areas of language. For example, some individuals with SLI may display deficits in syntax, but additionally display greater difficulties in phonology or lexical retrieval (Friedmann & Novogrodsky, 2008; van der Lely, 2005). Such individuals are appropriately classified as having Phonological-SLI (Pho-SLI) or Lexical-SLI (Le-SLI; Friedmann & Novogrodsky, 2008). Other individuals display difficulty primarily with higher order language processes and are labeled as having Syntactic-SLI (S-SLI) or Grammatical-SLI (G-SLI; Friedmann & Novogrodsky, 2004; van der Lely, 1996; van der Lely & Christian, 2000; van der Lely & Harris, 1990). This classification shows the variability within a diagnosis of "specific" language impairment. Further, even within the domain of syntax, different distinctions between symptomology have been identified. For instance, some individuals with SLI show deficits in both comprehension and verbal expression, while others show difficulties with spoken language production only (Aram & Nation, 1975). Such variability prompted some to argue that SLI is too broad of a category to classify such a heterogeneous group of individuals, and others to advocate for the creation of subgroups (Aram & Nation, 1975; Wilson & Risucci, 1986; Wolfus, Moscovitch, & Kinsbourne, 1980; Korkman & Hakinen-Rihu, 1994). Many studies, however, do not use categories to distinguish between possible types of SLI. Among those studies not distinguishing between the possible subcategories, one general trend seems to prevail. That is, individuals with SLI display difficulty with certain grammatical morphemes: third person singular -S, auxiliary and copula BE, auxiliary DO, and past tense -ED (Rice, Wexler, & Cleave, 1995; Leonard, Eyer, Bedore & Grela, 1997; Hadley & Rice, 1996).

The most prevalent error type seen in individuals with SLI is the error of morpheme omission (Rice et al., 1995). For example, while a typically developing child may say, "he is running," a child with SLI may say "he running," omitting the auxiliary BE form. For additional examples of how typically developing children and children with SLI differ in terms of morpheme omissions, refer to Table 1 below.

Table 1: Differences in Morpheme Usage between Typically Developing Children and Children with SLI

Auxiliary BE	
Typically Developing	*Children with SLI*
"Today, he is playing games"	"Today, he playing games"
Copula BE	
Typically Developing	*Children with SLI*
"Long ago, they were gamers"	"Long ago, they gamers"
Auxiliary DO	
Typically Developing	*Children with SLI*
"He does play games"	"He play games"
Past Tense -ED	
Typically Developing	*Children with SLI*
"Yesterday, he played games"	"Yesterday, he play games"
Third Person Singular -S	
Typically Developing	*Children with SLI*
"He plays games"	"He play games"

High omission rates of these particular morphemes have been observed in many tasks. Such tasks include naturalistic language samples (Rice, Wexler, & Hershberger, 1998), controlled production measures, such as elicitation tasks (Rice et al., 1998), and even grammaticality judgment tasks (Poll, Betz, & Miller, 2010; Lum & Bavin, 2007; Montgomery & Leonard, 1998), where omissions are often accepted as grammatical.

For comparison, other error types involving the abovementioned morphemes are present, but only at very low rates. For example, errors of agreement (e.g., substituting IS for ARE) were present in samples of children with SLI (n = 15; age *M* = 5;2) in only 10% of responses (Leonard, Deevy, Miller, Charest, Kurtz, & Rauf, 2003). This finding was mirrored in grammaticality judgment task performance, where both impaired children (age *M* = 6;0) and adults (18;0 to 25;11) were less likely to accept inappropriate forms of target morphemes, such as WAS for WERE (Rice, Wexler, & Redmond, 1999; Poll et al., 2010). The fact that other syntactic errors are not as prevalent as omissions implies that the primary difficulty for children with SLI lies in knowing when, and not necessarily how, to properly use certain morphemes.

Other morphemes, however, such as plural -S and progressive -ING, seem to be relatively unproblematic (Rice et al., 1998; Rice et al., 1999), as seen in a variety of tasks. For example, in a naturalistic language sample, plural -S was marked at rates above 90% in obligatory contexts for children with SLI, as well as for their age- and language-matched peers (Rice et al., 1998). Also, in grammaticality judgment tasks, children with SLI were significantly more likely to correctly reject the omission of a progressive -ING (88%) than the omissions of third person singular -S, copula BE, and auxiliary BE structures (82%; Rice et al., 1999).

In summary, two central points concerning the performance patterns of SLI emerge. First, it appears that individuals with SLI have problems with morphemes that specifically code for tense and agreement. Second, it can be seen that not all error types are equally problematic for individuals.

The reason why morphemes marking tense and agreement prove to be difficult and lead to errors of omission, while other morphemes appear relatively unaffected, remains unclear. Drawing from Brown's (1973) seminal work on grammatical morpheme acquisition, it is worth noting that morphemes which are earlier acquired proved to be less vulnerable in both impaired and unimpaired populations (McDonald, 2008a; Rice et al., 1998; Rice et al., 1999) than those which are acquired later. According to Brown (1973), who studied the order of acquisition for fourteen different morphemes in three children, both progressive -ING and plural -S were acquired earlier than copulas, auxiliaries, past tense, and third person singular -S forms. Although the exact order of acquisition slightly differed between studies, de Villiers and de Villiers (1973), James and Khan (1982), and Khan and James (1983) all echoed the finding that the structures which prove to be less problematic for children with SLI were acquired before those that are more difficult. An additional testament to the difficulty of morphemes coding for tense and agreement comes from a multi-phase study from Leonard and colleagues, in which children with SLI ultimately received 96 intervention sessions, at 4 sessions a week, targeting either third person singular -S or auxiliary BE (Leonard, Camarata, Brown, & Camarata, 2004; Leonard, Camarata, Pawlowska, Brown, & Camarata, 2006; Leonard, Camarata, Pawlowska, Brown, & Camarata, 2008). Although intervention was deemed successful, with lasting effects, intervention efficacy might have been intertwined with natural maturation, and mastery of these morphemes was still not achieved (Leonard et al., 2004; Leonard et al., 2006; Leonard et al., 2008).

1.2 Theories behind Specific Language Impairment

Many theories strive to explain the patterns of performance observed in children with SLI. One theory, the Agreement and Tense Omission Model (ATOM) specifically focuses on the trends of tense and agreement morpheme omissions within spoken language as a function of a potential grammatical deficit (Rice et al., 1995; Rice et al., 1998). Other theories claim that SLI performance is not a function of a specific deficit in grammar or language, but rather is the reflection of a broader impairment in cognitive processing, such as a deficit in one's short term memory, working memory, or otherwise overall processing ability (Archibald & Gathercole, 2006; Briscoe & Rankin, 2009). Such claims are rooted in the evidence that children with SLI underperform in tasks of phonological or verbal short term memory and working memory when compared to their typically developing counterparts (Pickering & Gathercole, 2001; Archibald & Gathercole, 2006; Briscoe & Rankin, 2009; Graf Estes et al., 2007), and further, moderate correlations (r = .29 to .43) exist between working memory performance and syntactic performance (Montgomery & Evans, 2009; Engel de Abreu, Gathercole, & Martin, 2011).

The current study was designed in an effort to further clarify the cognitive-based viewpoint discussed above. It stands to reason that if a cognitive deficit underlies the impairment, then manipulations of stimuli which serve to reduce one's available cognitive resources ought to lead to a performance profile akin to those seen naturally in individuals with SLI. In an attempt to support this argument, I endeavored to simulate a profile of SLI performance in typically developing children by manipulating the processing demands of a grammaticality judgment task.

1.3 Simulating Performance Profiles of Special Populations

In attempting to argue for or against the differing theories explaining SLI, most studies focus on testing impaired children. While this is intuitive, an alternative approach would be to shift the focus away from the clinical population and towards stimuli manipulations that may recreate a profile of impaired performance in a typically developing population. This perspective may be particularly useful when attempting to support claims that cite cognitive processes as a possible underlining cause of SLI. In the past, such an approach has been insightful in studying other disorders and unique populations, such as aphasics (Blackwell & Bates, 1995; Bates, Wulfeck, & MacWhinney, 1991) and second language learners (McDonald, 2006; McDonald & Roussell, 2010).

Similar to individuals with SLI, individuals with aphasia make inappropriate omissions or opt to use uninflected word forms, which are considered symptoms of agrammatism (Blackwell & Bates, 1995). Agrammatism also includes the more frequent tendency to make agreement errors, and to a lesser extent, transposition errors (Blackwell & Bates, 1995). Such a pattern of performance was simulated in typically developing adults by Blackwell and Bates when a digit load secondary task was added to a primary grammaticality judgment task targeting determiners and auxiliaries within multiple error types, including agreement errors (e.g., "the writer were holding a very big party"), omission errors (e.g., "Mrs. Brown working quietly in the church kitchen") and transposition errors (e.g., "Miss Hope sending was several green dresses that Lisa had ordered"). Although no analyses were computed on overall performance collapsed across error types, a general trend emerged showing that performance dropped as a function of increasing digit load (no load M = 98.0, 2 digit load = 97.5, 4 digit load M = 97.6, 6 digit load M = 96.8). More interestingly, formal analyses revealed that target structures showed performance drops at different points of processing strain, reflecting the production profile of individuals with agrammatic aphasia (Blackwell & Bates, 1995; Bates, Wulfeck, & MacWhinney, 1991). That is, digit load most impacted agreement errors, followed by omission errors, and, to a lesser extent, transposition errors (Blackwell & Bates, 1995). From these results, it can be concluded that cognitive factors such as WM capacity could be responsible for the syntactic errors made by aphasiacs.

Performance of second language learners was also simulated in native English speakers (McDonald, 2006; McDonald & Roussell, 2010). Typically, late second language learners display difficulty with rejecting ungrammatical sentences within a grammaticality judgment task (McDonald, 2000). This has been observed for multiple constructions, including articles, regular and irregular past tense, third person singular –S, regular and irregular plural, progressive –ING, wh-questions, and yes-no questions (McDonald, 2006). In an attempt to explore the possible causal factors for this, McDonald (2006) examined multiple constructions in a grammaticality judgment task given to both native speakers placed under a variety of types of processing loads including added noise and a memory load; stressed native speaker performance was then compared to that of late second language learners. The constructions tested included those listed above, as well as word order, which was not shown to differ between unstressed native speakers and second language learners (McDonald, 2006). Performance by native speakers operating under noise (r = .64), or memory load (r = .67) showed significant correlations with that of second language learners across all constructions tested (McDonald, 2006). This correlation was not observed when comparing the performance of unstressed native speakers to second language learners. When focusing on specific constructions, all constructions tested were significantly impacted by either the addition of noise or additional memory load except for word order, which was previously observed to not differ between native speakers and second language learners. Thus, a profile of a late second language learner was successfully simulated. Imposing a deadline strain on native English speakers also led to a performance profile similar to that of a second language learner (McDonald & Roussell, 2010). These findings implicate limitations on one's processing ability as a potential explanation for the poorer grammaticality judgments of second language learners (McDonald, 2006; McDonald & Roussell, 2010).

Concerning the research conducted with aphasiacs, as well as second language learners, it is important to note that the meaningfulness of the results lies not in the mere decrease in performance, even of target structures. If typically easy structures fail to be robust against increases in processing load, the result would only reflect the effectiveness of a particular load instead of a simulation of a disorder. Therefore, meaningfulness of a set of results lies in the specific patterns of performance that emerge under load, with unaffected structures being equally as telling as those which are affected.

1.4 Simulating a Profile of Specific Language Impairment

To date, only one study was identified that has attempted to simulate performance of SLI in a typical population. Hayiou-Thomas, Bishop, and Plunkett (2004) had typically developing 6-year-old children engage in a grammaticality judgment task featuring grammatical structures, which are both historically problematic (third person singular –S and past tense –ED) and unproblematic (plural –S and prepositions *in*, *on*, and *at*) for children with SLI. If a profile of SLI were to emerge with an increase in processing load demands—i.e., if the first two structures suffer, while the latter two are relatively unaffected—it would lend support to those theories of SLI, which focus on a cognitive-based explanation. Processing load in Hayiou-Thomas et al.'s design was manipulated in two ways. The first manipulation involved the load of the sentence itself. Low load sentence versions, with a mean of 11.3 syllables, were transformed into high load versions, with a mean of 20.0 syllables, via the addition of irrelevant information to increase sentence length (Hayiou-Thomas et al., 2004). The second manipulation in Hayiou-Thomas et al.'s design focused on presentation rate of the sentence. Each sentence version was featured in both a natural and compressed state (Hayiou-Thomas et al., 2004). By compressing speech, participants were given less time to process and encode incoming information, making memory more susceptible to interference and decay. This manipulation was successful in the past for taxing processing abilities (McDonald, 2006). In fact, compressed speech has been shown to negatively impact the performance of multiple grammatical structures, including those included in Hayiou-Thomas et al.'s design (third person singular –S, past tense –ED, and regular plurals; McDonald, 2006). From these manipulations, four possible stimuli conditions emerged: short sentences-normal paced, long sentences-normal paced, short sentences-fast paced, long sentences-fast paced. Due to the between-subjects design used, each participant received sentences in only one of these four conditions.

Results of Hayiou-Thomas et al.'s (2004) study showed that both manipulations, increasing sentence length (η^2 = .37) and compressing speech (η^2 = .49), reduced performance, particularly for the structures historically seen as problematic for individuals with SLI (third person singular –S, past tense –ED). Also as predicted, plural –S proved to be resistant to both forms of stress (Hayiou-Thomas et al., 2004). Even so, a clean simulation of SLI was not obtained. While the effects of dual cognitive strain taxed third person singular –S and past tense –ED, errors involving the omission of prepositions were not as resistant to increasing processing load as expected (Hayiou-Thomas et al., 2004). As a result, the overall findings show only a partial profile of SLI performance.

The work of Hayiou-Thomas et al. (2004), however, does not go without criticism. Three specific points will be discussed below. The first two critical observations involve the methodology of Hayiou-Thomas et al.'s (2004) study, specifically the between-subjects design and the unsystematic lengthening of the sentences. The third criticism focuses on the specific findings concerning the control structures used, and the implications for interpreting the overall results.

The first concern revolves around the fact that processing load manipulations were treated as between-subjects variables, with each participant receiving only one of four possible combinations of length and speed (Hayiou-Thomas et al., 2004). Because participants' cognitive abilities were not measured a priori, group differences could exist between the four experimental conditions, possibly influencing the performance trends seen across load combinations. Not accounting for individual differences is problematic since a subset of typically developing children with lower cognitive abilities would theoretically require a lesser load than children with higher cognitive abilities to simulate the same SLI performance profile. For a design that aims to investigate the role of cognitive load manipulation, being able to account for a child's cognitive abilities is invaluable when interpreting differences in language task performance. Therefore, a stronger argument supporting the role of processing load in SLI could have been achieved if such fluctuations in performance were observed while manipulating load within-subjects, where each participant acts as his or her control subject.

A second concern is the way in which sentences were lengthened in this study. While Hayiou-Thomas and colleagues 2004) manipulated sentence length roughly by the same degree, it was done in an unsystematic fashion. A review of the example stimuli offered in the appendix showed that increases of sentence length feature multiple types of manipulations, including but not limited to, changing a pronominal subject (e.g., "we") to a lexical subject (e.g., "my sister"), word substitutions (e.g., "big" vs. "enormous"), adding adjectives to the subject (e.g., "the monster" vs. "the gigantic, wild, green monster"), and adding adjectives to the direct object (e.g., "kicks a big football" vs. 'kicks a big, round, yellow, plastic football"; Hayiou-Thomas et al., 2004). While the additional information surfaced in multiple areas of the sentence, some sentence phrases may have received more additional wording than others; these differences can be observed not only across sentence types, but within a sentence type as well. Table 2 offers two pairs of plural –S sentences from Hayiou-Thomas et al.'s appendix for comparison. Among the differences between these two sentences, it is noteworthy that in one sentence, added content focused on increasing the final prepositional phrase (e.g., "in the forest" vs. "in the big, dark, scary forest") while in the other sentence, added content was added between the numerical adjective and the direct object (e.g., "six pigs" vs. "six fat, pink, happy pigs").

Table 2: Example Plural –S Stimuli from Hayiou-Thomas et al. (2004)

Short Plural -S:	Yesterday, **we** saw three bears in the forest
Long Plural -S:	Yesterday, **my sister** saw three **brown** bears in the **big**, **dark**, **scary** forest.
Short Plural -S:	Last week, Tom saw six pigs in a **big** muddy field
Long Plural -S:	Last week, Tom saw six **fat**, **pink**, **happy** pigs in an **enormous**, muddy, **smelly** field.

From the literature, it is known that not all sentences are the same in terms of their processing demands, and introducing new information can add more or less cognitive load, pending on the length and location of the added information (Bock & Miller, 1991; Hartsuiker & Barkhuysen, 2006). For example, increasing cognitive load is particularly successful when the additional verbiage is interjected between the subject and verb for sentences focusing on subject-verb agreement, or when the information to be added is longer rather than shorter (Bock & Miller, 1991; Hartsuiker & Barkhuysen, 2006). When new information is added without strict control as to placement, it could theoretically result in some sentences presenting a greater increase in cognitive load compared to others. If sentences were more systematically lengthened, it would have offered greater assurance that performance fluctuations between structures were driven by the target structures themselves and not influenced by the position or nature of the additional information.

The last, and arguably most important, potential concern with Hayiou-Thomas et al.'s (2004) findings has to do with the effect of load on the control prepositional errors stimuli. Hayiou-Thomas et al. (2004) hypothesized that performance on both control structures, plural –S and prepositions, would to be resilient to increases in stress. Indeed, plural -S and prepositions behaved similarly robust under compressed speech in a two-way interaction between structure and speed. However, when the effects of speed and length were combined in a three-way interaction with structure, an unanticipated pattern of performance emerged. For a clean simulation to occur, both control structures (plural -S and prepositions) should have remained robust, even against the compounded load. While this was the case with plural -S, performance on the prepositions was affected more similarly to the target structures, since these three structures all displayed a significant interaction between speed and length (Prepositions: η^2 = .12; Third person singular -S: η^2 = .22; Past tense -ED: η^2 = .19). Therefore, while speed influenced both control stimuli types similarly, when length and speed were combined, the preposition control group no longer behaved like the robust plural control group, leading to only a partial SLI profile replication (Hayiou-Thomas et al. 2004).

This pattern of results raises an interesting question: would other nonimpacted morphemes, such as progressive -ING, act similarly to the pattern of performance seen for plural -S, or more akin to that seen for the omitted prepositions? Without this information, one could argue that perhaps Hayiou-Thomas et al.'s (2004) findings only suggest that increases in processing load potentially lead to overall performance decreases across different morphemes, with plural -S alone rising as a unique resilient structure, which would not be reflective of an SLI performance profile. Thus, while the overall finding from this article initially supports the role of taxing the processing system, the question remains whether a clean simulation of SLI can be achieved via increases in processing load. To address this concern, the current study examined a subsample of the structures tested by Hayiou-Thomas and colleagues, as an experimental check, as well as additional structures to gauge the reliability of their findings.

To address these concerns, the current study featured three notable differences from Hayiou-Thomas et al.'s (2004) methodology. First, the current study featured processing load manipulations conducted within-subjects, while additionally measuring individual differences in cognitive abilities (phonological short term memory and working memory). Second, the sentence stimuli used was lengthened systematically to ensure that performance differences will be due to the target structures and not potential item effects. Finally, the current study examined a subsample of the structures tested by Hayiou-Thomas and colleagues (vulnerable: third person singular -S, resilient: plural -S) as an experimental check, as well as additional structures to gauge the generalizability of their findings (vulnerable: auxiliary BE, resilient: progressive -ING). Compared to Hayiou-Thomas et al.'s design, the current study's adjustments afforded a more systematic and controlled way of gaining insight into the relationship between individual differences in cognitive abilities and language task performance.

1.5 Grammaticality Judgments

One strong aspect of Hayiou-Thomas et al.'s (2004) design was the use of a grammaticality judgment task, which has been listed as one of a few types of tasks that serve as clinical markers in identifying individuals with SLI (Poll et al., 2010). A "clinical marker" refers to performance on a particular task, shown to aid in the diagnosis of a disorder because it is based in the behaviors of the targeted impairment (Poll et al., 2010; Rice & Wexler, 1996). Ideally, the performance on a clinical marker task, such as grammaticality judgment, should successfully distinguish between impaired and unimpaired populations with little performance overlap. However, much of the research to date focusing on SLI has concentrated on measures of spoken language production, such as naturalistic language samples or imitation tasks. While these experimental tools have provided a strong foundation for our understanding of SLI, like all tasks, including grammaticality judgment, they are not free of criticism or limitations. More importantly, they do not directly inform us about an individual's language comprehension ability or his acceptance of certain grammatical structures.

