The Pennsylvania State University
The Graduate School
The College of Health and Human Development

SENTENCE PROCESSING FACTORS IN ADULTS
WITH SPECIFIC LANGUAGE IMPAIRMENT

A Dissertation in
Communication Sciences and Disorders
by
Gerard H. Poll

© 2012 Gerard H. Poll

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May, 2012
The dissertation of Gerard H. Poll was reviewed and approved* by the following:

Carol A. Miller  
Associate Professor of Communication Sciences and Disorders and Linguistics 
Dissertation Co-Advisor 
Chair of Committee

Janet G. van Hell  
Professor of Psychology and Linguistics 
Dissertation Co-Advisor

Robert A. Prosek  
Professor of Communication Sciences and Disorders

James T. Herbert  
Professor of Education

Krista M. Wilkinson  
Professor of Communication Sciences and Disorders

Kathryn D. R. Drager  
Associate Professor of Communication Sciences and Disorders 
Professor-in-Charge of the Graduate Program

*Signatures are on file in the Graduate School
ABSTRACT

Sentence imitation effectively discriminates between adults with and without specific language impairment (SLI). Little is known, however, about the factors that result in performance differences. This study evaluated the effects of working memory, processing speed, and argument status on sentence imitation. Working memory was measured by both a storage and processing and a scope of attention measure. Performance differences on arguments versus adjuncts also tested the predictions of the Procedural Deficit Hypothesis (PDH) of specific language impairment. The PDH predicted that adjunct processing would be more difficult for adults with SLI because it would depend on a deficient procedural memory system. Alternatively, adults with SLI may compensate by engaging an intact declarative memory to process adjuncts.

Twenty-three adults with SLI and 23 typical language controls participated in three experiments. First, sentence processing times were recorded for argument and adjunct conditions in a self-paced listening task. Second, working memory and processing speed were measured, and used to predict accuracy in a sentence imitation task. The sentence imitation materials varied in length and argument versus adjunct composition. Finally, procedural and declarative memory systems were assessed for all participants.

In the online sentence processing task, there was a processing time advantage for noun arguments as compared to verb adjuncts, but no group difference. Correlations of argument and adjunct frequency with processing times were found for the group with SLI, but only argument frequency effects for the typical language group. Group differences as well as argument differences were found in the sentence imitation task. Participants’ working memory also contributed to sentence imitation accuracy, but processing speed did not. Argumenthood, working memory and sentence length interacted, with greater argument status effects in short
conditions, and greater working memory effects in long conditions. Both the storage and 
processing and the scope of attention measures predicted sentence imitation. Declarative memory 
performance was poorer in the group with SLI than in the group with typical language, whereas 
there was limited support for a procedural memory deficit in the group with SLI.

Overall support for the PDH was mixed. There was evidence of compensatory 
processing of adjuncts as predicted by the PDH, but the basis for the compensation was unclear 
given the unexpected poor declarative memory performance. The factors of group classification, 
working memory and argumenthood were confirmed as factors affecting sentence imitation 
performance, and the complexity of the task was revealed by interactions of these factors with 
sentence length. The association of a scope of attention measure to language performance 
suggested that the storage aspect of working memory was as crucial to adult SLI performance as 
abilities for concurrent storage and processing or rehearsal.
# Table of Contents

**List of Figures** ........................................................................................................... viii

**List of Tables** ............................................................................................................... ix

**Acknowledgements** ..................................................................................................... x

Chapter 1  Introduction ...................................................................................................... 1

Chapter 2  Review of the Literature ................................................................................ 4
  - Specific language impairment in adulthood .............................................................. 4
  - Spoken language abilities of adults with specific language impairment .................. 5
  - Cognitive deficits in adults with specific language impairment .............................. 6
  - Assessment of adults at risk for specific language impairment ............................... 8
  - Theoretical accounts of specific language impairment ............................................ 9
    - Linguistic accounts .............................................................................................. 10
    - Information processing accounts ....................................................................... 12
    - The Procedural Deficit Hypothesis .................................................................... 19
  - Argument preference in sentence processing .......................................................... 32
    - Argument effects in sentence processing ........................................................... 33
    - Argument effects and compensation in adults with specific language impairment ... 38
    - Self-paced listening ........................................................................................... 40
  - Sentence imitation ................................................................................................. 42
    - Working memory and sentence imitation ......................................................... 43
    - Syntactic structure effects in sentence imitation ............................................... 44
    - Processing speed and sentence imitation .......................................................... 45
  - Research questions and experiment summaries .................................................... 46

Chapter 3  Experiment 1: Self-Paced Listening ............................................................... 52
  - Participants ........................................................................................................... 52
  - Materials for the self-paced listening task ............................................................ 55
    - Corpus analyses ............................................................................................... 56
    - Norming studies .............................................................................................. 58
    - Confirmation of argument status ..................................................................... 60
    - Procedures ....................................................................................................... 60
    - Analytical approach ........................................................................................ 63
  - Results and discussion ......................................................................................... 65
    - Models of argument status and group effects .................................................. 69
    - Models of the noun argument and verb adjunct contrast .................................. 70
    - Correlations of co-occurrence probability and response times ......................... 72
    - Discussion ....................................................................................................... 78

Chapter 4  Experiment 2: Sentence Imitation ................................................................. 82
  - Materials ............................................................................................................ 82
  - Corpus analyses ................................................................................................ 83
Norming studies ................................................................. 85
Procedures ........................................................................ 85
Working memory tasks ......................................................... 87
Processing speed tasks .......................................................... 90
Procedures ........................................................................ 92
Analytical approach ............................................................ 93
Results: Working memory, speed of processing and sentence imitation ........... 95
  Working memory and processing speed ........................................ 95
  Sentence imitation performance ............................................ 97
  Correlation of predictors to sentence imitation accuracy .............. 99
  Models of factors affecting sentence imitation accuracy .......... 100
  Comparing the Competing Language Processing Task and Running Span .......... 105
Discussion ................................................................. 106
  Predictions of the Procedural Deficit Hypothesis ......................... 106
  Working memory and language processing speed .................. 108
  Factors related to sentence imitation accuracy .................. 109
Chapter 5 Experiment 3: Visual Procedural and Declarative Memory Tasks ............ 113
  The Weather Prediction Task ............................................ 113
    Procedures and scoring .............................................. 116
  The Paired Associate Recognition Task ................................ 117
    Procedures and scoring .............................................. 118
Results: Visual procedural and declarative memory tasks .................. 119
  The Weather Prediction Task ............................................ 120
  The Paired Associate Recognition Task ................................ 121
Discussion ................................................................. 123
Chapter 6 General Discussion ................................................................. 125
  Limitations and differences with prior studies .................. 125
    Methodological differences with prior studies .................. 126
    Participant differences .............................................. 128
  Predictions of the Procedural Deficit Hypothesis ......................... 130
    Evidence consistent with the Procedural Deficit Hypothesis ............. 130
    Evidence less consistent with the Procedural Deficit Hypothesis ............. 132
    Refinements to the Procedural Deficit Hypothesis ..................... 134
  Support for other accounts of specific language impairment .................. 135
    Refinements to the limited working memory account of specific language impairment .............................................. 137
  Understanding sentence imitation ........................................ 141
    The meaning of sentence imitation varies by sentence length ............. 142
  Argument preference effects in sentence comprehension .................. 144
  Implications for professional practice .................................... 147
    Slower processing not always a hindrance to language task performance .................. 147
    Working memory deficits in adults with specific language impairment .................. 148
    An explanation of Specific Language Impairment incorporating compensation ....... 150
    Understanding factors affecting sentence imitation performance .................. 151
Future directions ................................................................. 152
  Further testing of the procedural deficit hypothesis.................................153
  Extending new findings on working memory in specific language impairment.....154
  Processing speed and sentence imitation.....................................................155
  Online sentence processing of adults with specific language impairment ........156
  General conclusions ....................................................................................157

Appendix A  Self-Paced Listening Materials ...............................................159

Appendix B  Sentence Imitation Materials .....................................................161

References .................................................................................................163
LIST OF FIGURES

**Figure 1** Sentence structures for noun argument (left) and verb adjunct (right) conditions showing that the noun argument condition has more nodes ............................................. 35

**Figure 2** The purpose, predictions and evidence required to support the predictions for the sentence comprehension task. .................................................................................... 48

**Figure 3** The purpose, predictions, and evidence required to support the predictions for the sentence imitation task. .............................................................................................. 49

**Figure 4** The purpose, predictions and evidence required to support the predictions for the visual declarative and procedural memory tasks........................................................ 50

**Figure 5** Mean difference time in milliseconds by sentence region, by argument versus adjunct conditions ............................................................................................................ 67

**Figure 6** Mean difference times in milliseconds by sentence region for noun argument versus verb adjunct conditions. Group with typical language top panel, group with SLI bottom panel .............................................................................................................. 68

**Figure 7** Scatterplots of adjusted listening times and the probability of co-occurrence of a prepositional phrase with the same preposition and sense following the associated noun argument (top panel) or verb adjunct (bottom panel) ................................................. 74

**Figure 8** Mean proportion of words repeated correctly for each sentence imitation condition, by group ....................................................................................................................... 99

**Figure 9** Mean proportion of predictions aligned with probabilities by block and group for the Weather Prediction Task ...................................................................................... 120
LIST OF TABLES

Table 1 Mean (SD) demographic, performance intelligence and language measures for participants ....................................................................................................................... 55

Table 2 Mean (SD) difference times and adjusted listening times by condition, group, and sentence region .............................................................................................................. 66

Table 3 Correlations (Spearman's Rho) of prepositional phrase region adjusted listening times and the co-occurrence probabilities of the argument or adjunct phrases ............... 77

Table 4 Mean (SD) performance, tests of group mean differences, and effect sizes for the working memory and language processing speed tasks ......................................................... 96

Table 5 Mean (SD) performance by condition and group for the sentence imitation task. Scores for both the ordinal (0-4) and percentage correct scoring methods .................. 98

Table 6 Correlations (Pearson) of sentence imitation grand mean percentage correct with scores of working memory and processing tasks ......................................................... 100

Table 7 Final mixed effects regression model predicting sentence imitation accuracy. The model also included random intercepts by subject, and random slopes by length by participant ................................................................................................................. 102

Table 8 Card patterns, associated outcomes, and probabilities of sun for each card combination presented in the Weather Prediction Task ................................................................. 115
ACKNOWLEDGEMENTS

This research would not have been possible without the support of members of the language science community at Penn State. First, I would like to thank my advisor, Dr. Carol A. Miller, for many years of excellent guidance. Dr. Janet van Hell provided countless insights as a co-sponsor of this project, and Dr. Guili Dussias played a major role in the initial proposal for this research. Dr. Judith Kroll was generous with her time and lab resources in support of this research, and the Center for Language Science provided excellent opportunities for me to share this research in formative stages.

In addition to my sponsors, I would like to thank members and prior members of my doctoral committee. These include Drs. Elina Mainela-Arnold, Maya Misra, Carol Hammer, Robert Prosek, Krista Wilkinson, and James Herbert. I would also like to thank Dr. Gordon Blood for his role in encouraging me to pursue clinical and research training in speech-language pathology.

Beyond Penn State, Dr. Bruce Tomblin’s research team at the University of Iowa was exceptionally generous in sharing their time and resources to help me recruit for this study. I would like to particularly thank Ms. Marlea O’Brien, Ms. Connie Ferguson, and Ms. Marcia St. Clair for their intensive work on the Iowa recruitment. Dr. Tomblin, Dr. Amanda Owen-Van Horne, and Dr. Melissa Duff were all very helpful during discussions when this project was still in the formative stage.

This work was supported by award number F31DC010960 from the National Institute on Deafness and Other Communication Disorders. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

Finally, this work would have been impossible without the steadfast support of Kathleen Poll.
Chapter 1

Introduction

A considerable body of evidence indicates that specific language impairment (SLI) often persists into adulthood (Hall & Tomblin, 1978; Johnson et al., 1999; Tomblin, Freese, & Records, 1992; Whitehouse, Line, Watt, & Bishop, 2009). SLI is a developmental disorder characterized by unusual difficulty learning and processing language in the absence of hearing impairment, neurological injury, intellectual disability, or other likely causes for the language difficulties (Leonard, 1998). The functional effects of the disorder are considerable, and include more limited educational attainment and circumscribed career prospects (Johnson, Beitchman, & Brownlie, 2010). Despite their language limitations and significantly higher risk for learning disabilities, many young adults with SLI pursue post-secondary education to prepare for the job market (Carroll & Dockrell, 2010; Young et al., 2002). The factors that allow individuals with SLI to succeed in post-secondary education are not well understood, but for adults with more broadly defined learning disabilities, it is clear that post-secondary success is associated with self-awareness, self-advocacy, and a willingness to take extraordinary measures to compensate for learning limitations (Reis, McGuire, & Neu, 2000). In short, accurate and efficient identification of language impairment in adulthood, as an enabler of self-awareness and compensatory services, is one important element in mitigating the effects of the disorder.