While a language sample offers a genuine fragment of a child's linguistic and grammatical abilities, conversations may differ between participants and their experimenters, leading to a lack of experimental control. This lack of experimental control may result in differences in the rate of target morpheme production, with certain structures potentially not appearing frequently enough in a language sample to analyze. When morpheme tokens are produced at rates too low to examine, experimenters are forced to question whether the lack of morpheme production reflects the role of context or the ability of the child to produce the target structure. Additionally, it may be difficult to impossible for an experimenter to manipulate or introduce cognitive load within a naturalistic language sample.

More controlled tasks, such as sentence imitation, better allow for possible manipulation of cognitive load of stimuli. However, there is disagreement in the literature as to what sentence imitation tasks truly measure. While some believe that imitation tasks accurately reflect a child's grammar (Morehead, 1975; Santelmann, Berk, Austin, Somashekar, & Lust, 2002), others argue only surface processing occurs, suggesting that imitation tasks function more as a measure of short term memory (Dale, 1976). That is, if a child is required to repeat a complicated sentence and misses a crucial element during recall, the question arises whether this indicates systematic strain on the grammatical system or simply an overloading of short term memory. Additionally, due to the taxing cognitive demands on such high-processing load sentences, some children may only be able to repeat a few words, if anything at all. As a result, poor sentence recall for complex sentences only allows experimenters to comment on the overall success of the increase in load manipulation, but offers no specific information as to how the load impacted the target grammatical structures.

Language samples and sentence imitation tasks are appropriate for answering certain questions, such as how often does a child produce a target morpheme within a natural context or how reliably can a structure be produced even after primed with a target to repeat. For being able to scrutinize the cognitive based theories behind SLI, however, a measure is needed which offers maximum experimental control. A forced-choice grammaticality judgment emerges as a superior alternative to language production tasks since it enables all participants to be exposed to the exact same stimuli, and requires a simple response before continuing. First, in being able to examine the trends of syntactic violation acceptance within a controlled context, grammaticality judgment tasks can be used to test the relative difficulties of different grammatical structures, as will be discussed in 1.5.1. Second, as will be discussed in section 1.5.2, grammaticality judgments can determine which kinds of errors are most problematic for children with SLI, and under which conditions.

1.5.1 Grammaticality Judgments: Structure Difficulty

One function of grammaticality judgment tasks is to compare the relative performance of grammatical structures. Findings from grammaticality judgment tasks performed by children with SLI confirm what has been previously documented in earlier literature using production measures. That is, not all morphemes are consistently problematic for children with SLI, but those morphemes which are frequently problematic often involve tense and agreement, such as third person singular –S, auxiliary BE, and past tense -ED (Rice et al., 1995; Rice, et al., 1998; Rice et al., 1999; Montgomery & Leonard, 1998). Two studies in particular offer support for this assumption.

In the first study conducted by Montgomery and Leonard (1998), school aged children with SLI (*M* = 8;6) and both language and age matched controls engaged in a grammaticality judgment task focusing on omissions of third person singular –S, past tense –ED, and progressive -ING structures. Montgomery and Leonard's results showed that the combined performance accuracy of third person singular –S and past tense –ED differed between children with SLI (*M* = 82.4%) and their age matched counterparts (Age matched: *M* = 91.9%; Language matched: *M* = 85.5%); however, groups did not differ on the progressive –ING structure (SLI: *M* = 89.7%; Age matched: *M* = 87.8%; Language matched: *M* = 85.1%).

These trends of morpheme difficulty surface in even younger children (SLI: *M* = 6;0), as seen in a second grammaticality judgment study that featured previously examined (problematic third person singular –S and unproblematic progressive –ING; Montgomery & Leonard, 1998) and novel (problematic auxiliary BE) structures (Rice et al., 1999). Rice et al. offered an outline of performance during the study, including information on false alarm rates, when ungrammatical sentences were reported as being grammatical. When judging ungrammatical sentences featuring an omitted problematic morpheme such as third person singular –S or auxiliary BE, the false alarm rate for children with SLI was 32% (language-matched: *M* = 15%; age-matched: *M* = 5%; Rice et al., 1999). However, when judging an ungrammatical sentence featuring an omitted progressive –ING, the false alarm rate for children with SLI dropped to 13% (language-matched: *M* = 5%; age-matched: *M* = 0%), highlighting the relative ease of the progressive –ING structure for both the SLI and typically developing groups tested (Rice et al., 1999).

The overall findings from the studies above suggest a similar conclusion: structures involving tense and agreement (auxiliary BE, third person singular –S, past tense –ED) are especially problematic for children with SLI, while other structures, such as progressive –ING show less difficulty. Further, this trend was documented within a grammaticality judgment task in children as young as 6;0 (Rice et al., 1999).

Interestingly, most grammaticality judgment tasks routinely select the same select structures to examine. As would be expected, most designs include some of the structures long identified as being problematic, such as third person singular –S, past tense –ED, copula BE, or auxiliaries BE or DO. Progressive –ING and Plural –S frequently appear in grammaticality judgment designs as control structures since they are widely accepted as non-problematic for children with SLI (Rice et al., 1998). While being able to verify spoken language trends through a grammaticality judgment paradigm is informative, the examination of less researched structures offers additional information by which we can forward the collective understanding of this impairment.

Three structures, outside those listed above, have been examined. The additional structures tested include (1) determiners (i.e., "that", "which"; Wulfeck, Bates, Krupa-Kwiatkowski, & Saltzman, 2003) (2) Comparative –ER (Montgomery & Leonard, 2006), and (3) Possessive –S (Miller, Leonard, & Finneran, 2008; Montgomery & Leonard, 2006). Although empirical evidence has highlighted the particular difficulty of morphemes that code for tense and agreement, it is important to note that cognitive based theories do not limit problematic structures to any particular subset. Therefore, difficulty with additional morphemes, particularly when placed under cognitive strain, would lend support for a cognitive-based model.

In a study comparing performance on auxiliaries versus determiners, results indicated that children with SLI displayed significantly lower performance on errors (omissions, substitutions, movement) involving auxiliaries than determiners (Wulfeck et al., 2003). This supports previous empirical research showing that children with SLI are particularly sensitive to structures marking tense and agreement.

Although results from the previous study continued to show the difficulty of verbal morphology for individuals with SLI, some surprising results emerged when considering performance on comparative –ER. In a grammaticality judgment task, it was observed that both impaired and unimpaired children displayed greater difficulty with comparative –ER than progressive –ING (Montgomery & Leonard, 2006). Further, Montgomery and Leonard (2006) found that children with SLI underperformed compared to their typically developing peers on comparative –ER, but not on progressive –ING. These findings not only suggest that comparative –ER is a potentially difficult structure, but one that may pose exceptional problems for children with SLI (Montgomery & Leonard, 2006).

For additional consideration, two grammaticality judgment tasks found Possessive –S to also be an unusually problematic structure. In the first study conducted with both impaired (*M* = 9;0) and unimpaired (*M* = 8;11) children, performance on possessive –S and third person singular –S, in both natural and acoustically enhanced stimuli recordings, was compared (Montgomery & Leonard, 2006). Results showed that performance on possessive –S did not significantly differ from third person singular –S for either group—i.e., they were equally problematic (Montgomery & Leonard, 2006). Again, this indicates that morphemes outside the realm of those that mark for tense and agreement may be just as problematic for children with SLI. The comparative difficulty of possessive –S was later found in another study conducted on adolescents with SLI (age *M* = 15;9) and their age-matched peers (age *M* = 15;8; Miller et al., 2008). When performance on possessive – S was compared to progressive –ING, third person singular –S, and past tense –ED, results showed that omitted possessive –S displayed significantly lower performance than both omitted possessive –ING and omitted third person singular –S sentences (Miller et al., 2008). Further, although no group by structure interaction surfaced, within each individual structure, including both possessive –S and progressive –ING, adolescents with SLI performed worse compared to their age-matched counterparts (Miller et al., 2008). These findings suggest that, similar to comparative –ER, possessive –S may pose particular difficulty to children with SLI.

Most grammaticality judgment studies confirm the empirical research demonstrating the difficulty of structures coding for tense and agreement for children with SLI. However, it has come to light through using grammaticality judgment tasks that additional morphemes, which do not code for tense or agreement, and have also not been shown to be difficult for children with SLI may also pose a problem when placed in a grammaticality judgment task.

Besides testing how structures measure against each other at a given point in time during childhood, an additional way to test structure difficulty is to measure for how long structures remain problematic. A longitudinal study focusing on omissions of problematic BE and DO suggested that impaired children fail to catch up to their younger, language matched counterparts over time (Rice, Hoffman, & Wexler, 2009). Focusing on grammaticality judgment task performance, Rice et al. (2009) tested individuals with and without SLI over a period of 7 years on sentences which featured omissions of BE and DO. Growth curve modeling was then employed using initial testing to predict future performance. For the language match group, performance was predicted to fall within the .90-.95 range over time, while the range of performance for those with SLI was predicted to be between .75 and.80 (Rice et al., 2009). It was noted that these predictions were closely aligned with the actual observed results (Rice et al., 2009).

More recent research using grammaticality judgments affords us the knowledge that some structures remain problematic even past adolescence and into adulthood (Poll et al., 2010). In a rare study focusing on adults with and without SLI (age *M* = 21;0), Poll and colleagues (2010) examined subject-verb agreement errors (auxiliary ARE for auxiliary IS) and omission errors (omitted auxiliary IS), as well as progressive –ING in both simple and complex (e.g., embedded relative clause) sentences using an A’ statistic. This statistic takes into consideration both hit rates and false alarms, and ranges in value from .5 (chance performance) to 1.0 (ceiling performance; Stanislaw & Todorov, 1999). Echoing prior research, Poll et al. showed that adults with SLI were just as sensitive to violations of subject-verb agreement (Typically Developing Median *A′* = 1.00; SLI Median *A′* = 1.00) and progressive –ING (Typically Developing Median *A′* = 1.00; SLI Median *A′* = 1.00) as their typically developing peers within complex sentences. For historically problematic, structures, however, Poll et al. found the increase of sentence load was able to differentiate between clinical groups. While both groups of adults were equally able to reject ungrammatical structures featuring a dropped problematic morpheme (Typically Developing Median *A′* = 1.00; SLI Median *A′* = 1.00; Poll et al., 2010) for simple sentences, a significant group difference emerged for complex sentences featuring a dropped problematic morpheme (Typically Developing Median *A′* = 1.00; SLI Median *A′* = .95; Poll et al., 2010). This was supported by a large effect size (*r* = .54; Cohen, 1992; Poll et al., 2010). While increases in sentence complexity did not correspond with decreasing performance for unimpaired adults, adults with SLI were more likely to accept ungrammatical sentences with a problematic omission as correct (Poll et al., 2010). Although the performance of the impaired adults is almost at ceiling, the point to be gleaned from this study is that statistical differences in performance remain even in adulthood for problematic structures.

1.5.2 Grammaticality Judgments: Errors

Another purpose of grammaticality judgments is to test which kinds of errors most often go undetected and what conditions promote poor performance. From measures of production, it is known that frequently dropped markers of tense and agreement are the hallmark of children with SLI (Rice et al., 1995). Evidence from grammaticality judgments is consistent with these findings. For example, children with SLI were more likely to accept an ungrammatical sentence as grammatical when the error in question was a dropped morpheme, such as a dropped third person singular –S (e.g., "He jump"), rather than an agreement (substitution) error (e.g., "I jumps"; Rice et al., 1999).

However, recent research has suggested that agreement errors (e.g., "was" for "were") may be more problematic than previously thought. In an elaborate grammaticality judgment design given to children with and without SLI (age ranges: 7-8 years, 9-10 years, 11-12 years), performance on auxiliaries and determiners (demonstrative adjectives, numerals) was examined as a function of both error type and error location (Wulfeck et al., 2004). The three error types Wulfeck et al. (2004) examined included errors of agreement or substitution (e.g., "The writer were…" or "A boys are…"), errors of movement (e.g., "Miss Hope sending was…" or "Helicopter a was…"), and errors of auxiliary or determiner omission (e.g., "Mrs. Brown working…" or "Girl was working…"). Results showed third person agreement errors (*A′* = .77) to be the most difficult, movement errors (*A′* = .82) to be the least difficult, and omission errors (*A′* = .79) to not differ from either (Wulfeck et al., 2004). This finding, however, was qualified by an upper level interaction, driven by the impaired sample, such that the rate of performance on movement errors increased faster than that of verb and determiner agreement errors as children got older (Wulfeck et al., 2004). This is curious as other research on SLI indicated that omission errors, not agreement errors, are the most problematic (Rice, Wexler, & Redmond, 1999; Poll et al., 2010).

In addition to manipulating error type, Wulfeck and colleagues (2004) manipulated error location within the sentence, showing that performance is not solely dependent on the type of syntactic violation. For all sentence types, errors were either placed early in the sentence (e.g., "Mrs. Brown working in the church kitchen") or later (e.g., "She had written that mystery novel that her mother reading"; Wulfeck et al., 2004). Globally, it was observed that errors appearing earlier in the sentence (*A′* = .77) appeared to be more problematic (later errors: *A′* = .81; Wulfeck et al., 2004). Upon further inspection, Wulfeck et al. found that syntactic error location appeared to be especially influential for agreement error performance. This finding is meaningful as it proves that location within the sentence can play a vital role in the degree to which a structure appears problematic.

When including evidence from other methodologies, the traditional stance that omission errors are the *most* problematic error type for children with SLI appears to be upheld. However, findings from Wulfeck et al. (2004) indicate that omissions may not be the only problematic error type worthy of investigating. From this research, it can also be gleaned that special consideration must be paid not only to the morpheme in question, or the type of error involved, but also to the syntactic context surrounding the error and the subsequent effects on cognitive load it contributes.

From the findings gleaned through grammaticality judgment tasks, two general points surface. First, even though morphemes involved in tense and agreement marking are exceptionally problematic for children with SLI, they may not be exclusively problematic. Second, while omissions may still be the most prevalent error type observed for children with SLI, other error types, such as errors of agreement, may be more problematic than once thought.

1.5.3 Grammaticality Judgments and Cognitive Processes

Interestingly, the grammaticality judgment task is arguably strongly linked to the control of cognitive processes, making this methodology especially relevant by which to examine alternative theories of SLI rooted in more cognitive explanations (Bialystok & Ryan, 1985). It has been proposed that a grammaticality judgment task is the combination of two operations: analysis and control (Bialystok & Ryan, 1985). First, when an individual encounters a sentence to be judged as acceptable, he must reflect on his knowledge of syntax, and in essence, explicitly review the naturally implicit knowledge of acceptable grammar; this is referred to as analysis (Bialystok & Ryan, 1985). Second, he must inhibit all irrelevant information such as superfluous adjectives, prepositional phrases, or semantic violations, and solely focus on the syntactic content; this is referred to as control (Bialystok & Ryan, 1985).

Lum and Bavin (2007) conducted a grammaticality judgment task with school aged (8;6 to 10;5) children with SLI to test Bialystok and Ryan's (1985) theory of analysis and control. It was predicted that children with SLI would display more difficulty with the process of analysis, particularly as it relates to historically problematic structures (Lum & Bavin, 2007). Sentences featured in the grammaticality judgment task used morphemes both problematic (third person singular -S and past tense -ED) and unproblematic (progressive -ING) for impaired children in both plausible and implausible sentential contexts (Lum & Bavin, 2007). To examine the process of analysis, Lum and Bavin assessed accuracy on the judgments of only semantically plausible sentences. Because only semantically plausible sentences were used to test "analysis," participants had to make grammaticality judgments on syntactic structure without having to additionally inhibit conflicting semantic information within the sentence. First, there was a main effect of clinical status, with SLI children performing lower than their typically developing counterparts (partial η^2 = .199; Lum & Bavin, 2007). There was also a main effect of structure, with progressive -ING proving to be the easiest across both groups of participants (partial η^2 = .197; Lum & Bavin, 2007). Although a statistically significant interaction between group and structure did not emerge as expected, the performance differences between the problematic structures, third person singular -S (A' = .82) and past tense -ED (A' =

.84), and the easier progressive -ING structure (*A′* = .92) were more polarized for children with SLI than their typically developing peers (third person singular -S: *A′* = .94; past tense -ED: *A′* = .92; progressive -ING: *A′* = .98; Lum & Bavin, 2007)

To investigate control, Lum and Bavin (2007) examined performance on both semantically plausible and implausible sentences. By including implausible sentences, participants would be required on some trials to additionally inhibit contradicting semantic information while honing in on any pertinent syntactic violations. In certain working memory models (Cowan, 1988), the mechanisms of working memory have been described as the ability to keep certain information within the focus of attention while inhibiting distracting information. It stands to reason that if working memory deficits influence language task performance in children with SLI, we would anticipate the additional strain of inhibiting semantic information to prove exceptionally difficult. As expected, it was found that implausible sentences resulted in more errors for children with SLI (Third Person Singular -S: *A′* = .70; Past Tense -ED: *A′* = .73; Progressive -ING: *A′* = .74) than the typically developing control group (Third Person Singular -S: *A′* = .92; Past Tense -ED: *A′* = .87; Progressive -ING: *A′* = .93), indicating that the children with SLI were less able to inhibit semantic distraction (Lum & Bavin, 2007). By requiring the additional cognitive process of control, the ability to focus on syntactic violations (i.e., analysis) was negatively affected (Lum & Bavin, 2007). The typically developing children, on the other hand, were more successful at performing both analysis and control processes simultaneously (Lum & Bavin, 2007). To summarize, performance dropped when encountering problematic structures in plausible contexts for all children (Lum & Bavin, 2007). When implausible sentences were also included, forcing children to tap into the additional process of control, children with SLI in particular had a significant performance decrease for all structures, including the historically unproblematic progressive -ING (Lum & Bavin, 2007). From this finding, it can be assumed that working memory, or some broader cognitive ability, may be partially responsible for the performance differences between typically developing and impaired children in grammaticality judgment tasks. Unfortunately, cognitive individual differences were not measured in this study; without knowing the potential disparity in WM abilities between the impaired and unimpaired samples, the degree of WM's influential role is left to speculation.

Even though significant structure differences emerged in both conditions, it should be noted within the SLI group that not only did performance decrease overall as a function of dual cognitive loads (Analysis and Control) but smaller performance gaps between the different structures were observed (Lum & Bavin, 2007). For example, the widest performance gap in the analysis condition was between the performance on progressive –ING (*A′* = .92) and third person singular –S (*A′* = .82). For comparison, the widest performance gap in the analysis and control condition was less than half of the previous difference (progressive –ING: *A′* = .74; third person singular –S: *A′* = .70). These smaller performance gaps may be a reflection of the interaction between inherent load of the structure and the external demands of the task. From the literature, it is known that certain structures repeatedly show lower performance than others on language tasks. Montgomery and Leonard (2006) discuss a list of possible reasons for these discrepancies including when certain structures are acquired, the nature of the structure, or even the phonological saliency of the structure. When certain structures are then put under cognitive load, even some "easier" structures could theoretically become less robust. In the case of the above study by Lum and Bavin (2007), target structures were placed in implausible sentences, requiring the participants to exercise control. That is, participants had to block their knowledge of semantics and plausibility and hone in on the syntactic information alone. As a result of this extra cognitive load, performance on seemingly less problematic progressive –ING failed to differ from the historically more difficult structures. This pattern was previously seen by Hayiou and colleagues (2004), where performance on prepositions mirrored that of problematic third person singular –S and past tense –ED when placed under dual load.

Lum and Bavin (2007) did not analyze the performance gap between impaired and unimpaired children in plausible versus implausible sentences; however, the large numerical trends should be noted. The A′ performance gap between typically developing children and those with SLI ranged from 6 to 12 for plausible sentences only requiring the process of analysis (Lum & Bavin, 2007). For comparison, when implausible sentences were introduced, thus requiring the additional process of control, the A′ gap range increases from 14 to 22 (Lum & Bavin, 2007). It can be speculated that juggling two concurrent processes, one of which is inhibiting information, is more taxing for children with SLI in grammaticality judgment tasks.