Diagnosis of SLI in adulthood is hampered by the lack of appropriate norm-referenced measures (Barry, Yasin, & Bishop, 2007) and by the fact that many children with SLI are not identified during the primary and secondary school years (Hagaman, Trout, DeSalvo, Gehringer, & Epstein, 2010; Tomblin et al., 1997). Furthermore, the language difficulties of affected adults may be relatively subtle in casual conversation (Tomblin et al., 1992), further reducing the
likelihood of identification. Tasks that involve comprehending and repeating sentences, however, have been shown to be among the most promising ways to distinguish adults with SLI from those with typical language abilities (Fidler, Plante, & Vance, 2010; Poll, Betz, & Miller, 2010; Tomblin et al., 1992). These tasks have long appeared on standardized tests of language ability designed for children and adolescents, as well as those designed to identify adults at risk for acquired language disorders (Benton & DeS Hamsher, 1978; Hammill, Brown, Larsen, & Wiederholt, 1994; Semel, Wiig, & Secord, 2003).

Sentence imitation and sentence comprehension tasks appear to tap a set of abilities that fundamentally differ in individuals with SLI when compared to individuals with typical language. Sentence imitation has been proposed as clinical marker of SLI in several studies, suggesting that it reveals a defining behavior of the disorder in children and adults (Archibald & Joanisse, 2009; Conti-Ramsden, Botting, & Faragher, 2001; Poll et al., 2010). The validity of the task has been extended to a language quite different from English, Cantonese (Stokes, Wong, Fletcher, & Leonard, 2006). Performance on the task differentiates individuals with SLI not only from typical individuals, but also from those with attention deficit/hyperactivity disorder (Redmond, 2005; Redmond, Thompson, & Goldstein, 2011).

Despite the evident diagnostic value of sentence imitation and sentence comprehension, little is known about the cognitive or linguistic factors that result in the less accurate performance of adults with SLI. Most researchers suggest that memory limitations are one factor, but others suggest that grammatical factors must also play a role (Clegg, Hollis, Mawhood, & Rutter, 2005; Helenius, Parviainen, Paetau, & Salmelin, 2009; Stokes et al., 2006). The purpose of this study is to evaluate the contributions of several candidate factors that may better explain performance on these sentence-level tasks.

A clearer understanding of sentence imitation and sentence comprehension performance has several benefits. From a clinical perspective, an understanding of the factors that result in
performance differences will provide information for professionals concerned with identification and treatment of adults with developmental language impairments. Better understanding of factors related to performance on sentence-based tasks can provide information to guide interventions or accommodations.

From a research perspective, a better understanding of performance on sentence-level tasks provides a window into the nature of SLI in adulthood. This study tested the predictions of a theory of SLI, the Procedural Deficit Hypothesis (PDH) (Ullman & Pierpont, 2005). The PDH is among the most comprehensive theories of SLI to date. It integrates existing evidence on the language and cognitive characteristics of individuals with SLI, and proposes a neurological basis for the disorder. According to the PDH, a deficit in implicit, procedural learning underlies many of the language and cognitive deficits observed in individuals with SLI. This study sought to advance the understanding of SLI by testing the predictions of the PDH.

The predictions of the PDH were evaluated in part by application of theories and research methods from the sentence processing literature. This large body of literature has focused on the cognitive and linguistic mechanisms involved in converting written and spoken sentences into meaningful messages. The findings on sentence processing have been largely confined to the performance of typical language adults. A final purpose of the study is to extend the findings of the sentence processing literature to adults with developmental language disorders.
Chapter 2
Review of the Literature

Specific language impairment in adulthood

Specific language impairment is a developmental disorder characterized by unusual difficulties in the ability to learn, produce, or comprehend language (L. B. Leonard, 1998). Poor language abilities are often documented by scores well below the mean on norm referenced tests (Plante, 1998). Alternatively, disordered language can be identified by tasks or batteries of tasks that reveal behaviors characteristic of the disordered population, and therefore discriminate between affected and unaffected groups (Fidler et al., 2010; Rice & Wexler, 1996; Spaulding, Plante, & Farinella, 2006). Researchers identifying individuals with SLI rule out obvious factors that can contribute to impaired language. These include hearing loss, intellectual disability, neurological injury, or autism spectrum disorders (L. B. Leonard, 1998).

After initial identification of SLI in kindergarten, 73-78% of individuals remain language impaired at young adulthood (Johnson et al., 1999; Whitehouse et al., 2009). Those affected with SLI in adulthood have lower educational attainment than typical peers (Carroll & Dockrell, 2010). This may be related to their higher risk for learning disabilities. Young adults identified in kindergarten with language impairment and followed to adulthood were more than 11 times more likely to have learning disabilities in reading, spelling and math when compared to typical language controls (Young et al., 2002). While dyslexia and SLI are likely separate disorders, the presence of SLI puts individuals at considerable risk for word reading and reading comprehension difficulties (Catts, Adlof, Hogan, & Ellis Weismer, 2005).
Spoken language abilities of adults with specific language impairment

Studies of adults with SLI have shown that their spoken language deficits are wide ranging, spanning receptive and expressive language tasks, and multiple domains of language (Tomblin et al., 1992). At the level of speech sound perception, adults with SLI perform more poorly than typical peers on tasks involving perceiving, analyzing and manipulating sound segments. Pig latin and nonsense word repetition tasks are more difficult for affected adults, particularly when the sound sequences are more than 1 or 2 syllables (Clegg et al., 2005; Poll et al., 2010). They are also less able to perceive rapidly changing sound sequences (Tomblin et al., 1992).