Literature focusing on the ability of children with SLI, ranging in age from 4;0 to 5;4, to inhibit information in a stop-signal task offers some enlightenment (Spaulding, 2010). In the stop-signal task, preschool children were required to click a corresponding picture button when hearing the words "butterfly" or "dinosaur", but to inhibit a response when the target stimuli were followed by the word "stop" (Spaulding, 2010). It was observed that children with SLI, compared to typically developing children, displayed lower levels of inhibition and resistance to distractor information (Spaulding, 2010). This trend persisted even after contributions of nonverbal cognition were controlled for (Spaulding, 2010). The fact that children with SLI may have difficulty with inhibition offers a potential explanation for why the introduction of sentences requiring control in Lum and Bavin's (2007) design may have functioned as such a successful cognitive load.

1.6 Deficits in Cognitive Processes in Individuals with SLI

Speculation has long existed that a deficit in cognitive abilities, in one area or another, may be the root cause of SLI. Two cognitive functions in particular—verbal short term memory and working memory—have been examined as potential factors which may greatly influence the impairment (Archibald & Gathercole, 2006; Briscoe & Rankin, 2009). Verbal short term memory refers to the simple storage of auditory information (Baddeley, 1986). In contrast, working memory is the ability to not only store, but also manipulate information (Baddeley, 1986).

According to a modular model of working memory, the abilities to store and manipulate would represent independent processes, not drawing from a common pool of resources (Baddeley, 1986; Baddeley, Allen, & Hitch, 2010). Therefore, it would be quite possible to display deficits in one area, while appearing fully functional in another. In line with a more dynamic perspective of working memory, however, it seems intuitive that having a deficit in one area may translate to a deficit in another. This is because in alternative models of working memory, one's capacities are not divided into individual stores, but rather represent a pool of shared resources (Cowan, 1988; Bunting & Cowan, 2005). For example, if one cannot appropriately store information in short term memory, one would speculate that this would later be reflected in a measure where the information needs to be both stored and manipulated (Briscoe & Rankin, 2009).

The majority of articles which investigate verbal short term memory and working memory in SLI do so by examining these processes separately. Therefore, the following two sections will be devoted to reviewing the evidence for and against verbal short term memory and working memory as potentially influential factors of SLI.

1.6.1 Verbal Short Term Memory and Nonword Repetition in Individuals with SLI

As addressed earlier, most verbal short term memory tasks require simple storage and repetition of the given information (Baddeley, 1986). Examples of such tasks include digit recall or word list recall tasks, as featured in the *Working Memory Test Battery* (*WMTB-C*; Pickering & Gathercole, 2001). Yet, the majority of verbal short term memory tasks draw upon stored lexical information. Therefore, arguably, the cleanest measure of verbal short term memory would be the nonword repetition task, which has been identified as a clinical marker of SLI (Poll et al., 2010). Due to the nature of this task, some refer to this task not as measuring "verbal" short term memory but rather "phonological" short term memory (Gathercole, Tiffany, Briscoe, Thorn, & ALSPAC Team, 2005) since nonword repetition features phonological sequences most likely never encountered before. Therefore, nonword repetition maximizes being able to measure one's abilities to perceive, encode, and retrieve speech information, void of major contributions from lexical knowledge, aside from phonological probabilities.

A plethora of research has shown that performance on a nonword repetition task can distinguish between individuals with and without SLI (Montgomery & Evans, 2009; Weismer, Tomblin, Zhang, Buckwalter, Chynoweth, & Jones, 2000; Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990), including children as young as preschool aged (Deevy, Weil, Leonard & Goffman, 2010). A few versions of the nonword repetition tasks appear in the literature, including the *Children's Test of Nonword Repetition* (*CNRep*; Gathercole & Baddeley, 1996) and the *Nonword Repetition* (*NRT*) task by Dollaghan and Campbell (1998). The differences between these two tasks, and others, such as wordlikeness, word length, and articulatory complexity, were reviewed in a meta-analysis by Graf Estes, Evans, and Else-Quest (2007). The *CNRep* (Gathercole & Baddeley, 1996), for instance, consists of 40 two- to five-syllable nonwords of mixed wordlikeness and mixed articulatory complexity (Graf Estes et al., 2007). That is, nonwords feature both single consonants and consonant clusters (Graf Estes et al., 2007). The *NRT* of Dollaghan and Campbell (1998), however, consists of only 16 nonwords spanning in length from one to four syllables and was designed specifically to have no consonant clusters, no repeating vowels or consonants, and low wordlikeness, with consonants having a phonotactic probably of less than 25% for their given position (Graf Estes et al., 2007). From a glance, it is clear that these two tasks are quite different. The *CNRep* (Gathercole & Baddeley, 1996) offers more exemplars and a greater number of syllables, therefore, may appear to be a more sensitive measure of one's phonological short term memory ability, or lack thereof. However, by including nonwords of high wordlikness, or nonwords, which may even contain small English words within them, the *CNRep* calls to question whether only phonological short term memory is being measured (Graf Estes et al., 2007). One could argue that a participant may use high wordlikeness or embedded English words to aid in recall, thus potentially confounding a measure of pure phonological memory (Graf Estes et al., 2007). The *NRT*, although featuring fewer stimuli and with less syllables, has addressed these concerns by reducing wordlikeness (Dollaghan & Campbell, 1998).

Although the two tasks above greatly differ, all nonword repetition tasks examined within a meta-analysis, including the *CNRep* and *NRT*, were able to distinguish between individuals with and without SLI albeit to different extents (Graf Estes et al., 2007). In these studies, phonological short term memory, as measured by various nonword repetition tasks, in individuals with SLI was depressed compared to typically developing peers. However, Graf Estes and colleagues (2007) warn that just because typically developing individuals outperformed impaired individuals across tasks does not mean that the given tasks are completely analogous due to differences in design, discussed above, and corresponding effect sizes.

Although literature trying to unravel the relationship between SLI individuals' phonological short term memory ability and language performance is scarce, a few studies focusing on understanding how the two are intertwined have led to mixed results. One study in particular focused on the relationship between nonword repetition performance and performance on the *Clinical Evaluation of Language Fundamentals-Revised* (*CELF-R*; Semel, Wiig, & Secord, 1987), which is divided into two components: the *CELF-RLS* and *CELF-ELS*, respectively measuring receptive (e.g., following directions, understanding conceptual relationships) and expressive language (e.g., sentence repetition, ability to produce grammatical sentences, ability to produce appropriate morphemes given context; Montgomery & Windsor, 2007). The *CELF-RLS* and *CELF-ELS* do not exclusively test any specific structure or syntax element, but rather examine language ability within a broader context (Semel et al., 1987). Results showed that even after the effects of age were removed, significant positive correlations persisted between nonword repetition task performance and both expressive and receptive language measures of the *CELF-R* for children with SLI (age *M* = 8;9) but not for typically developing children (age *M* = 8;8; Montgomery & Windsor, 2007).

Other studies focused on the relationship between nonword repetition performance and sentence comprehension. Although one study (Montgomery, 2004) failed to find significant correlations between nonword repetition performance and comprehension in either impaired or unimpaired samples, both prior and more recent research suggest that phonological short term memory may play a role in sentence comprehension. In an earlier study, a significant positive correlation (r = .62) was observed between nonword repetition task performance and sentence comprehension (not focusing on any specific morpheme structure) when they collapsed across typically developing and impaired children (Montgomery, 1995). It should be noted, however that by failing to investigate each group separately means that it is possible that one group, SLI or typically developing, was driving the significant finding. In a more recent study, the relationship between nonword repetition and sentence comprehension was analyzed separately for impaired children (age M = 9;1) and their language and age-matched counterparts (Montgomery & Evans, 2009). While nonword repetition performance did not correlate with comprehension of simple or complex sentences in either typically developing group, a significant correlation emerged for children with SLI (Montgomery & Evans, 2009). Specifically, simple sentence comprehension (M = 80.6), but not complex sentence comprehension (M = 74.5), correlated with nonword repetition performance for impaired children (Montgomery & Evans, 2009). This finding reinforces the idea that perhaps phonological short term memory may play a role in language task performance in children with SLI.

In summary, when focusing solely on nonword repetition task performance, the finding that children with SLI display less accurate nonword recall compared to their typical counterparts is consistent (Graf Estes et al., 2007). This performance difference is seen regardless of which nonword repetition task is used (Graf Estes et al., 2007). When investigating the relationship between phonological short term memory and language task performance, however, two trends seem to emerge. First, nonword repetition task performance does not seem correlated with language task performance in typically developing children (Montgomery & Windsor, 2007; Montgomery & Evans, 2009). Second, there is some evidence that nonword repetition task performance positively correlates with language measures in children with SLI, even after removing the effects of age (r = .29 to .53; Montgomery & Windsor, 2007; Montgomery and Evans, 2009). It is important to note, however, that in Montgomery & Evans' (2009) study, complex sentence performance did not correlate with PSTM for the SLI group. Therefore, the degree of influence PSTM plays in language task performance may be overshadowed by other cognitive factors, such as working memory, discussed below.

1.6.2 Working Memory in Individuals with SLI

Working memory, as reviewed earlier, is the ability to manipulate stored information (Baddeley, 1986). As can be expected, many tasks exist which strive to quantify this ability. Two tasks in particular frequently appear in the SLI literature. The first is the *Competing Language Processing Task* (*CLPT*; Gaulin & Campbell, 1994), which is an adaptation of Daneman and Carpenter's (1980) listening span task. In this task, participants listen to a string of statements, judging their truthfulness, and remembering the final word of each statement (Gaulin & Campbell, 1994). An additional measure of working memory seen in the literature is the size judgment task (Montgomery, 2000a, 2000b). In the size judgment task, individuals are presented with a list of concrete nouns that they are required to recall (Montgomery, 2000a, 2000b). In a no load condition, participants are asked to engage in free recall (Montgomery, 2000a, 2000b). In essence, this is comparable to a verbal short term memory task, as no manipulation of the information is required. In the case of the single-load condition, participants must simply relist the words they hear from smallest physical object to largest (Montgomery, 2000a, 2000b). In the case of dual-load condition, participants must first divide the words into semantic categories, such as animacy, and then sort the items from smallest to largest within each category (Montgomery, 2000a, 2000b). Because information is both being retained as well as manipulated, the size judgment task in either of the load conditions provides a measure of one's working memory span. In addition, size judgment is an appropriate working memory measure to use alongside experiments measuring language performance as the task itself is linguistic in nature but not syntactic. This contrasts with the listening span task (Daneman & Carptenter, 1980) in which children must use comprehension skills to judge sentences as true or false.

Results focusing solely on working memory task performance support the speculation that children with SLI may suffer from a deficit in working memory. This is because children with SLI display lower levels of performance than typically developing counterparts on multiple working memory measures. Using the *CLPT* (Gaulin & Campbell, 1994), Montgomery and Evans (2009) demonstrated that individuals with SLI (age M = 9;1) differed in performance from age-matched, but not language-matched peers. Using the size judgment task, specifically focusing on the dual load condition, Montgomery (2000a, 2000b) showed that individuals with SLI were outperformed by their age-matched counterparts. Working memory differences between impaired and unimpaired populations have also been found by Archibald and Gathercole (2006) using original complex span measures found in the *Working Memory Test Battery* (*WMTB-C*; Pickering & Gathercole, 2001), such as listening recall, counting recall, and backward digit recall.

Aside from investigating the differences between impaired and unimpaired children on tasks of working memory, researchers have also explored the relationship between working memory and linguistic task performance, leading to mixed results. The results for, and then against, the positive relationship between working memory and SLI and language task performance is discussed below.

Generally speaking, there appears to be a positive trend between one's working memory span and one's ability to perform successfully on language tasks, regardless of clinical status. Using a comparatively large sample size (N = 58), it was found that working memory performance on the *CLPT* and sentence comprehension correlated for both children (age M = 9;1) who are impaired (r = .43) and their language-matched counterparts (r = .31), even after the effects of age were removed (Montgomery & Evans, 2009).

Two additional studies focusing solely on typically developing children also documented positive correlations between language task performance and working memory, as measured by the *CLPT*. In the first study, results from 112 third graders (age M = 8;9) documented that performance on a listening span working memory measure correlated with grammaticality judgments (r = .44) and syntactic corrections (r = .47; Gottardo, Stanovich, & Siegel, 1996). In addition, working memory also explained the largest amount of unique variance seen for reading comprehension (12.5%) compared to syntactic processing ability (1.5%) and phonological sensitivity (5.0%; Gottardo et al., 1996). In the second study focusing on 65 children ranging in age from 6 to 12 (age M = 8;6), sentence comprehension performance positively correlated with both an easier (r = .46) and harder version (r = .35) of the listening span task, even after removing the effects of age (Magimairaj & Montgomery, 2012). The processing demands of the listening span task were manipulated by including both easier sentences, featuring traditional subject-verb or subject-verb-object sentences, and more difficult object clefts (Magimairaj & Montgomery, 2012). A follow-up regression analysis even indicated that the easier listening span task was more predictive of sentence comprehension (Magimairaj & Montgomery, 2012). Magimairaj and Montgomery (2012) believed this to be because the easier listening span task appears to capture a more pure measure of processing and attentional control capabilities, while the more complex listening span task may have inadvertently involved verbal short term memory as well. Taken together, this research indicates that in both impaired and unimpaired populations, positive links may exist between language task performance and working memory.

However, not all research investigating the association between working memory and language task performance results in positive relationships. In one study of children with SLI (age *M* = 8;6), sentence comprehension and performance on the size judgment working memory measure were not significantly correlated for those with SLI or their age-matched or language-matched controls (Montgomery, 2000b). This was assumed by Montgomery (2000b) to be because of small sample sizes ($n = 12$), which would lead to low statistical power. Specifically for the SLI group, Montgomery (2000b) suggested the lack of a significant correlation could be due to the overall difficult nature of the task, which could have exceeded the children's processing abilities. Although another study, using the same sample size (12 participants per group), did report a positive significant correlation between size judgment performance and off-line sentence comprehension for the typically developing control group ($r = .47$), more curiously, an unexpected negative correlation was observed for those impaired with SLI ($r = -.43$; Montgomery, 2000a). Although the effect size of this negative correlation was not reported, Montgomery (2000a) suspects this negative correlation was due to factors aside from working memory ability, such as trace decay or rapid phoneme identification, which contribute to poor comprehension performance. Although this explanation seems plausible, the observed negative correlation should be viewed with some skepticism due the small sample sizes of this study, which was presumed in the previous study (Montgomery, 2000b) to possibly carry some responsible for the complete absence of a correlation.

To summarize, it is well documented that children with SLI display lower performance on various measures of working memory than their typically developing counterparts (Archibald & Gathercole, 2006; Montgomery, 2000a; Montgomery, 2000b; Montgomery & Evans, 2009). Studies focusing on the relationship between working memory and language task performance, however, lead to mixed results. Some studies support the notion that working memory shares a positive relationship with language task performance, both for typically developing and impaired children (Montgomery & Evans, 2009; Gottardo et al., 1996; Magimairaj & Montgomery, 2012). Other studies fail to show such a positive correlation (Montgomery, 2000a; Montgomery, 2000b). There are a few differences, which ought to be mentioned between the studies that find positive correlations and those that do not. First, the studies that find positive relationships almost exclusively use a listening span task to measure working memory ability. The use of a listening span task is worthy to note because this measure involves syntactic processing to judge whether sentences are true. This kind of syntactic processing is also being tested in the dependent measure language tasks examined. Therefore, it comes to little surprise that this particular working memory measure is more likely to correlate with language task performance than a size judgment measure, which is void of syntax. The second difference between the studies focuses on the size of the sample being tested. In studies that failed to find a correlation, or found a negative correlation, the sample studied was very small. Therefore, results from those studies should be viewed in light of their sample size limitations.

1.6.3 Verbal Short Term Memory versus Working Memory in Individuals with SLI

Research has been conducted looking at the relationships between language task performance and verbal short term memory or working memory separately. However, it has not been until recently that both verbal short term memory and working memory, as represented by the executive control portion of Baddeley's (1986) working memory model for purposes of this paper, have been explored together in impaired individuals. Two articles have been identified which do so, arriving at similar results, but conflicting conclusions.

Archibald and Gathercole (2006) administered the *Working Memory Test Battery* (*WMTB-C*; Pickering & Gathercole, 2001) to school aged (6;11 to 11;10) children with SLI. The *WMTB-C* (Pickering & Gathercole, 2001) comprises three subsets, each with multiple tasks designed to test a particular dimension of Baddeley's (1986) working memory model: verbal short term memory (digit recall, word list recall, non-word list recall, and word list matching), visual short term memory (block recall, mazes memory, visual patterns test), and executive control (listening recall, counting recall, and backward digit recall). Comparing the performance of the SLI sample to the norms set by typically developing children, results showed that children with SLI displayed the greatest impairments on tasks engaging both verbal short term memory and executive control (Archibald & Gathercole, 2006). This finding indicates that the impairment in this population may not stem from a deficit in verbal short term memory or executive control alone, but rather the combination (Archibald and Gathercole, 2006).

However, the interpretation that SLI stems from deficits in both short term and executive control has been challenged by others. Also administering subtests from the WMTB-C (Pickering & Gathercole, 2001) on school aged (7;2 to 9;8) children with SLI, as well as typically developing language and age matched controls, Briscoe and Rankin (2009) arrived at similar findings to those of Archibald and Gathercole (2006). That is, individuals with SLI were outperformed by age-matched controls on both measures of verbal short term memory (digit recall, word list recall, nonword list recall, CNRep) and executive control (listening recall, backwords digit recall; Briscoe & Rankin, 2009). However, group differences for the short term memory tasks persisted even after the variance from the executive control tasks was removed (Briscoe & Rankin, 2009). Contrarily, group differences on the executive control tasks were eliminated after removal of the variance from the short term memory tasks (Briscoe & Rankin, 2009). From these analyses, which differed from those conducted by Archibald and Gathercole (2006), Briscoe and Rankin (2009) argued that only verbal short term memory is impaired in the SLI population, and that this impairment, in turn, is reflected in lower performance of executive control measures.

The works of Archibald and Gathercole (2006) and Briscoe and Rankin (2009) are among the few that include both phonological short term memory and working memory within the same design in an attempt to shed light on a possible underlying factor of SLI. From these two studies it is observed that individuals with SLI display deficits in both short term and central executive tasks, albeit the relationship between the two remains unclear. However, without including a measure of language task performance, these studies only succeed at addressing whether phonological short term memory or working memory may display a greater degree of deficit. For the purposes of this study, the more interesting question is which of the two discussed cognitive factors more greatly impacts language studies. For further insight into how these factors are related, and how they interact with linguistic task performance, one can reference additional studies focusing on typically developing populations.

One study in particular, conducted on a young typically developing sample (age M = 6;3), highlights the relationship between cognitive and linguistic abilities (Engel de Abreu et al., 2011). Results showed that verbal short term memory (nonword repetition, digit recall) was related to syntactic comprehension, although this relationship was strongly mediated by vocabulary knowledge (Engel de Abreu et al., 2011). The links between working memory tasks (counting recall, backward digit recall) and syntactic comprehension, however, persisted even without contributions from vocabulary, rhyme awareness, or short-term storage (Engel de Abreu et al., 2011). Provided that individuals with SLI also reflect this pattern, we would expect to see language task performance correlating with individual differences in cognitive ability, but in particular, that of working memory.

The research focusing on the relationship between cognitive abilities and language task performance for children with SLI supports the influential role of working memory on language task performance documented by Engel de Abreu et al. (2011; Gottardo, Stanovich, & Siegel, 1996; Montgomery & Evans, 2009; Magimairaj & Montgomery, 2012). Additional research, however, suggests that phonological short term memory may influence performance as well (Montgomery & Windsor, 2007; Montgomery & Evans, 2009). After a comprehensive review of the literature, it appears that as language task demands increase, the influence of phonological short term memory is overcome by the role of working memory. Support for this conclusion stems from one study, reviewed above, in which working memory correlated with complex sentence comprehension for both impaired and unimpaired children, while phonological short term memory only correlated with simple sentence comprehension for the SLI group (Montgomery & Evans, 2009). While the PSTM deficit observed in children with SLI may influence performance, it appears that working memory ability becomes more predictive of performance for not only children with SLI, but also for typically developing children. For this reason, the current study focused on the impact of working memory, while intending to additionally control for individual differences in phonological short term memory.

1.7 Proposed Structures

Hayiou-Thomas et al. (2004) simulated an almost ideal profile of SLI performance in typically developing children via a grammaticality judgment task featuring four different structures: third person singular –S, past tense –ED, plural –S, and prepositions. As predicted by Hayiou-Thomas and colleagues (2004), performance on third person singular –S and past tense –ED decreased as a function of increasing load, while plural –S remained robust. However, preventing a clean simulation of SLI, performance on the preposition control group also decreased as a function of increasing load. One way to experimentally check Hayiou-Thomas et al.'s work, as well as further expand this body of research, would be to construct a similar grammaticality judgment task that includes morphemes that were both previously tested by Hayiou-Thomas et al. as well as novel structures. In the current study, the target problematic structures include third person singular –S, which was previously examined by Hayiou-Thomas et al., and auxiliary BE which was not. Similarly, the control structures to be used include one previously featured in Hayiou-Thomas et al.'s study, plural –S, and one that was not, progressive –ING. The individual structures will be reviewed in more detail below.

1.7.1 Target Structures: Third Person Singular –S

Third person singular –S has been shown to be a difficult structure for both typically developing children and, especially, those with SLI. As such, third person singular -S has been selected as one of the proposed morphemes to test. In a grammaticality judgment task on typically developing children and adults, it was found that even the oldest children tested (9;6-11;0) did not reach adult performance on third person singular –S structures (McDonald, 2008a). For comparison, other structures, such as plural –S and progressive –ING, achieved adult-like mastery between the ages of 8;0-9;6 and 9;6-11, respectively (McDonald, 2008a). This finding was paralleled by Hayiou-Thomas et al. (2004), who showed that third person singular –S, unlike the easier plural –S, was affected by additional processing strains, such as increases in stimuli speed or sentence length.

Children with SLI in particular have shown difficulty with third person singular –S. Evidence for this statement comes from both grammaticality judgment tasks (Montgomery & Leonard, 1998) and measures of production (Leonard et al., 2003), in which children with SLI underperform compared to their typically developing counterparts on third person singular –S. However, it should be noted that at least one production task showed no difference in third person singular –S performance between children with SLI (age *M* = 2;11) and typically developing peers, possibly due to the younger age of the subjects tested (Conti-Ramsden & Windfuhr, 2002).

A potential reason that third person singular –S may be problematic comes from the fact that this structure appears to be more demanding of individuals' working memory capacities (McDonald, 2008a). The fact that even older typically developing children (9;6 – 11;0) have not reached adult-like mastery indicates that the processing of this morpheme may not come as automatically as it would for plural –S or progressive –ING (McDonald, 2008a). Therefore, the amount of additional effort needed to process third person singular -S, or errors involving this morpheme, may begin to draw upon one's working memory capacity. Evidence for this possibility stems from a regression analysis calculated on the grammaticality judgments made concerning third person singular –S (McDonald, 2008a). Results indicated that working memory proved to be a significant predictor of third person singular –S performance, even beyond the effects of the other included predictors: age and phonological ability (McDonald, 2008a). Thus, it comes as no surprise that this structure may be especially taxing for children with SLI, who additionally have possible deficits in working memory (Montgomery, 2000a, 2000b).

1.7.2 Target Structures: Auxiliary BE

Alongside third person singular –S, auxiliary BE has been identified as one of the problematic structures for both individuals with SLI, and even those without the impairment. In a story completion task that varies target responses by complexity, it has been shown that as the sentence grows in complexity, auxiliary BE forms (IS and ARE) are omitted more frequently from production for both children with (age *M* = 5;3) and without SLI (ages *M* = 3;10 and *M* = 5;3; Grela & Leonard, 2000). This highlights the overall difficulty of this structure.

However, similar to third person singular –S, evidence suggests that individuals with SLI may be especially weak to the auxiliary BE structure. Support for this claim can be seen in naturalistic language samples, where language-matched typically-developing children correctly mark BE more than children with SLI (age *M* = 4;8; Cleave & Rice, 1997). The difficulty with this structure can also be seen through more controlled elicitation probes targeting BE, where typically-developing peers outperform children with SLI (Rice et al., 1998). It appears that while typically developing children tend to overcome the difficulty of this structure with age, individuals with SLI continue to display difficulty with auxiliary BE into adulthood, as evidence from a grammaticality judgment task over time shows (Rice et al., 2009). Because of this, it does not come as a surprise that an inherently difficult structure is even more difficult for impaired individuals.

1.7.3 Target Structures: Plural –S

Plural –S, unlike its phonologically identical counterpart, third person singular –S, has been historically seen as an easy structure. Not only is plural –S acquired earlier in development (Brown, 1973; de Villiers & de Villiers, 1973; James

& Khan, 1982), but also working memory was not found to play a role in the detection of plural -S omission errors in either typically developing children or adults (McDonald, 2008a, 2008b). As stated earlier, although performances differences appeared for third person singular -S, typically-developing children ranging in age from 8;0 to 9;6 did not differ from adults in their performance on plural -S (McDonald, 2008a). This highlights the relative ease of this structure. Further, Hayiou-Thomas et al.'s (2004) grammaticality judgment task with typically-developing children showed that this structure was resistant to the effects of increased load.

Although all the aforementioned examples concerning plural -S performance focus on typically developing individuals, studies focusing on individuals with SLI also confirm the relative ease of this structure (Rice et al., 1998). In particular, one study looked at the acquisition of plural -S in children with SLI (age *M* = 5;0) and compared their elicitation task performance to that of language-matched and age-matched peers (Oetting & Rice, 1993). Results highlighted that children with SLI correctly pluralized both frequently pluralized and infrequently pluralized regular nouns to the same degree as their language-matched peers (Oetting & Rice, 1993). For the reasons listed above, plural -S has been chosen as a control structure for the proposed study. Also, continuing to implement this structure, alongside third person singular -S, served as an experimental check on the findings of Hayiou-Thomas et al. (2004).

1.7.4 Target Structures: Progressive -ING

Like plural -S, progressive -ING has been used as a control structure by which to compare performance on problematic morphemes in a variety of tasks (Poll et al., 2010; Montgomery & Leonard, 1998; Lum & Bavin, 2007). One reason for this is that when progressive -ING is compared to other structures, it becomes evident that progressive -ING is less demanding of one's processing ability. In a word recognition task measuring on-line processing, target words were more quickly detected after the present -ING morpheme than after the third person singular -S and past tense -ED morphemes for typically developing participants and those with SLI (Montgomery & Leonard, 1998). Further evidence for the relative ease of progressive -ING comes from a grammaticality judgment task where both higher performance and faster reaction times were seen for the -ING structure in comparison to third person singular -S and past tense -ED structures for both unimpaired (age *M* = 9;5) and impaired (age *M* = 9;3) children (Lum & Bavin, 2007).

When children with SLI are compared to typically developing counterparts, performance of progressive -ING does not differ between the two populations either in a grammaticality judgment task (Poll et al., 2010) or in measures of production (Leonard et al., 2003). Also when word-detection RTs are examined for in a word recognition task featuring correct and incorrect sentences, results show that typically developing children show faster RTs for correct sentences over incorrect sentences featuring all three morpheme types (third person singular -S, past tense -ED, and progressive -ING; Montgomery & Leonard, 1998). For children with SLI, however, this trend was only seen for progressive -ING sentences, with no observable difference in RTs between incorrect and correct versions of the more difficult third person singular -S and past tense -ED sentences (Montgomery & Leonard, 1998).

The comparative ease in processing of progressive -ING is evident both (1) when comparing performance on progressive -ING to other morphemes in studies on typically developing children and (2) when observing the lack group differences in performance between individuals with SLI and their typically developing peers on this structure. For these reasons, progressive -ING emerges as a likely and logical choice for a second control structure.

1.8 Goals of the Current Study

The primary goal of this study was to explore what influential role, if any, working memory may play in SLI. Past research attempting to recreate an SLI performance profile focused on taxing the working memory ability of typically developing participants by means of manipulating stimuli (Hayiou-Thomas et al., 2004). The current study also aimed to do this, but in addition, accounted for individual differences in working memory, as measured by a size judgment task. Provided all assumptions for an ANCOVA were met, it was intended that phonological short term memory, as measured by a nonword repetition task, would be included within the analysis as a covariate. Performance is not a function of stimulus load alone. Individual differences in working memory may be just as important as external stimulus load, if not more so, in driving one's test performance. However, only a few studies examined the relationship between language task performance and individual differences for individuals with SLI (Montgomery, 2000a; Montgomery, 2000b; Montgomery & Evans, 2009). If processing difficulties are responsible for the performance seen in SLI, one should be able to determine the specific amount of load necessary to achieve a profile of SLI given one's individual differences in cognitive ability. More accurately, a simulation of SLI performance should be the result of an interaction between external stimuli load and individual differences in cognitive ability. In this study, external stimuli load was manipulated by altering length of the sentences to be judged within a grammaticality judgment task, featuring both historically problematic (third person singular -S and auxiliary BE) and unproblematic (plural -S and progressive -ING) structures. Additionally, this study measured individual differences in both phonological short term memory and working memory with the expectation that problematic structures would pose greater problems for individuals with lower working memory abilities, even after controlling for phonological short term memory.

2. PREDICTIONS

Studies have consistently shown that children with SLI have lower performance on tests of phonological short term memory and working memory when compared to typically developing children. These performance discrepancies between impaired and unimpaired children have led to a debate over whether phonological short term memory or working memory may more significantly influence the decreased language task performance observed in children with SLI. Focusing on this question, two different studies (Archibald & Gathercole; Briscoe & Rankin, 2009), discussed above in section 1.6.3 (Verbal Short Term Memory Versus Working Memory in Individuals with SLI), administered similar batteries of tests, including measures of executive function and verbal short term memory to children with SLI, and then compared their performance to typically developing children. One resulting theory is that SLI stems from deficits in both working memory and phonological short term memory (Archibald & Gathercole, 2006), while an alternative viewpoint argues that phonological short term memory alone is responsible (Briscoe & Rankin, 2009). Yet, neither of these two studies examined working memory and phonological short term memory in relation to a measure of language ability.

As stated earlier, it stands to reason that working memory, which involves both manipulation and storage, would be more implicated in a grammaticality judgment task than phonological short term memory, which solely involves storage. Evidence for this reasoning comes from studies conducted on both impaired and unimpaired children showing the correlations between working memory abilities and language task performance (Engel de Abreu et al., 2011; Gottardo et al., 1996; Montgomery & Evans, 2009; Magimairaj & Montgomery, 2012). Not denying the contributions that phonological short term memory may offer, for the current study, I hypothesized that working memory, beyond any contributions from phonological short term memory, would play a more significant role in grammaticality judgment task performance.

In this design, manipulations of sentence length was treated within-subjects, while individual differences in working memory represented the between-subjects factor. Provided the assumptions needed to perform an ANCOVA were met, variations in phonological short term memory would be included as a co-variate so as to test the influential role of working memory without the interference from the effects of phonological short term memory. My specific hypotheses included three main effects, further qualified by upper-level interactions. First, I expected to see a main effect of sentence load, such that as the length of the sentence increases, performance across all structures decreases. Second, I predicted a main effect of working memory ability, such that individuals with a higher working memory ability will outperform individuals with a lower working memory ability on the grammaticality judgment task. Third, I expected a main effect of sentence structure. Specifically, I anticipated lower performance for the historically problematic structures (third person singular -S and auxiliary BE) compared to the unproblematic structures (plural -S, progressive -ING). Further, I expected a series of 2-way interactions qualified by a 3-way interaction, such that individuals with lower working memory spans are particularly taxed by the compounding effects of high load sentences and problematic structures. Therefore, the lowest performance should be seen for individuals with low working memory spans for high load third person singular -S and auxiliary BE structures.

3. METHODS

3.1 Power Analysis

A power analysis was run with G*Power (Buchner, Erdfelder, Faul, & Lang, 2009) to determine the recommended sample size. Given the multiple within and between variables in this study, an exact test to determine recommended sample size was unavailable, so substitutions were made. A sample size was estimated for a repeated measures ANOVA using the between subjects design analysis within G*Power. This should offer a conservative estimate since the experimental variable of sentence length in the proposed study is to be conducted within subjects, which, in turn, would require comparatively fewer participants. A medium effect size of f = .25 is assumed, which would, again, be a conservative estimation given the large effect sizes observed by Hayiou-Thomas et al. (2004) for main effects of both sentence length ($\eta 2$ = .37) and structure ($\eta 2$ = .37). When also assuming an alpha of .05 and power of .80 for 2 groups (high WM ability vs. low WM ability) and 10 measures (long and short versions of 5 structures, including both filler structures as one structure), with the default correlation among repeated measures of .5, G*Power yielded a recommended total sample size of 72.

However, this power analysis is assuming that all factors are manipulated between-subjects, therefore a slightly smaller sample size for a partially within-subjects design would be expected. For comparison, Hayiou-Thomas et al. (2004) tested 120 participants. However, it should be noted that that both manipulations of speed and length were conducted between-subjects. Therefore, only a total of 30 children were used in any one condition in Hayiou-Thomas et al. (2004). Like Hayiou-Thomas et al. (2004), the sentence length in the current study is manipulated. However, this is the only stimuli manipulation, and further, will be conducted within-subjects. For this reason, we would expect a smaller requisite sample size. Thus, aiming for 30 observations per cell appears adequate to mirror Hayiou-Thomas and colleagues. Across the high and low WM span groups, this would yield a total of 60 participants to ensure adequate power.

3.2 Participants

The targeted population for the study was typically developing kindergarten children. Parental consent forms were sent out at one public school located in Louisiana’s East Baton Rouge parish, which reports a kindergarten through twelfth grade enrollment of 1360 and a teacher to student ration of 1:23 ("About LSU university", 2012). Of the 100 parental consent forms distributed, 70 were returned. Of the 70 eligible children, 9 were excluded from participation in the study due to being bilingual (1) or currently being seen by a speech language pathologist (8), as indicated on the returned parental consent forms. Unlike Hayiou-Thomas et al. (2004), the participants' hearing was not tested. However, from the consent form, all children were reported by their caregivers to have normal hearing. Thus, a total of 61 kindergarteners completed all parts of the study after giving their signed assent to participate. This sample size is in line with the sample size per cell used by Hayiou-Thomas et al. (2004).

The race of the participant sample was primarily Caucasian with one child classified as African American and one classified as Asian. Roughly equal numbers of males (N = 29) and females (N = 32) were tested, ranging in age from 5;3 to 6;8 years of age at first testing (age M = 6;1). Maternal education was also requested, and ranged from 12 (high school graduate) to 17 (graduate degree), with a mean of 16.3 (college degree).

Participants were spread across four different kindergarten classrooms. Testing always took place in the mornings between 7:30am, just prior to school officially starting, and continued until 9:00am. Children were removed from their class settings for approximately 10 minutes at a time. Testing took place in a separate room within the child's homeroom classroom. This room was either a walk-in closet or teacher's office.

3.3 Standardized Tests

To additionally ensure a typically developing sample, a series of standardized tests were given, and used to potentially exclude select participants' data from the formal statistical analyses. These tests mirror, and expand upon, the precautions taken by Hayiou-Thomas et al. (2004). The standardized tests that were administered included the *Primary Test of Nonverbal Intelligence* (*PTONI*; Ehrler & McGhee, 2008) and the *Peabody Picture Vocabulary Test-IV* (*PPVT*; Dunn & Dunn, 2007). These two tests conceptually replicate Hayiou-Thomas et al. (2004), who also screened participants based on nonverbal IQ (*Raven's Progressive Matrices*; Raven, Court, & Raven, 1986) and vocabulary knowledge (*British Picture Vocabulary Scale*; Dunn, Dunn, Whetton, & Pintilie, 1982) to ensure participants were typically developing. In addition, the syntax portion of the *Diagnostic Evaluation of Language Variation* (*DELV-NR*; Seymour, Roeper, & de Villier, 2005), and portions of the *Test of Early Grammatical Impairment* (*TEGI*; Rice & Wexler, 2001a) were given.

The standardized tests were administered after the experimental procedures, described below, were completed. Standardized testing took place across two separate sessions in a semi-randomized order due to a limited number of testing protocols available. Of the 61 participants tested, 21 failed to reach the set criteria for at least one of the standardized measures. Table 3 below provides an overview of the standardized test performance.

Table 3: Standardized Test Overview

Test	Mean	Range	Criteria for Exclusion	Did Not Meet Criteria
PTONI	*M* = 12.2 (*SD* = 18.21)	75 - 140	< 85	8
PPVT	*M* = 112.8 (*SD* = 10.5)	90 - 134	< 85	0
DELV-NR [Syntax]	*M* = 9.6 (*SD* = 2.50)	4 - 18	< 7	2
TEGI [3rd Person Singular -S]	*M* = 96.7% (*SD* = 6.51)	70% - 100%	Criterion score depending on child's age	13
TEGI [BE from BE/DO]	*M* = 94.5% (*SD* = 7.23)	64% - 100%	Criterion score depending on child's age	6

The following sections describe each of the standardized tests, as well as their respective scoring methods, in more detail.

3.3.1 Standardized Tests: Primary Test of Nonverbal Intelligence (PTONI)

The *PTONI* (Ehrler & McGhee, 2008) is a standardized measure of nonverbal intelligence, and was given to ensure that all possible participants fall within normal ranges of intelligence.

In this task, children were shown a set of pictures and were asked to identify the image that does not belong (Ehrler & McGhee, 2008).

This tool is graded based on a bell curve with the average set to 100 (Ehrler & McGhee, 2008); participants with scores lower than 85, or 1 standard deviation below the mean, failed to meet the criteria for eligibility. This value was chosen to mirror the cut-off value of used by Hayiou-Thomas et al. (2004), albeit on a different test of nonverbal intelligence.

3.3.2 Standardized Tests: Peabody Picture Vocabulary Test (PPVT)

The *PPVT* is used to measure receptive vocabulary knowledge (Dunn & Dunn, 2007). In this task, children were shown four pictures and were asked to point to the picture depicting a target word (Dunn & Dunn, 2007). Similar to the *PTONI* (Ehrler & McGhee, 2008), scores lower than 85 signified below average performance (Dunn & Dunn, 2007).

3.3.3 Standardized Tests: Diagnostic Evaluation of Language Variation (DELV-NR)

A subsection of the DELV-NR focusing on syntax was also given to screen eligible participants. This portion of the DELV-NR examines performance on wh-questions (e.g., "This father and this baby were having lunch together. Who ate what?"), passives (e.g., [point to] "The elephant was pushed"), and article usage (e.g., "Think of a police officer. What does he have?" [A gun, badge, etc.]; Seymour et al., 2004). Since the proposed study featured a grammaticality judgment task, which focuses on being able to identify errors in syntax, screening children with a language measure specifically focusing on syntax was appropriate. Although there are other tests that measure syntactic ability, the *DELV-NR* emerged as a strong option because it does not test features that overlap with those targeted in the current study. The inclusion of a syntax-based language measure serves to potentially screen out children with possible language weaknesses or impairments, which were undocumented on the consent form. Scores below 7 indicate below average syntactic performance (Seymour et al., 2004).

3.3.4 Standardized Tests: Test of Early Grammatical Impairment (TEGI)

The *TEGI* (Rice & Wexler, 2001a) is composed of multiple parts, two of which, focusing on verb morphology, were given. The subsections to be administered involve eliciting responses featuring morphemes that were also experimentally tested. The elicited morphemes include (1) the third person singular -S and (2) structures BE and DO (Rice & Wexler, 2001a). The purpose of administering this test was to ensure that the children have acquired the target morphemes to be later examined in their grammar.

In the third person singular -S task, participants are asked to describe what a target person, such as a police officer, does (e.g., "a teacher teaches"; Rice & Wexler, 2001a). In the BE/DO task, toys are used in addition to a story script to elicit questions (e.g., "are the moon guys resting?") or making statements (e.g., "the bug is tired") targeting either a BE (auxiliary or copula) or DO structure (Rice & Wexler, 2001a). For the BE/DO portion of the TEGI, the manual is unclear as to whether or not items may be repeated or whether additional prompting may be used (Rice & Wexler, 2001b). To ensure the maximum possible scoreable responses for each target item, the experimenter reprompted until a scoreable response was obtained. Scoreable responses could include appropriate marking of the desired morpheme (e.g., "is the bug hungry?", morpheme omissions (e.g., "the bug hungry?"), or incorrect forms (e.g., "are the bug hungry?") being used (Rice & Wexler, 2001b). For example, if the target was "is the bug hungry?" targeting the form IS, and the child, instead, asked, "are the moon guys hungry?", the experimenter would reprompt "ask the puppet if the *bug's* hungry?" from the TEGI (Rice & Wexler, 2001a) script, or "ask about the bug". Without such reprompting, a child could theoretically only give unscoreable utterances (e.g., "the bug ate"), or only utterances involving singular or plural forms, not affording an complete picture of whether that child has acquired BE or DO in their multiple forms. In the few cases that the child did not give a scoreable response specifically tailored to the target structure, the last utterance was scored as is.