Lexical processing differences in adults with SLI have included less fluency in word association tasks and slower word naming speeds (Tomblin et al., 1992). A recent magnetoencephalography (MEG) study found that adults with SLI have a normal electrophysiological response to lexical processing—the N400 (Helenius et al., 2009). There was evidence, however, that participants with SLI processed words more like non-words, with a more pronounced, longer-lasting brain response as compared to controls. Adults with developmental language weaknesses were also less likely to show distinctions between categorically related versus unrelated words in event-related potentials (Plante, Van Petten, & Senkfor, 2000). Overall, the evidence suggests that lexical representations may be less clearly defined, and retrieval of lexical items may be slower in adults with SLI as compared to typical controls.

Evidence that adults with SLI have deficits in syntax and morphosyntax include findings of depressed performance on standardized tests of grammar (Johnson et al., 1999) as well as poor performance on sentence comprehension and sentence imitation tasks (Miller & Poll, 2009; Poll et al., 2010; Tomblin et al., 1992). Adults with SLI have also been shown to be less sensitive to
errors in tense and agreement marking when judging the grammaticality of sentences (Poll et al., 2010).

**Cognitive deficits in adults with specific language impairment**

As a group, adults with SLI have been found to have a range of cognitive deficits when compared to typical language peers. Despite selection criteria that ensure non-verbal abilities in the typical range, adults with SLI have differences in processing speed, short-term and working memory, and in implicit learning (Miller & Poll, 2009; Plante, Gomez, & Gerken, 2002; Tomblin et al., 1992).

Implicit learning involves acquiring knowledge of a pattern or sequence without awareness of that learning (Evans, Saffran, & Robe-Torres, 2009; Squire, 1992). Implicit learning has been widely tied to language learning, both in terms of word learning (Evans et al., 2009) and syntactic learning (Plante et al., 2002), but is also viewed as a domain-general learning ability (Tomblin, Mainela-Arnold, & Zhang, 2007). Two studies have documented that adults with SLI are more limited in their ability to implicitly learn patterns found in language. In one study, adults with SLI were able to learn how to identify gender categories based on cues in the sound patterns of an unfamiliar language, but their accuracy was lower than typical language peers (Richardson, Harris, Plante, & Gerken, 2006). In a second study, adults with SLI and controls were briefly exposed to an artificial language (Plante et al., 2002). They were then asked to judge whether new strings conformed to the rules of the artificial language. Typical language adults were able to make these judgments at a level that exceeded chance, whereas adults with SLI were at chance or below. These studies suggest that implicit learning may be an area of weakness for adults with SLI, but the current evidence is limited.
Adults with SLI have been shown to differ from typical language adults in short-term and working memory tasks, but results depend on the demands of the task. For simpler short-term memory tasks, such as repeating lists of digits (digits forward), studies have found no group differences or moderate effect size differences (Isaki, Spaulding, & Plante, 2008; Poll et al., 2010). Group differences have been more consistent, and have had larger effect sizes, for tasks that demand not only storage of items, but also concurrent processing of items—working memory tasks. Examples include repeating lists of digits in reverse order (digits backward) or repeating sentences after answering questions.

Adults with SLI may also differ from their typical language peers in speed of processing for cognitive, linguistic, and motoric processes. Here again, group differences are more reliably found for more complex tasks rather than simpler tasks. For example, adults with SLI pressed a button in response to a “go” sign and distinguished whether words were the same or different just as fast as their typical language peers (Fisher, Plante, Vance, Gerken, & Glattke, 2007). However, other studies have found that adults with SLI have a slower speaking rate, and are slower than typical peers in completing an array of motor, cognitive, and language processing tasks of varying complexity (Miller & Poll, 2009; Tomblin et al., 1992).

Given the range of cognitive deficits for adults with SLI, the evidence does not seem to support SLI as a deficit specific to language. Processing speed is associated with intelligence measures throughout development in typical children, and the effect of speed differences appears to be mediated by working memory (Fry & Hale, 2000). In fact, there is evidence that cognitive ability and language ability interact over the course of development in SLI, resulting in a long-term decline in non-verbal intelligence relative to typical language peers over the course of development (Botting, 2005). Recognizing that lower non-verbal ability in adults with SLI as compared to typical peers is a likely result of a lifetime of more limited language ability, many
studies of this population have intentionally not matched the groups with SLI to typical language controls on non-verbal intelligence (Fidler et al., 2010; Tomblin et al., 1992).

In considering learning disabilities and developmental disabilities in general, intelligence measures may be better viewed as an outcome measure rather than representative of a causal construct for the disorder (Dennis et al., 2009). Any attempt to match or statistically equate a developmentally disordered group, such as adults with SLI, with a developmentally typical group on IQ would result in sample groups that are likely not representative of the disordered population. In a recent study of adults with SLI, the researchers took this perspective, and further argued that attempts to statistically control for group differences in non-verbal intelligence measures would result in adjusting language ability measures to more similar distributions across groups. This, in turn, would increase the likelihood of Type II error (Fidler et al., 2010).

Assessment of adults at risk for specific language impairment

There are few norm-referenced tests appropriate for the assessment of developmental language disorders in adults (Barry et al., 2007). The use of such norm-referenced tests is further complicated by the absence of evidence that they accurately discriminate between adults with typical language and those with language impairments. In particular, the use of cut-off scores at one or 1.5 standard deviations below the mean on standardized tests has been criticized as arbitrary (Spaulding et al., 2006). An alternative is the use of empirically validated approaches for identification of adults with language impairment. Two studies have developed discriminant functions that accurately classify adults into typical and language impaired groups (Fidler et al., 2010; Tomblin et al., 1992). Fidler and colleagues developed and evaluated a discriminant function using scores from written spelling, defining words, and following directions tasks, and compared the results to a battery of tasks developed by Tomblin and colleagues (1992). Both sets
of tasks were administered to the same set of participants. The battery developed by Fidler and colleagues had higher classification accuracy than did the battery developed by Tomblin and colleagues. For young adults, the sensitivity of the battery was 78% and the specificity was 83%, compared to 69% sensitivity and 78% specificity for the Tomblin and colleagues (1992) battery. The battery of tasks developed by Fidler and colleagues (2010) currently has the best documented ability to classify adults as developmentally language impaired.

A task common to the most discriminating batteries of tasks evaluated by Fidler and her colleagues (2010) was sentence comprehension—the ability to follow increasingly complex, sentence-level directions. Neither the Fidler and colleagues nor the earlier Tomblin and colleagues (1992) studies provided a theoretical account for the distinct performance of adults with SLI on such sentence-based tasks. This study will evaluate the predictions of the Procedural Deficit Hypothesis (PDH) for sentence comprehension and sentence imitation tasks. The PDH integrates several other theoretical views on SLI.

Theoretical accounts of specific language impairment

While there is some consensus on the language and cognitive traits of individuals with SLI, there is little consensus on the causes of the disorder (L. B. Leonard, 1998). Leading theories have tended to come from either a linguistic or an information processing perspective. Those taking a linguistic perspective are rooted in theories of syntax, and consider gaps in linguistic knowledge as fundamental to the disorder (Rice, Wexler, & Cleave, 1995; van der Lely, 1998). Information processing theories have argued that SLI results from general cognitive factors that are important not just to language, but also to non-verbal tasks (Ellis Weismer, Evans, & Hesketh, 1999; Kail, 1994). Both linguistic and information processing accounts provide insight into sentence processing and sentence imitation performance. The Procedural Deficit
Hypothesis (PDH) attempts to integrate the two perspectives (Ullman & Pierpont, 2005), and so may provide a more complete explanation of sentence processing in SLI. After a brief review of selected linguistic and information processing theories, I will turn to the PDH.

**Linguistic accounts**

A well-documented linguistic account of SLI is the Agreement Tense Omission Model (ATOM) (Wexler, Schutze, & Rice, 1998). This theory is based on the observation that children with SLI will omit tense and agreement markers in sentences at much later ages than will children with typical language. For example, children with SLI are more likely to produce the sentence, *Him stand_ on chairs* for *He stands on chairs*. This difference is attributed to a gap in language knowledge. Children with SLI do not recognize the requirement to use the appropriate form of the pronoun (subject case (*he*) rather than accusative case (*him*)), and to include the –s affix to mark the tense and third person singular agreement with the singular subject. The knowledge gap for tense and agreement marking is proposed to be specific to the linguistic system, and determined by a genetic, maturational process (Wexler, 1996).

There is considerable evidence that children with SLI have the predicted problems with tense and agreement marking, and that signs of these deficits are present into adolescence and adulthood. Adolescents with SLI at age 15 performed well below their peers on tense and agreement marking (Rice, Hoffman, & Wexler, 2009). Adults with SLI are also less accurate in their ability to detect errors of tense and agreement (Poll et al., 2010). Omissions of tense and agreement may account for some, but likely not all errors made in sentence imitation tasks.

A second theory of SLI with a linguistic basis is the Computational Grammatical Complexity (CGC) hypothesis (van der Lely, 2005). This theory is consistent with the findings of the ATOM, but accounts for a wider array of sentence processing phenomena (van der Lely,
The fundamental problem in SLI by the CGC account is a deficit in the ability to compute complex hierarchical relationships found in the grammar and phonology of language. Beyond marking tense in a sentence, an example of a hierarchical dependency is the link between a pronoun and its referent. In the sentence *Daffy Duck says that Bugs Bunny is hugging him*, the pronoun *him* must be linked to *Daffy Duck* for the sentence to be fully comprehended. Structures involving these dependent relations across words in sentences are among the forms that children with SLI have difficulty comprehending, supporting the CGC hypothesis (Montgomery & Evans, 2009; van der Lely, 1998).

An important element of these linguistic theories of SLI is that they assert that grammatical ability is isolated from other aspects of cognition, and they view grammatical computations as central to this domain-specific view (van der Lely, 2005). Other aspects of language, such as learning the arbitrary associations of words and referents, may involve domain-general abilities. An example of co-existing domain-specific (grammatical) and domain-general (word learning, associative memory) mechanisms is the formation of past-tense verbs. Domain-specific accounts, such as the CGC, suggest that regular past tense formation is done by the application of a grammatical rule (verb stem + *-ed*), whereas irregular past tense formation is done by recall of associations (the past tense of *buy* is *bought*), a lexical process (Pinker, 1991; Pinker & Ullman, 2002; van der Lely, 2005). In children with SLI, the deficit in the grammatical system results in the use of the associative mechanism for the formation of all verbs, both regular and irregular. Studies of children with SLI have shown that they are more accurate producing the past tense for more frequent regular verbs when compared to less frequent regular verbs, and that such frequency effects do not exist for typical language children (Van der Lely & Ullman, 2001). These researchers argue that typical language children apply a grammatical rule to form regular past tense, resulting in similar performance across more and less frequent verbs.
Linguistic theories of SLI have focused on grammatical deficits as the central factor in the language difficulties of children with SLI, and view language processing as involving both rule-based processes relying on grammar knowledge, and associative processes relying on domain-general learning and memory mechanisms. Processing-based theories of SLI take a fundamentally different perspective.

**Information processing accounts**

In contrast to the linguistic theories of SLI, information processing accounts of the disorder assert that grammatical processing is not isolated from other cognitive abilities. Cognitive abilities such as auditory perception, speed of processing, and working memory capacity can affect grammatical abilities, and therefore the sentence processing of children with SLI (L. B. Leonard, 1998). Furthermore, language processing is proposed to take a single path, rather than the dual path (rule-application and associative memory) proposed by linguistic accounts (Joanisse, 2004). Two information processing accounts are particularly relevant to sentence processing and production abilities: the generalized slowing and working memory limitations accounts.

**Generalized slowing**

Individuals with SLI have been shown to have a slower response time across many cognitive and language operations, and such slowing may be fundamental to their difficulties in language production and comprehension (Kail, 1994). This observation has led to the generalized slowing hypothesis (L.B. Leonard, 1998). This theory asserts that the response times of individuals with SLI are slower by a common multiple across domains (motor, language,
cognition) and task types. So whether the task is a simple motor response or a complex cognitive operation, the SLI group tends to be slower by some multiple of the response time of individuals without SLI. Researchers have suggested that such results may be consistent with differences in synaptic transmission rates across the neural networks of individuals with SLI as compared to those with typical language, or differences in the amount of experience with the task elements (Cerella & Hale, 1994; Kail & Miller, 2006). These general neural network or experience differences clearly contrast with the grammar-specific proposals of the linguistic accounts.