Unlike the *PTONI* and *PPVT*, which compute standard scores, separate criterion scores are used to determine whether a child passes the subsections of the *TEGI* (Rice & Wexler, 2001b). The criterion scores are based upon the percent of third person singular -S or BE/DO marking for the child's age at testing (Rice & Wexler, 2001b). Percentage of marking was calculated as the number of times a child used third person singular -S or BE/DO in contexts which required the third person singular -S or BE/DO marker (Rice & Wexler, 2001b). Since the current study does not include any instances of DO, only performance on BE was considered.

The abovementioned subcomponents of the TEGI (Rice & Wexler, 2001a) focus on syntax. However, the purpose of administering this test was not to gauge a child's syntactic ability, as was the purpose of the DELV-NR, but rather to measure a child's mastery of certain grammatical morphemes. Without concrete evidence that a young child has acquired a particular grammatical structure, the driving force behind possible poor performance could be unclear. For instance, poor performance could be driven by the inherent difficulty of the structure, or, conversely, could be indicative of a structure not yet acquired.

It should be noted that out of all the morphemes to be experimentally examined, only two – third person singular -S and auxiliary BE – were formally tested to ensure structure mastery. The standardized test (*TEGI*; Rice & Wexler, 2001a) used to determine mastery of these morphemes does not offer sections focusing on plural -S or progressive -ING. This potentially raises the question of whether participants have also mastered plural -S and progressive -ING. Past documentation indicates that these structures, mastered by 3;1 if not sooner, are among the earliest acquired in typical language development (Brown, 1973), and pose little difficulty for both impaired and unimpaired children (Rice et al., 1998; Lum & Bavin, 2007). Given the age of the typically developing sample of children in the current study (age *M* = 6;1), it is highly probable that plural -S and progressive -ING have already been sufficiently mastered.

3.4 Experimental Tasks

Experimental testing took place across two days, and the order of these days was counterbalanced across participants. Testing for all eligible children included a nonword repetition task and a size judgment task to measure individual differences in phonological short term memory and working memory, respectively. Also, a grammaticality judgment task (administered over two days) was given, which focused on four grammatical markers: third person singular -S, auxiliary BE, plural -S, and progressive -ING. On one day, a participant received the short sentence grammaticality judgment condition followed by the size judgment task, while on a separate day, the child would receive the nonword repetition task followed by the long sentence grammaticality judgment condition.

All experimental audio stimuli were recorded in a sound proof booth using a Marantz PMD670 digital audiorecorder, and were subsequently administered using PowerPoint on a Dell Inspiron N5110 PC laptop computer. All experimental stimuli were normalized after being recorded within Audacity 1.2.5 (Mazzoni et al., 2006) to ensure no peak clipping had occurred.

During the grammaticality judgment task, both the experimenter and participants wore Panasonic RP-HTX7-K circumaural headphones connected to the laptop via a y-cable audio splitter. The nonword repetition task and size judgment task, however, were not presented via headphones so that the child could more clearly monitor his or her verbal responses without the noise reduction effect the headphones contribute. Instead, these tasks were presented over the laptop's internal loudspeakers at a comfortable listening volume.

3.4.1 Nonword Repetition

The nonword repetition task was given to assess the children's phonological short term memory. In the nonword repetition task, individuals were asked to repeat nonwords presented auditorily via a PowerPoint presentation to the best of their ability. The nonwords were the same used by Dollaghan and Campbell (1998), rerecorded by a native-English-speaking, African American female hailing from the southern United States region.

The task began with the experimenter reading the instructions to the participant from the PowerPoint experiment. Next, four practice items, taken from the nonword repetition portion of the *Diagnostic Evaluation of Language Variation Screening Test* (*DELV-ST*; Seymour, Roeper, & de Villiers, 2003) were spoken aloud by the experimenter. After completing the practice items, the experimenter pressed the laptop's spacebar to begin the formal task.

The task included a total of 16 nonwords, with four words presented per length. The task was not adaptive in that each child received all nonwords, however the words were presented in order of increasing length. Nonwords started at one syllable and extended up to four syllables in length, always with a CVC structure, and no consonant occupied a syllable position with a phonotactic probability of greater than 25%. For each trial, a blank PowerPoint slide would appear, accompanied by a novel nonword to be recalled. The child would repeat the perceived word aloud. All verbal responses were audiotaped using a portable, digital Edirol R-09HR audio recorder for offline scoring. After the child responded, the experimenter would press the spacebar, and the next word would immediately be presented aloud.

Scoring of the NWR task was carried out in the same way as Dollaghan and Campbell (1998). Any omissions or phoneme substitutions were marked as errors, but any distortions of the target phoneme or phoneme additions were not counted against the participant. For example, if an individual repeated the target word “t/ei/v/a/k” as “t/ai/v/a/k,” this would be marked as a distortion and not counted as an error. However, if an individual repeated “t/o/v/a/k,” the obvious phoneme substitution “o” for “ei” would be marked as an error.

3.4.2 Size Judgment Task

The size judgment task provided a means to assess an individual’s working memory ability. In this task, participants heard lists of multiple one and two syllable words, recorded by a different native-English-speaking, African American female from the southern United States region. All stimuli were presented over the laptop’s internal loudspeakers. Participants were then required to list these words from smallest physical object to largest physical object. The length of these lists gradually increased from two words to six words, with three sets per list length. The lists used are available in Table 4.

Table 4: Size Judgment Task Word Lists

Level 2	
Set 1:	Stove___Mouse___
Set 2:	Key___Squirrel____
Set 3:	Cat___Needle___
Level 3	
Set 1:	Rabbit___Bike___Tooth___
Set 2:	Coat___Goldfish___Book___
Set 3:	Seed___Guitar___Kitten___
Level 4	
Set 1:	Hat___Truck___Beetle___Fox___
Set 2:	Table___Lizard___Duck___Car___
Set 3:	Island___Lemon___Ant___Bear___
Level 5	
Set 1:	Bee___Whale___Parrot___Door___Apple___
Set 2:	Cow___Goat___Nail___Mountain___Pan___
Set 3:	Rooster___Lion___Chair___Shoe___Worm___
Level 6	
Set 1:	Pony___Ring___Wolf___Ocean___Chicken___House___
Set 2:	Planet___Rat___Fly___Dog___Bed___Airplane___
Set 3:	Giraffe___Purse___Cup___Bridge___Snail___Sheep___

The task began with the experimenter reading instructions to the participant from the PowerPoint experiment. Next, each child completed three practice items. The practice lists were two words in length, and were administered verbally by the experimenter to ensure the participant comprehended the task. Afterwards, the formal task was executed using PowerPoint. Similar to the nonword repetition task, each stimuli set was presented over the computer's loudspeakers immediately after the experimenter pressed the spacebar. The ISI for each word list was 500msec. At the end of a list's presentation, a circle would appear in the upper right corner of the monitor indicating the list was complete and the participant was free to begin recalling the items aloud in order of smallest to greatest. Unlike the nonword repetition task, this task was not audiotaped. Rather, responses were recorded online by the experimenter on an answer sheet. In case the child repeated words more than once, or falsely recalled a non-target, that word was documented on the paper along with the serial number in which it was said.

Although working memory ability has been measured using size judgment task performance in the past (Montgomery, 2000a, 2000b), this particular stimuli set has not been used in prior research. While the current study's working memory task and that of Montgomery (2000a, 2000b) are similar, they differ slightly. In term of stimuli, both the current task and the task used by Montgomery (2000a, 2000b) include words assumed to be familiar to a child. One difference to note is that Montgomery (2000a, 2000b) used only monosyllabic words, which are sometimes plural (i.e., "socks," "shoes"). When ranking items by size, presenting a plural object may contribute to confusion. For contrast, the current study includes only singular words, but words may be either one or two syllables in length. In terms of methodology, Montgomery's (2000b) lists ranged from three to seven words were created from a word bank of 25 words and presented randomly for each participant. This differed from the current study in which the same lists of non-repeated words were presented in an incrementally increasing fashion for all participants, beginning with the two word lists and ending with the six word lists. While Montgomery's (2000b) method may possibly ward against elevated performance as the child cannot anticipate the number of items he or she will need to recall, the comparatively small word bank and semi-randomization of the lists contribute to two potential issues. First, although restrictions are set so that no word repeats within a list, the small pool of eligible words guarantees that the same words will be used across the task, potentially increasing the chances of intrusion errors. Also, items closely related in size (i.e., "skates," "boots") may be generated in the same list, unintentionally increasing the demands of the task by introducing ambiguity into the ordering.

3.4.3 Grammaticality Judgment Task

Children were administered a grammaticality judgment task via PowerPoint which focused on morphemes which children with SLI historically struggle with (third person verbal –S, auxiliary BE) and also display little difficulty with (plural –S, progressive –ING). The grammaticality judgment task was administered across two separate sessions to guard against fatigue, with order of the sessions counterbalanced across participants. To conceptually parallel the design of Hayiou-Thomas et al. (2004), in which participants only received one of the four possible load manipulations, only sentences of a particular length (short vs. long) were presented for any one session within the current study. Using a y-cable headphone splitter, both the children and experimenter listened to the sentenced via circumaural headphones. The stimuli were presented at a comfortable listening volume. All sentence stimuli were recorded by a native-English-speaking, Caucasian female from Louisiana. In addition to normalizing the audio clips, 250msec of silence was added before and after each sentence.

The grammaticality judgment task started with the experimenter reading the instructions aloud from the PowerPoint experiment. After the instructions were given, four practice items were presented via the PowerPoint. The PowerPoint would automatically play the practice sentence aloud to be judged. The child was then asked to say if the sentence sounded "good" or "not so good". Additionally, the participants were asked to elaborate why an item may have sounded not so good. When a participant did not correctly identify the ungrammatical items, the experimenter would draw the participant's attention to the violation, and ask the participant how he or she would say the sentence or make it sound better, which often resulted in the child noticing the error and correcting the sentence. This guided learning was intended to highlight to the participant that the focus on whether the sentence sounded good or not so good was not based on semantics, but rather the syntactic content of the sentence. After the practice items, the formal task began. The same four practice items were given, albeit in a different order, for both the long and short sentence conditions. The practice items represented both grammatical and ungrammatical versions of two sentences focusing on structures, past tense –ED and article A, not targeted in the formal task. The ungrammatical sentence versions featured either a past tense –ED or article A omission.

For the formal task, a PowerPoint slide would appear playing the sentence stimuli. Then the participant responded aloud with "good" or "not so good". After the participant responded, the experimenter would record the participant's answer online on an answer sheet before pressing the laptop's spacebar, which would immediately present the next sentence stimuli.

The formal grammaticality judgment task consisted of 96 sentences, with 64 of those sentences focusing on one of the four target morphemes. The 32 remaining sentences were filler sentences, which either featured a subject-verb agreement error using BE ("am" for "is" or "is" for "am") or its correct sentence counterpart. Most past research conducted on agreement errors indicate they are not exceptionally problematic for children with SLI (Leonard et al., 2003; Rice et al., 1999; Poll et al., 2010). The positive aspect of using PowerPoint is that the experimenter could easily go back to an item in the instance that an external distraction occurred that prevented the child from hearing a sentence. Although the frequency of having to repeat an item was rare (.003%), this feature was particularly important as children were tested in school environments, which do not afford the same level of environmental control offered in a laboratory setting. The negative aspect of using PowerPoint is that stimuli were not presented randomly without replacement for each participant. Thus, all participants received the sentences to be judged in the same order, which potentially introduced order effects. To address this issue during stimuli creation, sentences were assigned to one of four blocks per session, with a small break occurring between the blocks to allow the child to rest. Additionally, for each block, as the task progressed, a visual bar on the screen would fill from red to green indicating to the child that a break was coming, which seemed to help curb fatigue. The sentence presentation order within each block was determined via a random number generator with a few limitations. First, each block consisted of half grammatical items, with only the grammatical or ungrammatical version of each sentence able to appear within any given block. Also, only one ungrammatical sentence from each type appeared within each block. While we were unable to present a randomized task for each participant, the presentation restrictions taken during stimuli creation afforded some experimental control to ensure that participants were not exposed to the same type of error in short succession.

The sentences, all present tense, featured an equal number of low and high processing load versions, and each incorrect sentence was balanced with a correct counterpart. To manipulate load, all sentences were systematically lengthened. Starting with a subject-verb-direct object base sentence, the low load versions contained one additional word and the high load sentences contained 6 additional words. Specifically, low load sentences featured a base sentence with an additional 3-syllable adverb at the end (e.g., "He is playing many games happily"; "He pays many bills lazily"), and high load sentences featured both the adverb from the low load sentence, followed by an additional prepositional phrase, as well as an additional word, "Today," that was added to the beginning of the sentence(e.g. "Today, he is playing many games happily in the old gym"; "Today, he pays many bills lazily at the new bank").

Each of the four grammatical structures was manipulated within one of two specific base sentence structures. Third person singular –S (e.g., "Today, he pay(s) many bills lazily") and plural –S (e.g., "Today, she sprays may plant(s) thoroughly") were manipulated within sentences featuring third person singular –S as the main verb. Auxiliary BE (e.g., "Today, he (is) buying many shoes eagerly in the large store") and progressive –ING (e.g., "She is say(ing) many things nervously") were manipulated within sentences featuring an auxiliary BE structure as the main verb.

The four specific grammatical structures to be examined were each featured in 16 sentences, with 8 sentences being grammatical and 8 sentences being ungrammatical due to the target morpheme omission. All grammatical and ungrammatical sentences were further divided into short, low load sentences and long, high load counterparts. Thus, each structure had 4 possible combinations consisting of four sentences each: grammatical-short, ungrammatical-long, ungrammatical-short, and ungrammatical-long. For a list of all of the base sentences proposed, as well as an example of how a sentence in lengthened, refer to Tables 5 and 6 respectively.

Table 5: Base Sentences

	Grammatical	**Ungrammatical**
Third Person Singular -S	1. She flies many planes fearlessly 2. She grows many plants secretly 3. He boos many teams angrily 4. He pays many bills happily	1. She fly many planes fearlessly 2. She grow many plants secretly 3. He boo many teams angrily 4. He pay many bills happily
Auxiliary BE	1. He is buying many shoes eagerly 2. She is throwing many balls playfully 3. He is chewing many chips noisily 4. She is laying many eggs cautiously	1. He buying many shoes eagerly 2. She throwing many balls playfully 3. He chewing many chips noisily 4. She laying many eggs cautiously
Plural -S	1. He ties many bows correctly 2. He rows many boats lazily 3. She screws many bolts forcefully 4. She sprays many plants thoroughly	1. He ties many bow correctly 2. He rows many boat lazily 3. She screws many bolt forcefully 4. She sprays many plant thoroughly
Progressive -ING	1. She is trying many foods hungrily 2. He is sewing many shirts quietly 3. He is stewing many pears hastily 4. She is saying many things nervously	1. She is try many foods hungrily 2. He is sew many shirts quietly 3. He is stew many pears hastily 4. She is say many things nervously

Table 6: Sentence Lengthening

	Base Sentence (Short)	**Lengthened Version (Long)**
Example of Third Person Singular -S	She grow(s) many plants secretly	Today, she grow(s) many plants secretly in the green house
Example of Auxiliary BE	He (is) buying many shoes eagerly	Today, he (is) buying many shoes eagerly in the large store
Example of Plural -S	He ties many bow(s) correctly	Today, he ties many bow(s) correctly with the pink string
Example of Progressive -ING	She is try(ing) many foods hungrily	Today, she is try(ing) many foods hungrily in the meat aisle

In addition to the target structures, there were 8 filler sentences featuring auxiliary agreement errors where "is" was replaced with "am" (e.g., "He am weighing many grapes easily"), and 8 filler sentences featuring auxiliary agreement errors where "am" was replaced with "is" (e.g., "I is playing many games skillfully"). Table 7 lists the filler base sentences used.

Table 7: Agreement Error Filler Sentences

	Grammatical	**Ungrammatical**
Agreement AM	1. I am weighing many grapes easily. 2. I am towing many trucks rapidly 3. I am crying many tears openly 4. I am gluing many stars cheerfully	1. He am weighing many grapes easily. 2. He am towing many trucks rapidly 3. He am crying many tears openly 4. He am gluing many stars cheerfully

Table 7 (Continued): Agreement Error Filler Sentences

	Grammatical	**Ungrammatical**
Agreement IS	1. He is mowing many lawns carelessly 2. He is playing many games skillfully 3. She is frying many eggs patiently 4. She is viewing many films carefully	1. I is mowing many lawns carelessly 2. I is playing many games skillfully 3. I is frying many eggs patiently 4. I is viewing many films carefully

These filler sentences are speculated to be relative easy for children with SLI for two reasons. First, even though auxiliary BE is considered a difficult structure for children with SLI, differences in sensitivity between the forms of BE may exist, with AM being less sensitive. Evidence for this assumption stems from one study focusing on eliciting first person auxiliary BE forms (Polite & Leonard, 2007). Results showed that although children with SLI (age *M* = 5;3) produced AM less frequently than their typically developing peers, one third of the children with SLI marked AM for every trial, indicating that, while AM may still be somewhat problematic, it may not be as problematic as other forms of BE (Polite & Leonard, 2007). Secondly, the filler sentences to be used feature an agreement error instead of a morpheme omission. According to the literature on SLI, it is known that errors aside from omission are infrequent in both typically developing and SLI populations (Leonard et al., 2003; Rowland, Pine, Lieven, & Theakston, 2005).

Performance on these filler sentences alone would be a particularly interesting contribution to the current design. For a successful simulation of SLI to occur, individuals, even under the hardest of loads, would still be able to accurately identify an agreement error and label that sentence as ungrammatical. If, however, individuals accept agreement errors as frequently as omission errors under load, this would fail to support the working memory theory of SLI, and simply reflect the detrimental nature of taxing working memory.

4. RESULTS

4.1 Nonword Repetition Task Performance

Nonword repetition was scored based upon percent accuracy, where the number of correctly produced phonemes is divided by the total number of phonemes for all the words. In the case that a child did not respond to an item, the number of phonemes for that nonword were not included within the total scoreable number. Percent accuracy on the nonword repetition task ranged from 64% to 98% (M = 82.7%, SD = 7.7). For comparison, on this same task, Dollaghan and Campbell (1998) found nonimpaired children (age range: 6;0 - 9;9) to perform at 84% accuracy.

4.2 Size Judgment Task Performance

Similar to the nonword repetition task, the size judgment task was scored based upon percent accuracy of links recalled, where the number of links recalled is divided by the maximum possible number of links across all attempted lists. Links are defined as successfully recalling a smaller word followed by a larger word, both of which must appear on the to-be-remembered list. For example, if a child recalls the words "house, dog, airplane, planet" from the 6 item list "planet, rat, fly, dog, bed, airplane," he would be given a percentage score of 40%, since out of a maximum of 5 possible links, a child listed 2 (dog < airplane = 1, airplane < planet =2).

This scoring method appears to ward against the effects of possible free or serial recall of items, as well as accounting for instances where a participant failed to respond to a given list, which would otherwise artificially deflate a participant's score. Percent accuracy on the size judgment task ranged from 13% to 78% (M = 42.8%, SD = 13.7).

4.3 Grammaticality Judgment Task Performance

4.3.1 Grammaticality Judgment Task Performance with All Participants

Initially, an A' statistic was calculated to determine performance on the target structures (Stanislaw & Todorov, 1999). Ideally, an A' value should range between .5, indicating chance performance, and 1, indicating ceiling performance. A' values for the current study, however, included values below .5 and missing values indicated by division by zero, suggesting that some participants displayed below chance performance, or that the false alarm rate exceeded the hit rate. Therefore, an alternative measure of judging performance—percent accuracy on ungrammatical items—was used. While it is unclear why a participant chooses to label a grammatical item as incorrect, one can more reasonably speculate that the reason for labeling an ungrammatical item as grammatical is due to the fact that the presented syntactic violation was not perceived as problematic.

First, a bivariate correlation was conducted including all standardized test measures, cognitive measures, and grammaticality judgment items; the results can be seen in Table 8, where S stands for sentences that were short in length and L stands for long sentences.

Concerning the relationship between phonological short term memory, as measured by the nonword repetition task, with the other items, two observations were made. First, nonword repetition failed to correlate with any of the syntactic measures, either within the standardized tests or the grammaticality judgment task. This supports our assumptions that phonological short term memory would be less implicated in grammaticality judgment task performance than working memory. The second observation revolved around the direction of the nonsignificant correlations between nonword repetition task performance and the grammaticality judgment items. Although not significant, the correlations between nonword repetition were positive for some target morphemes and negative for others. These conflicting negative and positive correlations, in turn, violated the homogeneity of regression slopes assumption that would be necessary for including nonword repetition task performance as a covariate within an ANCOVA design, as intended.