Evidence supporting the generalized slowing hypothesis comes from studies of children, adolescents, and adults with SLI (Miller, Kail, Leonard, & Tomblin, 2001; Miller et al., 2006; Miller & Poll, 2009). These studies demonstrate that response times of individuals with SLI are slower by a common proportion across domains. Tasks that are more complex, and involve more operations, result in a greater gap in response time between language ability groups. Other research suggests that response times may differ by domain (Montgomery, 2006). Studies of the association between processing speed and language abilities have had mixed findings. Some studies have not found direct relationships (Lahey, Edwards, & Munson, 2001) while others do, specifically between speed of language processing and sentence comprehension (Miller & Poll, 2009). For typical children, requiring them to perform a task at a faster rate results in a pattern of errors similar to that found in SLI (Hayiou-Thomas, Bishop, & Plunkett, 2004). Children with typical language were presented with speeded sentences, and were asked to judge their grammaticality. The pattern of errors was similar to that made by children with SLI at normal presentation rates, supporting the generalized slowing theory.

Slower processing has been proposed as an explanation for sentence production difficulties in SLI (Bishop, 1994; L. B. Leonard, 1998). Sentence production can be viewed as a series of operations. Operations include selection of concepts to express, selection of lexical items to communicate the concepts, ordering of lexical constituents and encoding the
phonological elements of the message (Bock & Levelt, 1994; Levelt, 1989). Error free speech production depends on rapid processing of each operation. If the processing of these linguistic operations is slowed, more errors would be expected as the demands of production increase. Evidence supporting the role of processing speed deficits in sentence production errors is limited.

In samples of spontaneous speech from children with SLI, inconsistent patterns of grammatical errors were observed (Bishop, 1994). Bishop proposed that the pattern was more consistent with slower processing than with gaps in grammatical knowledge. She found evidence for some children that longer, more complex sentences, which required more steps, resulted in more errors. This pattern was consistent with a processing speed limitation, but other children with SLI did not follow this pattern.

In typical adults, there is indirect evidence that processing speed may affect sentence production. In one study, adults listened to and imitated sentences and unrelated word lists (Allen & Baddeley, 2009). Participants also performed a visual task, indicating the location of a visual stimulus. The visual task was performed both alone and while participants listened to and repeated sentences and word lists. The response times on the visual task were slowed to a much greater degree in the production phase of the sentence repetition task than the listening phase. The authors suggested that comprehending sentences was relatively automatic. Production, however, appeared to be much more demanding of processing resources, which may suggest that production accuracy is affected by individual variation in speed of processing. A study in which typical adults completed sentences found that the addition of a concurrent processing load resulted in higher error rates for typical adults (Hartsuiker & Barkhuysen, 2006). The evidence from these studies may suggest that reduced speed of processing will independently affect sentence production accuracy in adults, or it may suggest that there is a trade-off of task difficulty and working memory capacity (Just & Carpenter, 1992).
**Working memory limitations**

A second line of research on information processing limitations in individuals with SLI focuses on working memory (Ellis Weismer et al., 1999; Montgomery, 1995). Working memory is the capacity to hold information in mind so that it can be immediately available for a cognitive functions (Cowan, 2010). For example, when a sentence is heard in a sentence imitation task, the meaning of the presented sentence must be maintained while it is reformulated (Bley-Vroman & Chaudron, 1994). When the demands of a task exceed working memory capacity, performance is degraded either in terms of the speed or accuracy of processing (Just & Carpenter, 1992). Despite the inter-related nature of working memory and processing speed, these factors are likely separable components, each independently affecting language ability in children with SLI (L. B. Leonard et al., 2007). Limitations in working memory capacity for individuals with SLI appear to be general, not limited to linguistic or auditory processing. Children with SLI have poorer recall for material presented visually (Gillam, Cowan, & Marler, 1998). Furthermore, children with SLI perform worse than typical peers in tasks presented in both the visual and auditory modalities (Evans, Selinger, & Pollack, 2011).

There are multiple theories on the structure of working memory (Cowan, 2010). Two that have contributed to understanding the limitations of children with SLI are the multi-component model of Baddeley (2000) and the storage and processing concept of Just and Carpenter (1992). Baddeley’s (2000) theory suggests that working memory has multiple components. A central executive component allocates attention and processing resources, and three sub-components are dedicated to the short-term storage of visual, episodic, and phonological information respectively. This last component, the phonological loop, enables storage and rehearsal of auditory information. Research on children with SLI has suggested that their primary deficit is an unusual limitation of the phonological loop (Baddeley, Gathercole, &
Papagno, 1998; Gathercole & Baddeley, 1990). The capacity of the phonological loop has been measured by non-word repetition tasks. Findings of poorer non-word repetition abilities in children with SLI as compared to typical language peers have been widely replicated (Graf Estes, Evans, & Else-Quest, 2007), suggesting that children with SLI have a more limited ability to store and rehearse the auditory signal of language.

The conception of working memory proposed by Just and Carpenter (1992) was originally focused on the task of language comprehension. To successfully derive the meaning from sentences, the listener or reader must simultaneously process and store information. Initial sound and words must be interpreted, and sentence constituents must be linked together to derive meaning. During sentence interpretation, there are preliminary results of these activities which must be held in mind to be linked to later components of the sentence. Both of these aspects of working memory, storage and processing, are thought to draw on a single pool of resources according to the Just and Carpenter model. Tasks designed to measure working memory according to the Just and Carpenter model have explicit storage and processing elements. The Competing Language Processing Task (CLPT) (Gaulin & Campbell, 1994), for example, requires participants to judge the truth value of a series of sentences (the processing component), then recall the final word of each sentence in the series (the storage component). Children with SLI have consistently been shown to have more limited working memory capacity than their typical language peers on storage and processing working memory tasks (Archibald & Gathercole, 2006; Ellis Weismer et al., 1999; Montgomery & Evans, 2009). Furthermore, performance on storage and processing measures for children with SLI is correlated with broad measures of language ability, and to the ability to comprehend sentences (L. B. Leonard et al., 2007; Montgomery & Evans, 2009).

The compelling evidence that children with SLI have deficits in working memory according to both the multicomponent (Baddeley, 2000) and storage and processing (Just &
Carpenter, 1992) theories leads to issues regarding the interpretation of these findings. The first issue is exactly what elements of working memory ability are impaired in SLI. Does poor performance on a task designed to measure the phonological loop indicate poor storage capacity, poor rehearsal ability, or both? Similarly, poor performance on a storage and processing measure may indicate limited storage capacity, limited processing capacity, or both (Cowan et al., 2005). To address this issue, Cowan et al. (2005) proposed tasks that measure the scope of attention.

The scope of attention concept (Cowan et al., 2005) is distinct from the control of attention represented by the central executive in Baddeley’s (2000) model. Whereas Baddeley’s model emphasizes executive function, the ability to maintain focus on a goal or switch the focus of attention, Cowan and colleagues (2005) emphasize the amount of information that can be maintained in the active focus of attention. They view this capacity as the number of items available for immediate recall and use in cognitive processes. Measures of the scope of attention suppress rehearsal strategies.

Rehearsal is a mechanism for maintaining items in memory, and is part of the phonological loop in Baddeley’s (2000) model. The dissociation of rehearsal and storage mechanisms for verbal information is supported by studies indicating each activity engages different neural structures (Awh et al., 1996). Rehearsal engages primarily frontal areas also shown to support speech production whereas storage in the absence of rehearsal engages bilateral parietal regions. A measure of the scope of attention, such as Running Span (Cowan, et al., 2005), requires the participant to listen to lists of digits. After the list ends, the participant is asked to recall the last 5, 6, or 7 digits in the list. The lists vary in length, so the participant is unable to predict the end, and unable to rehearse because they must constantly update which digits were heard most recently. Measuring performance on scope of attention tasks promises to clarify whether storage capacity, independent of rehearsal, differs for adults with and without SLI.
Unlike the Just and Carpenter (1992) conception of working memory, scope of attention measures have the goal of measuring capacity in terms of the number of information chunks, rather than the integrated ability to both process and store information (Cowan, et al., 2005). In Running Span, for example, the goal is to recall as many list-final digits as possible in the order they were presented. There is no separate processing element for the task. Measuring the scope of attention in adults with SLI could indicate whether the capacity of active attention is limited independent of a processing load, and whether individual differences in that storage capacity covary with language ability.

A second issue raised for the Just and Carpenter (1992) account of working memory is whether working memory capacity is unitary. An alternative view is that working memory mechanisms within individuals may differ for interpretive processing tasks as compared to post-interpretive processing tasks (Caplan & Waters, 1999). In a critique of the Just and Carpenter (1992) model of working memory, Caplan and Waters argued that the effects of working memory capacity on sentence processing had not been supported for online sentence comprehension—the interpretation of a sentence as it unfolds. They agreed that working memory limitations do have an effect on tasks involving the use of information gleaned from sentences after the immediate interpretation, the post-interpretive phase. Evidence for the interpretive versus post-interpretive distinction comes from studies of typical adults and adults with neurological injuries. The adults with neurological injuries were shown to have working memory limitations, but they did not differ from typical participants in their immediate comprehension of sentences of varied syntactic complexity (Caplan & Waters, 1999). It is not clear if the distinction between interpretive and post-interpretive working memory holds true for adults with SLI.

Information processing theories of SLI assert that grammatical processing is not independent of other cognitive abilities, and that language deficits result from domain-general factors such as processing speed or working memory. The Procedural Deficit Hypothesis (PDH)
(Ullman & Pierpont, 2005) has attempted to account for both linguistic and information processing limitations in SLI with one explanatory framework. The PDH may provide an integrated account of deficits in both sentence comprehension and sentence imitation among adults with SLI.

The Procedural Deficit Hypothesis

The PDH is based on studies demonstrating dissociations between procedural and declarative memory systems (Ullman & Pierpont, 2005). Procedural and declarative memory are components of long-term memory each thought to have a distinct neural basis (Lee, 2004). Procedural memory is an implicit memory system, meaning that learning takes place automatically, or unconsciously (Tulving, 1985). Procedural memory is involved in learning skills, habits, patterns and sequences over time. Examples include learning motor sequences like riding a bicycle or learning the word sequences consistent with an artificial grammar (Squire, 1992). Procedural memory is supported by a network of brain structures including the basal ganglia and their projections to the supplementary motor area and Broca’s area in the frontal cortex (Lee, 2004; Ullman & Pierpont, 2005).

Declarative memory is an explicit memory system, meaning that learning taking place as a conscious process (Tulving, 1985). Recall of facts and events depends on declarative memory. Assessment of declarative memory is by tasks that require observation of a new fact or event, followed by recall or recognition of the fact after a delay (Squire, 1992). Declarative memory is supported by multiple brain areas, but the hippocampus and temporal lobe regions are fundamental to explicit memory formation (Lee, 2004; Squire, 1992). Retrieval of lexical information from the declarative memory system based in temporal and hippocampal regions may
also involve a neural circuit including the thalamus, basal ganglia, and anterior portions of Broca’s area (Ullman, 2006).

The Paired Associate Recognition Task (PART) was developed to measure declarative memory ability (Ragland, Gur, Deutsch, Censits, & Gur, 1995). The tasks requires recognition and recall of arbitrarily associated visual patterns following a delay, in line with the tasks shown to rely on the brain areas supporting declarative memory (Squire, 1992). In the PART, participants are shown pairs of target cards and key cards with varied patterns of shapes and colors. After a two-minute delay, they are asked to identify which of four key cards were shown with each target card. Ragland and colleagues (1995) documented the validity of the PART as a declarative memory task by showing that performance on it correlated with other tests of visual recall and semantic memory, and diverged from tests of short-term memory, working memory and executive function.