In contrast, WM, as measured by the size judgment task, significantly correlated with measures of syntactic performance from both the standardized tests (*TEGI* BE) and grammaticality judgment items (short progressive -ING, long auxiliary BE, and long plural -S). While it was unexpected that WM would significantly correlate with performance on the control structures, the presence of significant correlations with items in the grammaticality judgment task in general supported our assumptions that working memory would significantly impact syntactic performance.

Table 8: Correlations between Standardized Test Measures, Cognitive Measures, and Target Structures for All Participants

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. PTONI														
2. PPVT	.42**													
3. DELV-NR	-.06	.18												
4. TEGI 3rd -S	.08	.15	.18											
5. TEGI BE	.14	.10	-.01	.07										
6. Size Judge	.33**	.26*	.19	.23	.33*									
7. Nonword Rep	.12	.32*	.03	.05	.14	.38**								
8. S 3rd -S	.23	.26*	-.08	-.13	.19	.21	.10							
9. S Aux BE	.20	.24	.12	-.00	.03	.17	-.06	.54**						
10. S Plural S	.26*	.16	.16	-.03	.11	.20	-.20	.41**	.41**					
11. S ING	.30*	.32*	-.11	-.06	.13	.34**	.13	.53**	.47**	.39**				
12. L 3rd -S	.10	.16	.16	.13	.21	.20	.15	.36**	.41**	.33**	.37**			
13. L Aux BE	.19	.28*	.33**	-.01	.24	.28*	.23	.24	.24	.14	.31*	.53**		
14. L Plural S	.28*	.37**	.32*	.01	.15	.32*	.19	.46**	.37**	.45**	.46**	.47**	.47**	
15. L ING	.05	.32*	.02	.09	.11	.15	.18	.31*	.36**	.26*	.43**	.57**	.32*	.52**

** correlation at the p < .01 level
* correlation at the p < .05 level

In an effort to clarify the more specific role of WM on grammaticality judgment task performance, further analyses were performed. Originally, a 2 (short, long) x 4 (third person singular -S, auxiliary BE, plural -S, progressive -ING) x 2 (low WM span, high WM span) ANCOVA was intended, including PSTM as a covariate. Due to the violation of one of the assumptions necessary to conduct an ANCOVA using phonological short term memory as a covariate, this element was eliminated from the design.

A 2 (short, long) x 4 (third person singular -S, auxiliary BE, plural -S, progressive -ING)_x 2 (low WM span, high WM span) ANOVA was performed for all participants, including working memory performance as a between subjects variable. Participants were divided into low and high working memory groups via median split of their size judgment task performance. Based on the hypotheses, we expected to see individuals with lower working memory spans displaying poorer performance on historically problematic structures (third person singular -S and auxiliary BE), especially in longer sentence contexts. First, main effects surfaced for both length, $F(1,59) = 17.18$, $p < .01$ (partial $\eta^2 = .23$), and working memory, $F(1,59) = 4.50$, $p < .05$ (partial $\eta^2 = .07$). These main effects were qualified by a significant two-way interaction between length and structure, $F(3,177) = 5.08$, $p < .01$ (partial $\eta^2 = .08$), and ultimately a significant three-way interaction between length, structure, and working memory, $F(3,177) = 3.42$, $p < .05$ (partial $\eta^2 = .06$). These results can be seen below in Figure 1.

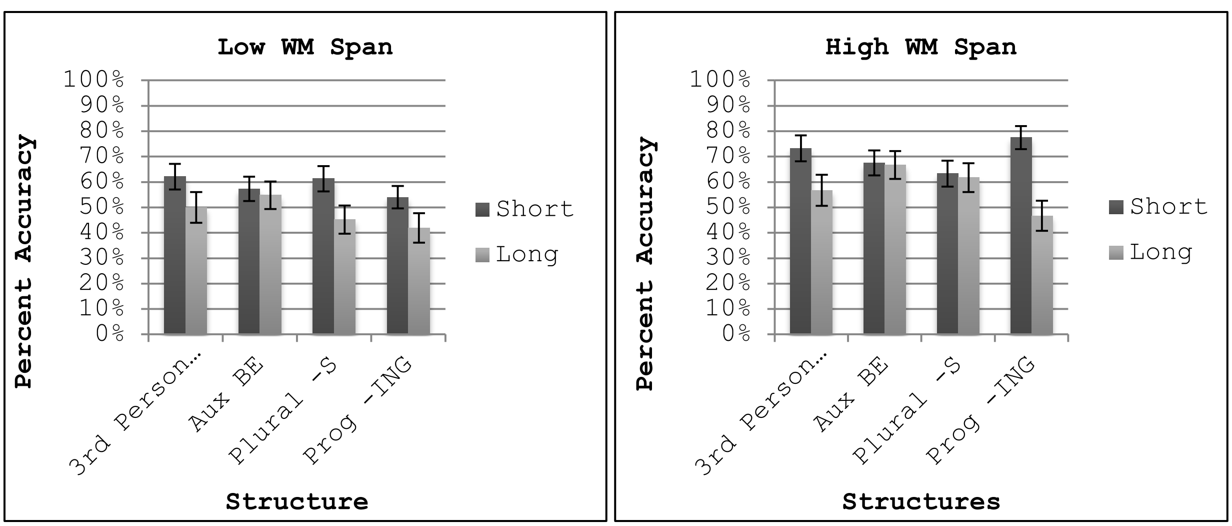

Figure 1: Target Structure Performance Split by Length for Individuals with Low and High Working Memory Spans (All Participants)

Follow up analyses for the significant three-way interaction between length, structure, and working memory showed that within the low WM span group, only a main effect of length emerged, $F(1,30) = 8.23$, $p < .001$ (partial $\eta^2 = .22$), where shorter sentences ($M = 58.7\%$) displayed higher performance than longer sentences ($M = 48.0\%$). Within the high WM span group, the main effect of length was also seen, $F(1,29) = 8.92$, $p < .01$ (partial $\eta^2 = .24$), with shorter sentences ($M = 70.4\%$) outperforming longer sentences ($M = 57.9\%$), but in addition, there was a significant interaction between length and structure, $F(3,87) = 8.86$, $p < .001$ (partial $\eta^2 = .23$). For individuals with higher WM spans, structure differences emerged in both the short, $F(3,87) = 3.31$, $p < .05$ (partial $\eta^2 = .10$), and long, $F(3,87) = 4.35$, $p < .01$ (partial $\eta^2 = .13$) sentence condition. In the short sentence condition, progressive -ING ($M = 77.5\%$) displayed higher performance than both auxiliary BE ($M = 67.5\%$) and plural -S ($M = 63.3\%$). In the long sentence condition, a reverse trend was seen such that performance on progressive -ING (M = 46.7%) was significantly lower than both auxiliary BE (M = 66.7%) and plural -S (M = 61.7%). When investigating the effect of length within each structure for individuals with high WM spans, it was observed that length most negatively impacted third person singular -S, $F(1,29) = 7.63$, $p < .05$ (partial $\eta^2 = .21$), and progressive -ING, $F(1,29) = 23.07$, $p < .001$ (partial $\eta^2 = .44$). Alternatively, when focusing on performance differences between individuals with low and high WM spans, it was noted that, contrary to our predictions, low and high WM span individuals differed specifically in their performance on short progressive -ING, $t(59) = -3.73$, $p < .001$, and long plural -S, $t(59) = -2.06$, $p < .05$.

These results indicate that our hypotheses were not completely supported. As predicted, individuals with higher working memory spans outperformed individuals with lower working memory spans, and length appeared to detrimentally impact performance regardless of personal differences in cognitive abilities. Contradicting our hypotheses, historically difficult structures did not systematically show lower performance than the selected control structures, which were predicted to remain robust. Although a three-way interaction emerged as predicted, follow-up analyses indicated that the specific patterns of performance were not reflective of specific language impairment. While length appeared to impact performance in the lower working memory span group, these main effects were qualified by a two-way interaction between the two in the high working memory span group. Why this interaction was observed for individuals with higher, and not lower, working memory spans could be due to the exceptionally low performance of the low working memory span group.

4.3.2 Grammaticality Judgment Task Performance with Selected Participants

The above analyses included all participants, regardless of their understanding of how to perform a grammaticality judgment task or their performance on measures from the standardized tests taken to ensure that they have age-appropriate performance in nonverbal intelligence, vocabulary knowledge, syntax, and third person singular –S and BE mastery. To omit all participants which failed to perform at an average or above level on any one of the given standardized tests would exclude 21 participants. However, just because a participant showed acceptable performance on the battery of standardized tests does not necessarily mean he or she was capable of performing a grammaticality judgment task successfully. To exclude participants who additionally failed to perform above chance on a composite measure of all the short (low load) grammaticality judgment sentences would lead to a total exclusion of 27 participants. This would severely reduce the statistical power for the subsequent analyses in which individual differences in phonological short term memory and working memory are explored. Therefore, the following measures were taken to exclude participants from the analysis while maximizing on the amount of data to analyze. First, participants (N = 10) were excluded if they failed to perform at age-appropriate measures for the *PTONI* (Ehrler & McGhee, 2008), *PPVT* (Dunn & Dunn, 2007), and *DELV-NR* (Seymour et al., 2004). This mirrors and expands upon the standards set by Hayiou-Thomas et al. (2004). By doing so, we can assume that our sample is "typically developing". Secondly, an additional 9 participants were removed from analysis if, on a composite measure of grammatical and ungrammatical short sentence performance across the entire experiment, they performed at chance (50%) or below. This decision was based on the assumption that, if participants are unable to correctly reject syntactic errors and accept grammatical sentences above chance within the baseline condition, they are unable to successfully perform a grammaticality judgment task. Including such data would add unnecessary noise. Between these two methods of participant selection, 4 participants, who failed to perform at the age-appropriate criterion score for both the third person singular –S and BE subsections of the *TEGI* (Rice & Wexler, 2001a) were also excluded.

The above bivariate correlation was re-run on the remaining 42 participants, seen in Table 9. Results continued to show

Table 9: Correlations between Standardized Test Measures, Cognitive Measures, and Target Structures for Selected Participants

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. PTONI														
2. PPVT	.37*													
3. DELV-NR	-.09	.10												
4. TEGI 3rd -S	.07	-.03	.19											
5. TEGI BE	.10	.18	-.01	-.26										
6. Size Judge	.42**	.33*	.05	.14	.16									
7. Nonword Rep	.34*	.50**	.06	.10	.19	.44**								
8. S 3rd -S	.25	.28	-.16	-.15	.11	.25	.15							
9. S Aux BE	.25	.26	-.00	-.07	.03	.19	.06	.54**						
10. S Plural S	.24	-.06	.02	-.16	-.11	.07	-.27	.21	.11					
11. S ING	.39*	.37*	-.25	.00	.14	.35*	.11	.51**	.59**	.24				
12. L 3rd -S	.25	.37*	.06	.01	.03	.04	.24	.52**	.42**	.15	.49**			
13. L Aux BE	.14	.32*	.30	-.11	.04	.03	.20	.30	.40**	-.05	.32*	.56**		
14. L Plural S	.34*	.47**	.27	.06	-.10	.15	.21	.35*	.42**	.32*	.40**	.49**	.37*	
15. L ING	.20	.49**	-.12	-.13	-.03	.09	.20	.40**	.33*	.07	.53**	.59**	.38*	.61**

** correlation at the p < .01 level
* correlation at the p < .05 level

that nonword repetition task performance failed to correlate with measures of syntax, either within the standardized or experimental tasks. The removal of selected participants, however, reduced the number of significant correlations between size judgment performance and syntactic measures, such that the only significant correlation that persisted was between WM and progressive –ING in the short condition.

To parallel the initial analyses, the 2 x 4 x 2 ANOVA was repeated on the reduced sample. Again, there was a main effect of length, $F(1,40) = 17.75$, $p < .001$ (partial $\eta^2 = .31$), qualified by a two-way interaction with structure, $F(3,120) = 5.35$, $p < .005$ (partial $\eta^2 = .12$), seen in Figure 2 below. When restricting the sample size, however, the main effect of working memory was no longer significant, and the three-way interaction observed between working memory, length, and structure reduced to only marginal significance, $F(3,120) = 2.65$, $p = .052$. This is possibly due to a decrease in power, attributed to a reduction in sample size. Failing to support the predicted trends of performance, the nature of this marginally significant three-way interaction was similar to the one previously observed when including all participants.

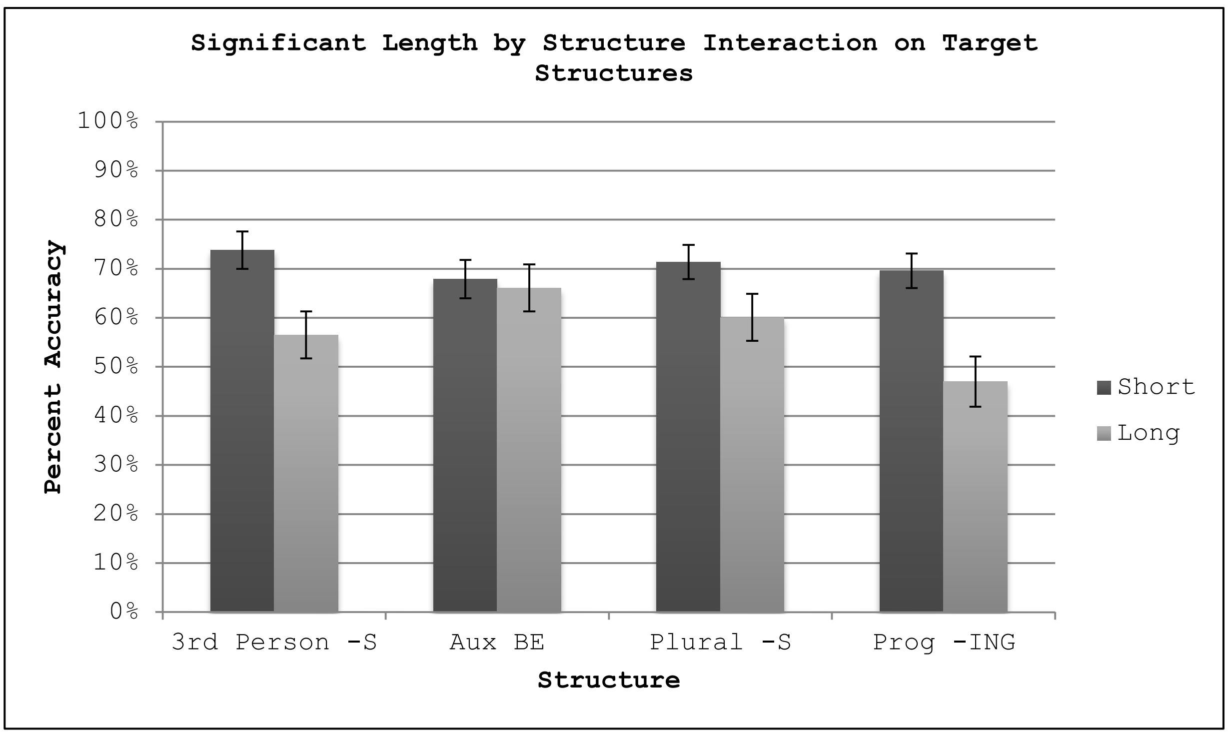

Figure 2: Target Structure Performance Split by Length (Selected Participants)

Exploration of the significant two-way interaction between length and structure analysis revealed that structure differences did not surface in the short sentence condition when collapsed across both WM span groups. In the long sentence condition, however, there was a significant main effect of structure, $F(3,123) = 5.50$, $p < .01$ (partial $\eta^2 = .12$), where progressive -ING (M = 47.0%) displayed lower performance than all structures (third person singular -S: M = 56.5%; auxiliary BE: M = 66.1%; plural -S: M = 60.1%), and third person singular -S displayed lower performance than auxiliary BE. Concerning the effects of load within each structure, three of the four structures examined showed a detrimental effect to increases in sentence length: Third person singular -S, $F(1,41) = 16.11$, $p < .01$ (partial $\eta^2 = .28$), plural -S, $F(1,41) = 5.31$, $p < .05$ (partial $\eta^2 = .12$), and progressive -ING, $F(1,41) = 26.29$, $p < .001$ (partial $\eta^2 = .39$). Using a restricted sample size, the results fail to more closely approximate SLI. First, the lowest performance in the long sentence condition is for the control structure progressive -ING. Second, results showed that while only one historically difficult structure displayed sensitivity to increasing processing load, both control structures were negatively affected by sentence length.

4.3.3 Target Structure Overview

For a successful simulation to occur, children with a lower working memory span would have displayed lower performance for third person singular -S and auxiliary BE structures, particularly in the long sentence condition. When exploring the effects of length within each structure, the global hypotheses were unsupported, even after selectively removing participants in order to achieve a cleaner sample. In the restricted sample, both plural -S and progressive -ING unexpectedly failed to be robust against the load manipulation. This indicates that both control structures did not behave as anticipated. Likewise, performance on only one of the historically problematic morphemes, third person singular -S, was successfully taxed by increasing sentence length. Because one of the four structures was affected by increases in processing load in the manner predicted, it could be argued that a partial simulation was successful. However, for a true simulation to occur, either in whole or in part, it is the pattern of performance across multiple structures, both experimental and control, that must be considered. In this case, since neither control structures and only one target structure behaved as predicted, it appears that a simulation was not achieved.

From these results, however, we can glean that not all structures function the same and may be vulnerable to different factors and to different degrees. For example, in the restricted sample, only third person singular –S, plural –S, and progressive -ING showed a performance difference when the sentence was lengthened. This suggests that, for auxiliary BE, simply increasing the information to be processed within the sentence was insufficient in taxing that structure's baseline performance.

The analyses focusing on the differences in performance between structures also suggested that the attempt at simulating SLI was unsuccessful. This is because the general trends between the structures are not reflective of a performance profile of SLI, even when taking into consideration individual differences in working memory. When restricting the sample size to children who are most likely typically developing, and who displayed understanding of a grammaticality judgment task, the lowest performance was seen unexpectedly for progressive –ING in the long sentence condition. The fact that progressive –ING's low performance remained after selected participants were removed from analysis suggests that low performance on this structure is less likely to be an artifact of possible clinical status or general mastery of grammaticality judgment.

Because progressive –ING has historically been shown to be an easier structure, one explanation is that such low performance for this structure in the long sentence condition may not stem from the nature of the structure itself. Instead, low performance may be driven by the fact that the syntactic violation occurs within a relatively more medial position within the stimuli. In fact, movement of syntactic error location has been recently used as a manipulation of cognitive load (Noonan, Redmond, & Archibald, 2013). In fact, some populations, such as individuals with both language and working memory impairments are more sensitive to syntactic errors occurring in more medial positions than early occurring violations (Noonan et al., 2013). Thus, the unintentional placement of progressive –ING in a comparatively medial position, when combined with load (lengthening), may have contributed additional processing demands that are then reflected in the performance for this structure.

4.4 Filler Structure Performance

4.4.1 Filler Structure Performance with All Participants

Aside from the four target structures, filler sentences featuring a subject-verb agreement error, and their grammatical counterparts, were included in the study. These filler sentences afford an additional way to test whether a simulation of SLI is possible. To recap, two types of agreement sentences were featured. The first type included an inappropriate use of AM for a context requiring IS (e.g., "Today, he am weighing many grapes easily"). This sentence type will be referred to as "He AM". The second type included an inappropriate use of IS for contexts requiring AM (e.g., "Today, I is playing many games happily"). This sentence type will be referred to as "I IS".

An initial bivariate correlation was computed focusing on the relationship between the filler agreement error sentences and the other standardized and experimental measures for the original 61 participants. From the results, seen in Table 10, it is observed that nonword repetition task performance only significantly correlated with one item, I IS in the long condition, while size judgment performance correlated significantly with all four subject-verb agreement error sentence variations. This is not particularly surprising as past research has shown subject-verb agreement error performance is influenced by working memory span, as measured by a size judgment, although the effect sizes of this finding was not reported (McDonald, 2008a).

Table 10: Correlations between Standardized Test Measures, Cognitive Measures, and Filler Structures for All Participants

Measure	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. **PTONI**										
2. **PPVT**	.42**									
3. **DELV-NR**	-.06	.18								
4. **TEGI 3rd -S**	.08	.15	.18							
5. **TEGI BE**	.14	.10	-.01	.07						
6. **Size Judge**	.33**	.26*	.19	.23	.33*					
7. **Nonword Rep**	.12	.32*	.03	.05	.14	.38**				
8. **S HE AM**	.35**	.32*	.16	-.01	.14	.39**	.13			
9. **S I IS**	.22	.22	-.08	.06	.23	.41**	.13	.58**		
10. **L HE AM**	.33**	.25	.26*	.04	.10	.40**	.19	.44**	.32*	
11. **L I IS**	.21	.36**	.12	.09	.13	.42**	.29*	.48**	.55**	.50**

** correlation at the p < .01 level
* correlation at the p < .05 level

To further explore the potential influences of WM and length on agreement error sentences, a repeated measures 2 (length) x 6 (structures) x 2 (low WM span, high WM span) ANOVA was conducted on the full 61 participant sample. The purpose of this analysis was to compare the performance of the two agreement error sentences to the four target structures, while controlling for individual differences in working memory. In line with a successful simulation of SLI, performance on agreement-error sentences would be more akin to the predicted performance of nonproblematic structures (plural -S, progressive -ING) due to the infrequent occurrence of agreement errors for children with SLI (Leonard et al., 2003; Rice et al., 1999). That is, sentences with an agreement error should show greater performance when compared to sentences featuring problematic third person singular -S and auxiliary BE. Results revealed main effects of all three variables: length, $F(1,59) = 9.70$, $p < .001$ (partial $\eta^2 = .14$), structure, $F(5,295) = 5.37$, $p < .001$ (partial $\eta^2 = .08$), and working memory span, $F(1,59) = 6.91$, $p < .05$ (partial $\eta^2 = .11$). In addition, there was a significant two-way interaction between length and structure, $F(5,295) = 5.80$, $p < .001$ (partial $\eta^2 = .09$), which was qualified by a three-way interaction between all three variables, $F(5,295) = 2.32$, $p < .05$ (partial $\eta^2 = .04$), seen in Figure 3.

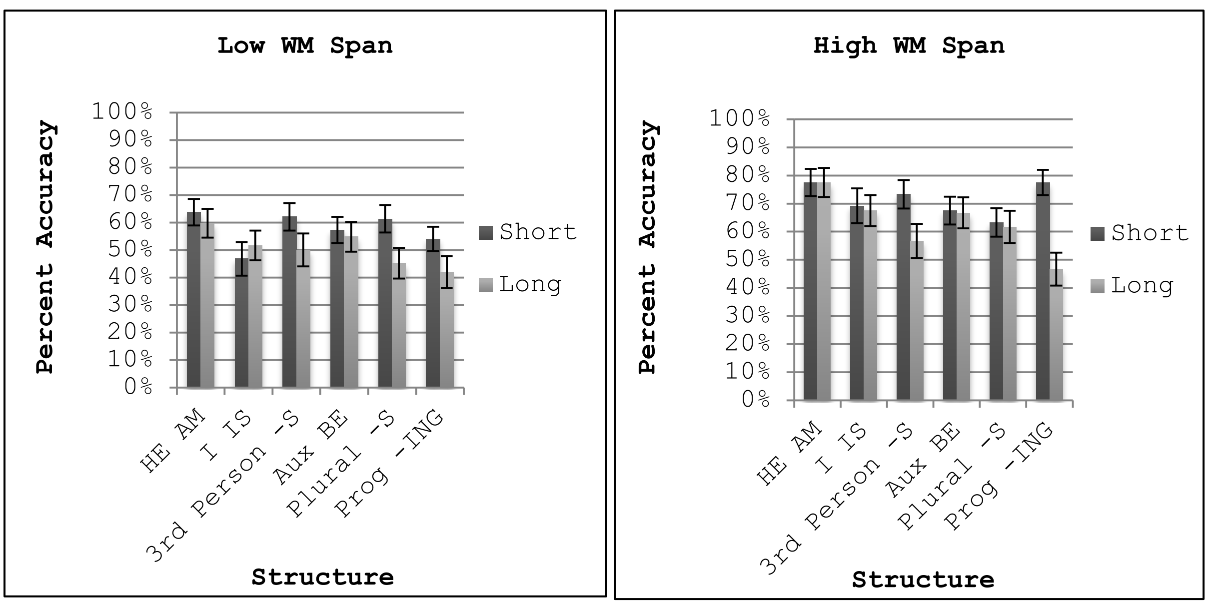

Figure 3: Target and Filler Structure Performance Split by Length for Individuals with Low and High Working Memory Spans (All Participants)

Exploration of this three-way interaction focusing on the subject-verb agreement error sentences revealed that, for the low WM span group, structure differences emerged in both the short, $F(5,150) = 2.68$, $p < .05$ (partial $\eta^2 = .08$) and long, $F(5,150) = 2.55$, $p < .05$ (partial $\eta^2 = .08$) sentence condition. In the short sentence condition, He AM ($M = 63.7\%$) significantly differed from I IS ($M = 46.8\%$), while I IS additionally differed from third person singular -S ($M = 62.1\%$) and plural -S ($M = 61.3\%$). In the long sentence condition, He AM ($M = 59.7\%$) only showed significantly higher performance than both control structures (plural –S: $M = 45.2\%$; progressive –ING: $M = 41.9\%$). For the high WM span group, structure differences only emerged in the long sentence condition, $F(5,145) = 6.42$, $p < .001$ (partial $\eta^2 = .18$). In the long sentence condition, He AM ($M = 77.5\%$) differed from third person singular –S ($M = 56.7\%$), plural –S ($M = 61.7\%$), and progressive –ING ($M = 46.7\%$), while I IS ($M = 67.5\%$) only differed from progressive –ING. Additional t-test analyses comparing subject-verb agreement error performance between participants with low and high WM spans revealed significant differences between low and high WM span individuals for all for length and filler structure combinations: short He AM, $t(59) = -2.03$, $p < .05$, long He AM, $t(59) = -2.42$, $p < .05$, short I IS, $t(59) = -2.58$, $p < .05$, and long I IS, $t(59) = -2.05$, $p < .05$.

Given that children with SLI tend to not often make agreement errors, it was hypothesized that these structures, similar to the selected control structures, would remain robust against the effects of length and individual differences in working memory. As a result, it was hypothesized that these structures would show significantly higher performance than the historically problematic structures. For individuals with lower and higher WM spans, multiple trends surfaced in which subject-verb agreement performance violated the hypotheses. For instance, performance on I IS in the long sentence condition displayed significantly lower performance than the historically difficult third person singular –S for individuals with lower WM spans. Also in the long sentence condition, structure I IS failed to differ from either historically problematic structure for individuals with higher WM spans. Results on the filler subject-verb agreement error sentence continue to compound the conclusion that a successful simulation of SLI was not obtained.

4.4.2 Filler Structure Performance with Selected Participants

The above analyses were conducted on the entire 61 participant sample. However, as discussed above, this sample included participants that may not be considered "typically developing," or otherwise were unable to grasp how to perform a grammaticality judgment task. Therefore, the above correlation and ANCOVA were repeated using the smaller 42 subject sample.

Results from the second correlation continued to implicate working memory as more influential than phonological short term memory in detecting subject-verb agreement errors. As seen in Table 11, after removing participants who may have contributed noise to the data, the single correlation observed between nonword repetition performance and I IS in the long sentence condition no longer surfaced as significant. For the same structure, the correlation with size judgment task performance no longer was significant.

Table 11: Correlations between Standardized Test Measures, Cognitive Measures, and Filler Structures for Selected Participants

Measure	1	2	3	4	5	6	7	8	9	10
1. PTONI										
2. PPVT	.37*									
3. DELV-NR	-.09	.10								
4. TEGI 3rd -S	.07	-.03	.19							
5. TEGI BE	.10	.18	-.01	-.26						
6. Size Judge	.42**	.33*	.05	.14	.16					
7. Nonword Rep	.34*	.50**	.06	.10	.19	.44**				
8. S HE AM	.26	.36*	.20	.03	-.02	.38*	.19			
9. S I IS	.30	.10	-.24	-.07	.14	.45**	.11	.51**		
10. L HE AM	.35*	.27	.27	.04	-.00	.41**	.25	.37*	.28	
11. L I IS	.30	.46**	-.05	.04	-.04	.27	.27	.43**	.51**	.46**

** correlation at the p < .01 level
* correlation at the p < .05 level

The previously performed 2 x 6 x 2 ANOVA was re-run on the smaller sample size in an effort to examine whether any changes in performance would result after restricting the participant sample. Main effects of length, $F(1,40) = 10.06$, $p < .005$ (partial $\eta^2 = .20$), and structure, $F(5,200) = 4.68$, $p < .001$ (partial $\eta^2 = .11$), continued to emerge, qualified by a two way interaction between the two variables, $F(5,200) = 5.44$, $p < .001$ (partial $\eta^2 = .12$), seen in Figure 4. When restricting the sample size, the previously observed main effect of working memory span, and three-way interaction between length, structure, and working memory span, failed to be significant.

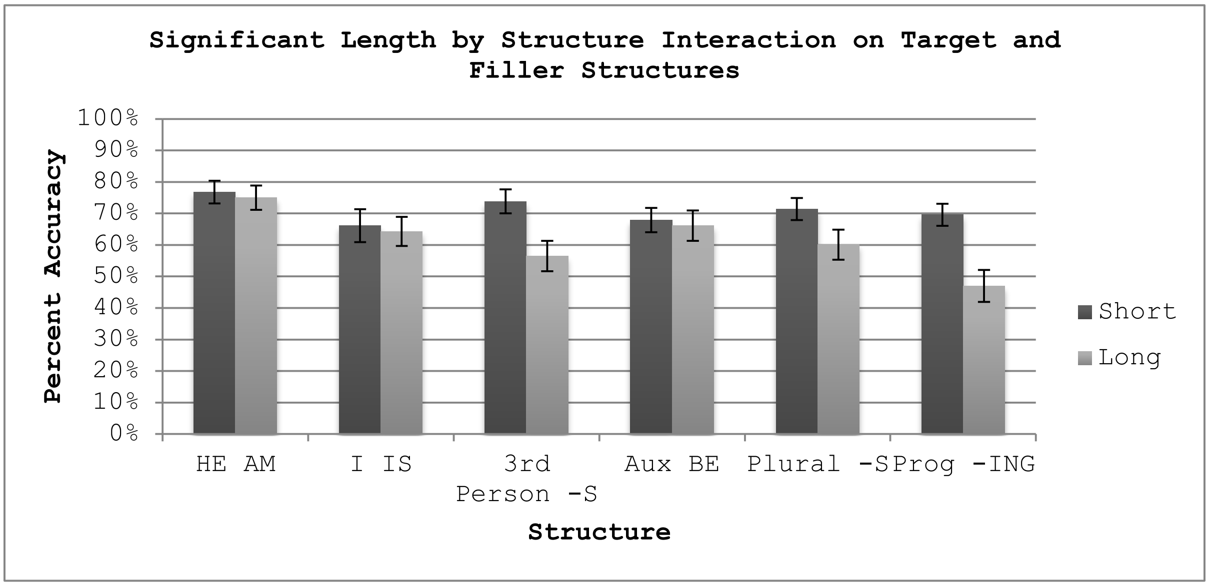

Figure 4: Target and Filler Structure Performance Split by Length (Selected Participants)

Follow-up analyses on the length by structure interaction focusing specifically on the subject-verb agreement sentences revealed that performance for neither He AM nor I IS sentence types was significantly affected by increasing sentence length. Focusing on the relationship of filler sentence performance to the target structures, it was indicated that within the short sentence condition, no structure differences exist. Within the long sentence condition, as previously noted, main effects of structures emerged, $F(5,205) = 8.14$, $p < .001$ (partial η^2 = .17). Specifically concerning the filler structures, He AM (M = 75.0%) was noted to have significantly higher performance than all other structures, while I IS (M = 64.3%) was noted to only differ from He AM and progressive -ING (M = 47.0%).

These results, focusing on the subject-verb agreement error sentences for the restricted participant sample, only partially support a simulation of SLI. As predicted, sentence length did not impact subject-verb agreement error performance, and, in the long condition, performance on He AM was noted to be significantly higher than both historically problematic structures. However, contrary to our hypotheses, in the long sentence condition, He AM also displayed higher performance than both of the control morphemes, and I IS failed to differ in performance from either third person singular –S or auxiliary BE. Even after restricting the participant sample, the findings indicated that a pattern of SLI performance was not obtained for the subject-verb agreement error sentences.

5. CONCLUSION AND DISCUSSION

Previously, Hayiou-Thomas et al. (2004) attempted to simulate a performance profile of SLI by both lengthening the sentence stimuli and compressing the speech stream within a grammaticality judgment task. Performance on both problematic structures (third person singular -S, past tense -ED), as well as performance on one supposed easy structure (prepositions), significantly dropped as a result of the combined load manipulations (Hayiou-Thomas et al., 2004). Only one of the control structures (plural -S) proved resistant (Hayiou-Thomas et al., 2004). Thus, three of the four structures tested responded in a manner that was predicted, resulting in a partial simulation of SLI. However, these results only indicated that some cognitive aspect may be deficient in children with SLI. Because no measures of individual differences were taken, it leaves readers speculating as to which aspect of one's system is potentially underlying the impairment.

The current study expanded Hayiou-Thomas et al.'s (2004) work in two ways. First, the current study examined different structures (auxiliary BE and progressive -ING). Second, individual differences in both phonological short term memory and working memory were assessed, although, due to statistical violation preventing PSTM from being included as a covariate, as intended, only working memory was accounted for in the analyses. From this, our design attempted to not only show if SLI is potentially caused by a cognitive deficit, but also enrich that explanation by suggesting "how".

In the current study, it was predicted that both weaker structures (third person singular -S and auxiliary BE) would display lower performance while both control structures (plural -S and progressive -ING) would remain robust as a result of increasing sentence length. Results from the current study showed although increasing sentence length lowered performance, it did not do so in a manner wholly consistent with our predictions. When including all participants, it was observed individuals with low WM spans were globally impacted by increases in sentence length, while sentence length specifically affected third person singular -S and progressive -ING structures for individuals with higher WM spans. These results only partially support our hypotheses. As predicted, third person singular -S was negatively affected by increases in sentence length, but, violating our hypotheses, progressive -ING was additionally affected. When reducing the sample size, results showed that length negatively impacted all structures except for auxiliary BE. This contradicts the hypotheses in two ways. First, performance on only one historically problematic structure decreased as a function of length. Second, both control structures failed to remain robust. Regardless of which sample was used, mixed results surfaced, which often partially supported one prediction, while violating another. Most curious is the finding that the chosen control structures, particularly progressive -ING, consistently failed to be robust against increases in length.

It was also predicted that, when compared to one another, historically difficult structures (third person singular -S and auxiliary BE) would display lower performance than the selected control structures (plural -S and progressive -ING). Analyses including all participants showed that for individuals with higher WM spans, after length was added to the sentence stimuli, the previously high performance of progressive -ING became significantly lower than both auxiliary BE and plural -S. When restricting the sample size, structure differences emerged in the long sentence condition such that progressive -ING displayed lower performance than all other structures, and third person singular -S additionally displayed lower performance than auxiliary BE. Regardless of which sample size was used, the pattern of results focusing on comparative structure performance failed to support the hypotheses. Contrary to the hypotheses, the lowest performance observed across structures, particularly in the long sentence condition, consistently appeared to be historically easy progressive -ING. Additionally, the fact that performance on plural -S also never statistically differed from either historically problematic structure indicated that it was not as robust as predicted.

These general findings conflict with the prior findings of Hayiou-Thomas et al. (2004) in two ways. First, when processing load increased, the structures that marked for tense and agreement were not consistently the most impacted. Second, plural -S failed to be robust against the increasing sentence demands.

From the findings, it can be concluded that our hypotheses were not supported. Further, a successful simulation of SLI was not accomplished. One finding in particular drove the unexpected results. When comparing performance across structure types, both problematic structures failed to consistently prove more difficult than the control structures selected, even when stimuli load was increased in an attempt to polarize performance between problematic and unproblematic structures. In fact, the lowest performance seen was on one of the control structures (progressive -ING). Reasons why the control structures, especially progressive -ING, proved to be unusually difficult will be explored further in section 5.1.

Due to our unsupported hypotheses, these findings would suggest that a simulation is not possible via increasing stimuli load, at least by increasing sentence length alone. However, the results can possibly explained by possible other influences, also discussed below.

5.1. Why Easy Structures May Be Problematic

When examining the performance differences across morphemes, even after increasing sentence length, neither control structure emerged as having significantly higher performance than the historically difficult structures. This trend of structure performance, at first, appeared to be exceptionally problematic. This was initially troubling as maximum experimental control was exercised during stimuli creation to focus on the error type. That is, all sentences were created equal in terms of sentence structure and location, length, and structure of padding. However, controlling stimuli creation so strictly in one area resulted in the unintentional systematic manipulation of error location. Therefore, errors involving third person singular -S or auxiliary BE omissions always preceded errors involving plural -S or progressive -ING omission.

From past research, we know that error location plays an influential role in performance, although findings are mixed as to which location – frontal or medial (or late) – is more problematic. In past research simulating a performance profile of aphasics in typically developing individuals, it was found that agreement errors towards the front of the sentence, defined as occurring within the first 1200 msec of the recording, led to lower performance than later agreement violations, described as occurring after the 1200 msec point (Blackwell & Bates, 1995). Using the same stimuli and parameters to define early and late errors, Wulfeck et al. (2004) also find that earlier errors lead to greater difficulty than those placed later in a grammaticality judgment task for both typically developing children and those with SLI.

More recent research has looked specifically at early versus medial, and not just "late," errors. Noonan et al. (2013) examined the effects of working memory load within a grammaticality judgment task for children with SLI and children with dual language and working memory impairments, as well as their typically developing peer controls. Sentence load was specifically manipulated by adjusting the error location within the sentences, which averaged approximately 11 words (Noonan et al., 2013). Low load sentences included an error in a frontal position, either 3 or 4 words into the sentence (e.g., "The girls are sit on the bench and giggling to each other"; Noonan et al., 2013). High load sentences contained a medial error located 7 to 9 words into the sentence ("Chris and George will learn to carved a pumpkin for Halloween"; Noonan et al., 2013). Children with SLI differed from their typically developing counterparts regardless of whether the error was in a frontal or medial position (Noonan et al., 2013). However, children with both language and working memory deficits differed from their typically developing counterparts only on sentences with a medially placed violation (Noonan et al., 2013). This suggests that, at least for some populations, medially occurring errors may be more difficult.

Because of how the various studies define early versus late or medial placement, these studies do not necessarily contradict one another. First, Blackwell and Bates (1995) and Wulfeck (2004) and colleagues compared early versus late, not medial, error locations. Although their late errors were defined as occurring after 1200 msec of the sentence recording, from the appendix (Blackwell & Bates, 1995), it can be noted that late errors are almost exclusively placed within the last few words of the sentence (e.g., "John had finished the candy that his mother were saving"). In contrast, Noonan et al. (2013) intentionally placed errors in a more medial position.

Although the research on the role of error location is inconclusive, it may be an influential and explanatory variable for why the control structures used showed lower levels of performance, particularly on the longer sentences. It can be speculated that having to identify more medial errors created an unintentional load that, when combined sentence lengthening, led to progressive -ING's low performance. A possible qualm with this logic is that even within the long sentences of current study, the latest violation type (plural -S) only occurs five words into the sentence. This violation is placed much earlier than the "medial" errors seen in Noonan et al. (2013)'s study and could arguably be more similar to even the less problematic frontal violations.

An alternative explanation for why progressive -ING in particular displays such low performance might lie in where the participant may be focusing his or her attention, anticipating an error. As a recap, two sentences structures were created from which our target morphemes were manipulated. The first sentence structure featured a third person singular -S as the main verb. Within this sentence, third person singular -S and plural -S morphemes were omitted. The second structure featured an auxiliary BE form as the main verb. Within this sentence, auxiliary BE and progressive -ING morphemes were omitted. However, there were also 32 filler sentences, half of which featured an incorrect auxiliary BE form. This creates an unequal proportion in which, for sentences featuring an auxiliary BE form, the syntactic violation revolves around the auxiliary. Because of this, one could speculate that participants may implicitly respond this ratio. Therefore, once a child passed the auxiliary BE form, anticipation for a possible error is lowered, allowing violations involving progressive -ING to be accepted as grammatical. Extending this logic, the addition of the subject-verb agreement filler sentences might draw more attention to verbs in general, which might include third person singular -S, and result in less focus to the information following the main verb (plural -S).

5.2 Is a Simulation of SLI Impossible?

To best answer whether a simulation of SLI using typically developing children is even possible, let's examine the evidence to date. In the case of Hayiou-Thomas et al. (2004), three of the four tested structures successfully simulated a profile of SLI performance under the combined load of increased sentence length and speed. In the current study manipulating only sentence length, even some of those effects were not replicated. Insightful comparisons can be made between the current study and that of Hayiou-Thomas and colleagues, from which to guide future research.

The most obvious contrast between the current study and that of Hayiou-Thomas et al. (2004) is the means by which a simulation was attempted, and, by extension, from which conclusions of simulation feasibility were drawn. In the current study, only one manipulation of stimuli load (sentence lengthening) was employed. By comparison, Hayiou-Thomas and colleagues employed two different means to increase stimuli load. When comparing Hayiou-Thomas et al.'s load manipulations, it becomes clear that performance is more negatively impacted when participants are simultaneously placed under two forms of load, rather than only one. This is because even though both length and speed emerged as main effects, they were both qualified by a three-way interaction with structure (Hayiou-Thomas et al., 2004). In this regard, the current study, even after considering individual differences, may not have commanded the necessary degree of load needed to replicate Hayiou-Thomas et al. While lengthening increases the amount of information to be processed, compressed speech affects the rate at which information is being processed. One could speculate that increasing information amount versus information rate may differentially stress separate aspects of one's cognitive system, with one form of stress potentially being overall more influential. To this end, the current study's design might have been more successful if, instead of focusing on sentence lengthening in particular, it focused on alternative, or even combined, forms of load.

Although the two studies may slightly differ on how the processing demands of the task were increased, they are similar in one regard. In both simulation attempts of SLI, load has been manipulated within the stimuli itself, either by lengthening or speeding up the sentence. However, an alternative measure potentially worth considering as a future direction is to manipulate load by adding an external component. Other simulation studies have manipulated load by introducing a secondary task, such as having participants remember and later recall a string of digits (McDonald, 2006). Such a direction may be appropriate for slightly older children, or even modified for a younger sample. By exploring alternative kinds of load, while continuing to measure individual differences, we can continue to explore if a simulation is possible.

From the two SLI simulation attempts, as well as studies conducted with impaired children, we can identify certain factors that may influence language task performance, and which should be measured or controlled for in any future simulation attempts. First, we know that cognitive abilities—both phonological short term memory (Graf Estes et al., 2007) and working memory (Montgomery & Evans, 2009; Pickering & Gathercole, 2001)—are lower in children with SLI, and have been linked with language task performance in the impaired population (Montgomery & Windsor, 2007; Montgomery & Evans, 2009). Second, the context of the error plays a vital role in performance. Performance should not be speculated solely based on the target morpheme selected. While some morphemes may have a documented history of being either problematic or unproblematic for children with SLI, exceptions exist, potentially driven by the context in which they appear. Some sentential variables, which have influenced past performance, include, but are not necessarily limited to, sentence length (Hayiou-Thomas et al., 2004), sentence speed (Hayiou-Thomas et al., 2004), syntactic error location (Wulfeck et al., 2004), semantic ambiguity (Lum & Bavin, 2007), and sentence structure (e.g., active sentences vs. reflexives; Montgomery & Evans, 2009). Even after considering all the variables above in further experimentation, it may be that a clean simulation of SLI via processing load manipulation is not possible. This information is still valuable in better understanding SLI, and would, instead, lend support for competing theories, which may see working memory deficits as potentially co-morbid with, rather than an underlying cause of, SLI (Archibald & Joanisse, 2009).

As stated earlier, SLI affects approximately 7% of the kindergarten population (Tomblin et al., 1997). Based on estimated rates of reading impairment (Catts, 1991; Wilson & Resucci, 1988), it is assumed that half of these children will later display difficulty with literacy (Tomblin et al., 1997). Through accurate, early diagnosis and successful therapy, some of the consequences of this impairment may be circumvented. However, without a clear consensus of what causes this impairment, the job of clinicians to both diagnose and treat earlier is more difficult. With the ultimate goal of being able to inform therapy efforts for children with SLI, this line of research strives to lay a foundation for ultimately discovering what it is we are actually treating: a language deficit, a cognitive deficit, deficits in both language and cognition, or something more nuanced.

REFERENCES

About LSU university laboratory school. (2012). Retrieved October 16, 2013, from http://www.uhigh.lsu.edu/about/about.html.

Aram, D., & Nation, J. (1975). Patterns of language behavior in children with developmental language disorders. *Journal of Speech and Hearing Research*, *18*, 229-241.

Archibald, L., & Gathercole, S. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders*, *41*(6), 675-693.

Archibald, L., & Joanisse, M.F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research*, *52*, 899-914.

Baddeley, A. (1986). *Working memory*. New York, NY: Oxford University Press.

Baddeley, A., Allen, R., & Hitch, G.J. (2010). Investigating the episodic buffer. *Psychologica Belgica, 50*(3-4), 223-243.

Bates, E., Wulfeck, B., & MacWhinney, B. (1991). Cross-linguistic research in aphasia: An overview. *Brain and Language*, *41*(2), 123-148.

Bialystok, E., & Ryan, E.B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly*, *31*, 229-387.

Blackwell, A., & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience*, *7*, 228-257.

Bock, K., & Miller, C.A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45-93.

Briscoe, J., & Rankin, P.M. (2009). Exploration of a 'double-jeopardy' hypothesis within working memory profiles of children with specific language impairment. *International Journal of Language and Communication Disorders*, *44*(2), 236-250.

Brown, R. (1973). *A first language: The early stages*. Cambridge: Harvard University Press.

Buchner, A., Erdfelder, E., Faul, F., & Lang, A. (2009). G*Power (Version 3.1.2)[Computer program]. http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/

Bunting, M.F., & Cowan, N. (2005). Working memory and flexibility in awareness and attention. *Psychological Research, 69*, 412-419.

Catts, H. (1991). Early identification of dyslexia: Evidence of a follow-up study of speech and language impaired children. *Annals of Dyslexia*, *41*, 163-176.

Cleave, P.L., & Rice, M.L. (1997). An examination of the morpheme BE in children with specific language impairment: The role of contractibility and the grammatical form class. *Journal of Speech and Hearing Research*, *40*(3), 480-492.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Conti-Ramsden, G., & Windfuhr, K. (2002). Productivity with word order and morphology: A comparative look at children with SLI and children with normal language abilities. *International Journal of Language and Communication Disorders*, *37*(1), 17-30.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, *104*, 163-191.

Dale, P. (1976). *Language development: Structure and function.* (2nd Ed.), New York: Holt, Rinehart, & Winston.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.

Deevy, P., Weil, L.W., Leonard. L.B., & Goffman, L. (2010). Extending use of the NRT to preschool-age children with and without specific language impairment. *Language, Speech, and Hearing Services in Schools*, *41*, 277-288.

de Villiers, J.G., & de Villiers, P.A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, *2*(3), 267-278.

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

Dunn, L.M., & Dunn, D.M. (2007). *Peabody Picture Vocabulary Test-IV.* Circle Pines, MN: American Guidance Service.

Dunn, L.M., Dunn, D.M., Whetton, C., & Pintilie, D. (1982). *British picture Vocabulary Scale – Long form.* Windsor, UK: NFER-Nelson.

Ehrler, D.J., & McGhee, R.L. (2008). *Primary Test of Nonverbal Intelligence.* Austin, TX: PRO-ED.

Engel de Abreau, 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.

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

Friedmann, N., & Novogrodsky, R. (2008). Subtypes of SLI: SySLI, PhoSLI, LeSLI and PraSLI. *Langauge Acquisition and Development*, 205-217.

Gathercole, S.E., Tiffany C., Briscoe, J., Thorn, A., & ALSPAC Team. (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *46*, 598-611.

Gathercole, S., & Baddeley, A. (1996). *The Children's Test of Nonword Repetition*. London: The Psychological Corporation.

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.

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

Gottardo, A., Stanovich, K.E., & Siegel, L.S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology*, *63*, 563-582.

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

Grela, B.G., & Leonard, L.B. (2006). The influence of argument-structure complexity on the use of auxiliary verbs by children with SLI. *Journal of Speech, Language, and Hearing Research*, *43*, 1115-1125.

Hadley, P.A., & Rice, M.L. (1996). Emergent uses of *be* and *do*: Evidence from children with specific language impairment. *Language Acquisition*, *5*, 209-243.

Hartsuiker, R., & Barkhuysen, P. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*, 181-204.

Hayiou-Thomas, M.E., Bishop, D.V.M., & Plunkett, K. (2004). Simulating SLI: General cognitive processing stressors can produce a specific linguistic profile. *Journal of Speech, Language, and Hearing Research*, *47*, 1347-1362.

James, S.L., & Khan, L.M.L. (1982). Grammatical morpheme acquisition: An approximately invariant order? *Journal of Psycholinguistic Research*, *11*(4), 381-388.

Khan, L., & James, S. (1983). Grammatical morpheme development in three language disorder children. *Journal of Childhood Communication Disorders*, *6*, 85-100.

Korkman, M., & Hakkinen-Rihu, P. (1994). A new classification of developmental language disorders. *Brain and Language*, *47*, 96-116.

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

Leonard, L.B., Camarata, S.M., Brown, B., & Camarata, M.N. (2004). Tense and agreement in the speech of children with specific language impairment: Patterns of generalization through intervention. *Journal of Speech, Language, and Hearing Research, 47*(6), 1363-1379.

Leonard, L.B., Camarata, S.M., Pawlowska, M., Brown, B., & Camarata, M.N. (2006). Tense and agreement morphemes in the speech of children with specific language impairment during intervention: Phase 2. *Journal of Speech, Language, and Hearing Research, 49*(4), 749-770.

Leonard, L.B., Camarata, S.M., Pawlowska, M., Brown, B., & Camarata, M.N. (2008). The acquisition of tense and agreement morphemes by children with specific language impairment during intervention: Phase 3. *Journal of Speech, Language, and Hearing Research, 51*(1), 120-125.

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*, *30*, 769-795.

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*, 741–52.

Lum, J., & Bavin, E. (2007). Analysis and control in children with SLI. *Journal of Speech, Language, and Hearing Research*, *50*, 1618-1630.

Magimairaj, B.M., & Montgomery, J.W. (2012). Children's verbal working memory: Role of processing complexity in predicting spoken sentence comprehension. *Journal of Speech, Language, and Hearing Research*, *55*, 669-682.

Mazzoni, D., Brubeck, M., Crook, J., Dannenberg, R., Johnson, V., Meyer, M., . . .Wyss, O. (2006). Audacity (1.2.6) [Computer program]. Retrieved July 18, 2011 from http://auacity.sourceforge.net/

McDonald, J.L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*, 395-423.

McDonald, J.L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381-401.

McDonald, J.L. (2008a). Grammaticality judgments in children: The role of age, working memory and phonological ability. *Journal of Child Language, 35*, 247-268.

McDonald, J.L. (2008b). Differences in the cognitive demands of word order, plural, and subject-verb agreement constructions. *Psychonomic Bulletin & Review, 15*(5), 980-984.

McDonald, J.L., & Roussel, C.C. (2010). Past tense grammaticality judgment and production in non-native and stressed native English speakers. *Bilingualism: Language and cognition, 13*, 429-448.

Miller, C.A., Leonard, L.B., & Finneran, D. (2008). Grammaticality judgments in adolescents with and without language impairment. *International Journal of Language and Communication Disorders, 43*(3), 346-360.

Montgomery, J.W. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research, 38*(1), 187-199.

Montgomery, J.W. (2000a). Relation of working memory to off-line and real-time sentence processing in children with specific language impairment. *Applied Psycholinguistics, 21*, 117-148.

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

Montgomery, J.W. (2004). Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. *International Journal of Language and Communication Disorders, 39*(1), 115-133.

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

Montgomery, J.W., & Leonard, L.B. (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.W., & Leonard, L.B. (2006). Effects of acoustic manipulation on the real-time inflectional processing of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *49*, 1238-1256.

Montgomery, J.W., & Windsor, J. (2007). Examining the language performances of children with and without specific language impairment: Contributions of phonological short-term memory and speed of processing. *Journal of Speech, Language, and Hearing Research*, *50*, 778-797.

Morehead, D. (1975). The study of linguistically deficient children. In S. Singh (Ed.), *Measurement procedures in speech, hearing and language.* Baltimore: University Park Press.

Noonan, N., Redmond, S., & Archibald, L. (2013). *Differentiating linguistic and working memory demands on children's grammaticality judgments*. Poster session presented at the Symposium on Research in Child Language Development, Madison, WI.

Oetting, J.B., & Rice, M.L. (1993). Plural acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *36*(6), 1236-1248.

Pickering, S.J., & Gathercole, S.E. (2001). *Working Memory Test Battery for Children*. London: The Psychological Corporation.

Polite, E.J., & Leonard, L.B. (2007). A method for assessing the use of first person verb forms by preschool-aged children with SLI. *Child Language Teaching and Therapy*, *23*(3), 353-366.

Poll, G.H., Betz, S.K., & Miller, C.A. (2010). Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language, and Hearing Research*, *53*, 414-429.

Raven, J.C., Court, J.H., & Raven, J. (1986). *Raven's Coloured Progressive Matrices.* London: H.K. Lewis.

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

Rice, M. L., & Wexler, K. (2001b). *Rice/Wexler Test of Early Grammatical Impairment Examiner's Manual*. San Antonio, TX: The Psychological Corporation.

Rice, M., Hoffman, L., & Wexler, K. (2009). Judgments of omitted BE and DO in questions as extended finiteness clinical markers of specific language impairment (SLI) to 15 years: A study of growth and asymptote. *Journal of Speech, Language, and Hearing Research*, *52*, 1417-1433.

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., & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and 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.

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

Rowland, C.F., Pine, J.M., Lieven, E.V.M., & Theakston, A. (2005). The incidence of error in young English children's wh-questions. *Journal of Speech, Language, and Hearing Research*, *48*, 384-405.

Santelmann, L., Berk, S., Austin, J., Somashekar, S., & Lust, B. (2002). Continuity and development in the acquisition of inversion in yes/no questions: Dissociating movement and inflection. *Journal of Child Language*, *29*, 813-842.

Semel, E., Wiig, E., & Secord, W. (1987). *Clinical Evaluation of Language Fundamentals-Revised*. San Antonio, TX: The Psychological Corporation.

Seymour, H., Roeper, T., & de Villiers, J. (2003). *Diagnostic Evaluation of Language Variation Screening Test.* San Antonio, TX: The Psychological Corportation.

Seymour, H., Roeper, T., & de Villiers, J. (2004). *The Diagnostic Evaluation of Language Variation*. San Antonio, TX: Harcourt.

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.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137-149.

Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., & Smith, E. (1997). Prevalence of Specific Language Impairment in Kindergarten Children. *Journal of Speech, Language, and Hearing Research*, *40*, 1245-1260.

van der Lely, H. (1996). Specific language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua*, *98*, 243-272.

van der Lely, H. (2005). Domain specific cognitive systems: Insight from grammatical specific language impairment. *Trends in Cognitive Science*, *9*, 53-59.

van der Lely, H., & Christian, V. (2000). Lexical word formation in children with grammatical SLI: A grammar-specific versus an input processing deficit? *Cognition*, *75*, 33-63.

van der Lely, H., & Harris, M. (1990). Comprehension of reversible sentences in specifically language impaired children. *Journal of Speech and Hearing Disorders*, *55*, 101-117.

Weismer, S.E., Tomblin, J.B., Zhang, X., Buckwalter, P. Chynoweth, J.G., Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, *43*, 865-878.

Wilson, B., & Risucci, D. (1986). A model for clinical-quantitative classification. Generation I: Application to language-disordered preschool children. *Brain and Language*, *27*, 281-309.

Wilson, B., & Risucci, D. (1988). The early identification of developmental language disorders and the prediction of acquisition of reading skills. In R. Masland & M. Maland (Eds.), *Preschool prevention of reading failure* (187-203). Parkton, MD: York.

Wolfus, B., Moscovitch, M., & Kinsbourne, M. (1980). Subgroups of developmental language impairment. *Brain and Language*, *10*, 152-171.

Wulfeck, B., Bates, E., Krupa-Kwiatkowski, M., & Saltzman, D. (2004). Grammaticality sensitivity in children with early focal brain injury and children with specific language impairment. *Brain and Language*, *88*(2), 215-228.

APPENDIX 1: IRB APPROVAL

ACTION ON PROTOCOL APPROVAL REQUEST

LSU
Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.6792
irb@lsu.edu | lsu.edu/irb

TO: Janet McDonald
Psychology

FROM: Robert C. Mathews
Chair, Institutional Review Board

DATE: August 23, 2012
RE: IRB# 3307

TITLE: Recreating SLI Performance in Typically Developing Children

New Protocol/Modification/Continuation: New Protocol

Review type: Full ___ Expedited _X_ **Review date:** 8/24/2012

Risk Factor: Minimal _X_ Uncertain ______ Greater Than Minimal______

Approved _X_ **Disapproved**_________

Approval Date: 8/24/2012 **Approval Expiration Date:** 8/23/2013

Re-review frequency: (annual unless otherwise stated)

Number of subjects approved: 100

Protocol Matches Scope of Work in Grant proposal: (if applicable)______

By: Robert C. Mathews, Chairman Robt C Mathews

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. SPECIAL NOTE:

**All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb*

APPENDIX 2: APPROVED IRB APPLICATION

Application for Approval of Projects Which Use Human Subjects

This application is used for projects/studies that cannot be reviewed through the exemption process.

LSU

Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/irb

– Applicant, Please fill out the application in its entirety and include two copies of the competed application as well as parts A-E, listed below. Once the application is completed, please submit to the IRB Office for review and please allow ample time for the application to be reviewed. Expedited reviews usually takes 2 weeks. Carefully completed applications should be submitted 3 weeks before a meeting to ensure a prompt decision.

– A Complete Application Includes All of the Following:
- **(A)** Two copies of this completed form and two copies of part B thru F.
- **(B)** A brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts 1&2)
- **(C)** Copies of all instruments to be used.
  - *If this proposal is part of a grant proposal, include a copy of the proposal and all recruitment material.
- **(D)** The consent form that you will use in the study (see part 3 for more information.)
- **(E)** Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project, including students who are involved with testing or handling data, unless already on file with the IRB. Training link: (http://phrp.nihtraining.com/users/login.php)
- **(F)** IRB Security of Data Agreement: (http://research.lsu.edu/files/item26774.pdf)

1) Principal Investigator*: Dr. Janet McDonald **Rank** Professor

*PI **must be** an LSU Faculty Member

Dept: Psychology **Ph:** (225)578-4116 **E-mail:** psmcdo@lsu.edu

2) Co Investigator(s): please include department, rank, phone and e-mail for each

Christy Seidel (doctoral student)
Psychology Department
(832)474-0803
cseide1@tigers.lsu.edu

IRB# 3307 LSU Proposal # ______

- Full
- Expedited
- ✓ Human Subjects Training
- ✓ Complete Application

3) Project Title: Recreating SLI Performance in Typically Developing Children

4) Proposal Start Date: 8/15/2012 **5) Proposed Duration Months:** 15

6) Number of Subjects Requested: 100 **7) LSU Proposal #:**

8) Funding Sought From: NA

Study Approved By:
Dr. Robert C. Mathews, Chairman
Institutional Review Board
Louisiana State University
203 B-1 David Boyd Hall
225-578-8692 | www.lsu.edu/irb
Approval Expires: 8/23/2013

ASSURANCE OF PRINCIPAL INVESTIGATOR named above

I accept personal responsibility for the conduct of this study (including ensuring compliance of co-investigators/co-workers) in accordance with the documents submitted herewith and the following guidelines for human subject protection: The Belmont Report, LSU's Assurance (FWA00003892) with OHRP and 45 CFR 46 (available from http://www.lsu.edu/irb). I also understand that copies of all consent forms **must be maintained at LSU for three years after the completion of the project.** If I leave LSU before that time, the consent forms should be preserved in the Departmental Office.

Signature of PI *JMcDonald* Date 8-9-2012

ASSURANCE OF STUDENT/PROJECT COORDINATOR named above. If multiple Co-Investigators, please create a "signature page" for all Co-Investigators to sign. Attach the "signature page" to the application.

I agree to adhere to the terms of this document and am familiar with the documents referenced above.

Signature of Co-PI (s) *Christy Seidel* Date 8-9-2012

VITA

Christy Seidel is a New Orleans native and high school graduate of Mount Carmel Academy. She received her Bachelors in Science degree in psychology in 2007, with minors in chemistry and Russian language, from Louisiana State University before being accepted into Louisiana State University's doctoral program for cognitive psychology to work with Dr. Janet McDonald.

In 2010, Christy received her Master of Arts in psychology and also began working with Drs. Janna Oetting, Janet McDonald, and Michael Hegarty as part of a National Institutes of Health-funded grant team looking at Specific Language Impairment in dialect speaking children, which was the inspiration for this dissertation.