The dissociation of procedural and declarative memory, and its relation to language processing is based on studies of individuals with neurological diseases or lesions (N. J. Cohen & Squire, 1980; Knowlton, Mangels, & Squire, 1996; Lieberman, 2002; Lieberman et al., 1992; Reber, Knowlton, & Squire, 1996). Adults with amnesia are able to learn and improve on skills requiring procedural learning, such as learning to read words presented in mirror image. They are much less able to recall facts, such as the words that were repeatedly used in the task (N. J. Cohen & Squire, 1980). Adults with amnesia have injuries to brain areas associated with the declarative memory system (Knowlton et al., 1996). Adults with Parkinson’s disease demonstrate the opposite pattern. In Parkinson’s, brain areas associated with declarative memory are thought to be intact, whereas there is degeneration in brain areas associated with procedural memory, the basal ganglia. In the Knowlton et al. (1996) study, adults with amnesia, Parkinson’s disease, and typical controls performed a probabilistic classification task (the Weather Prediction Task) in which they predicted “rain” or “shine” using a set of four cards as cues. Each cue was associated
with a weather outcome at different levels of probability. The task tested gradual procedural learning over the course of 50 trials in five blocks. The performance of adults with Parkinson’s was worse than that of adults with amnesia or controls, and their performance did not exceed chance by the fifth block. Adults with amnesia and controls improved significantly over the five blocks. Adults with amnesia, however, performed much worse than adults with Parkinson’s and controls on a multiple choice test requiring recall of facts about the Weather Prediction Task. These and other studies established that injury to hippocampal regions results in deficits in declarative memory, but spares procedural memory, whereas degeneration of the basal ganglia impairs procedural memory but spares declarative memory.

The Procedural Deficit Hypothesis (PDH) suggests that linguistic, working memory, and subtle motor deficits in SLI can be attributed to a deficit in procedural memory (Ullman & Pierpont, 2005). In contrast, declarative memory is proposed to be normal, or potentially enhanced through its use as a compensatory mechanism. The PDH rests on the declarative-procedural model of language, which views language processing as taking dual routes, one for rule-governed, grammatical processing, and a second for arbitrary associations supporting lexical processing (Ullman, 2001). In this regard the PDH is consistent with the linguistic accounts of SLI. Procedural memory is the primary system for learning and processing the rule-governed, hierarchical aspects of language, whereas declarative memory is the primary system for learning and processing the lexical aspects of language, relying on associative memory.

With deficiencies in procedural memory, the PDH predicts that individuals with SLI may compensate by using declarative memory systems in situations where typical language individuals would rely on procedural memory systems (Ullman & Pierpont, 2005). The declarative-procedural model predicts frequency effects for tasks carried out by associative memory (Ullman, 2001). Associative memory formation is a function of declarative memory (Tendolkar et al., 2007). Associations are expected to be stronger with more frequent exposure to
regularities in language. No frequency effects are predicted for tasks carried out by the procedural system. The accuracy or speed of carrying out a rule-governed action, such as the addition of –ed to a verb stem to form regular past tense, is expected be independent of the frequency of the given verb. Taken together, differences are expected for syntactic tasks carried out by individuals with SLI as compared to those with typical language. The typical group should have frequency effects for lexical operations, and no frequency effects for syntactic operations. Those with SLI, however, are predicted to compensate by use of the declarative system for both lexical and syntactic operations, resulting in frequency effects in syntactic tasks as well as lexical tasks.

The findings of procedural-declarative dissociations in the neuropsychological literature provide a neurological basis for the PDH (Ullman & Pierpont, 2005). If people with SLI have a procedural deficit, then there should be evidence of differences in brain circuits supporting the procedural memory in individuals with SLI. The PDH would then advance the understanding of SLI by providing a coherent account of evidence at the biological, cognitive, and behavioral levels (Morton, 2004), something few other theories of SLI have done (Thomas, 2005). Ullman and Pierpont (2005) supported the PDH with an extensive review of findings at all three levels, but few studies have provided independent evidence supporting the novel predictions of the PDH. I will next evaluate biological, cognitive, and behavioral evidence for the PDH.

**Biological evidence for the Procedural Deficit Hypothesis**

Ullman and Pierpont (2005) presented an array of findings from neuroimaging and electrophysiological studies supporting their hypothesis that brain circuits supporting procedural memory are impaired in individuals with SLI. Much of the evidence is based on studies of the KE family. The genetic mutation of gene **FOXP2** and its association with language was
identified through studies of the KE family, but the type of language disorder resulting from the mutation present in this family is very rare (Lai, Gerrelli, Monaco, Fisher, & Copp, 2003). As a result, neuroimaging findings based on the KE family may not be representative of the much larger population of individuals with SLI.

Basal ganglia involvement is a critical feature of procedural memory engagement (Knowlton et al., 1996; Lee, 2004). Ullman and Pierpont (2005) cited two neuroimaging studies not based on the KE family that supported differences in basal ganglia neuroanatomy for individuals with SLI. One was a group study based on 12 children with language and learning disabilities and a control group (Jernigan, Hesselink, Sowell, & Tallal, 1991). This study found evidence for reduced volume of right hemisphere basal ganglia structures. The study also found clearer evidence for reduced volumes in the left prefrontal areas for children with SLI. The second neuroimaging study was a case report of a subject from the group study reported in the Jernigan et al. (1991) paper just described. This subject was found to have bilateral lesions in the heads of the caudate nuclei, basal ganglia structures. Behaviorally, the child was reported as having mild difficulties with rapid motor movements, emotional disturbances, speech difficulties, and language functioning significantly below that of the control group of the study. None of the other study participants was found to have similar lesions.

Beyond the literature cited by Ullman and Pierpont (2005), evidence in support of basal ganglia differences in children with SLI includes a study of whole brain volumes involving children with and without SLI (Herbert et al., 2003). The primary finding was of larger total brain and larger white matter volumes for children with SLI as compared to controls. After adjusting for total brain size, the volume of the caudate nucleus was smaller for the group with SLI. The authors of this study viewed their findings on white matter volumes as more supportive of information processing accounts of SLI than linguistic accounts. A more recent study of children with SLI found reduced blood flow to selected basal ganglia areas primarily in the right
hemisphere in children with SLI (Hwang et al., 2006). In the Hwang and colleagues study, the researchers administered sedatives to the participants during scanning. It is unclear what result that might have had on resting blood flow to basal ganglia regions. Furthermore, this study had a control group with a mean age of 11 years, as compared to a group with SLI with a mean age of 4 years. Both of these factors urge caution in interpreting the findings of the Hwang and colleagues study.

Considered more broadly, the neuroimaging findings on individuals with SLI are varied and the support for the proposals of the PDH is tentative. Findings that are not easily reconciled with the PDH include evidence of a greater risk for language impairment for children with smaller auditory cortices, and less leftward asymmetry in superior temporal regions (C. Leonard, Eckert, Given, Berninger, & Eden, 2006). These findings are at odds with the view of the PDH in that these brain regions are not associated with the procedural memory system (Ullman & Pierpont, 2005). Other studies have also found a tendency for the planum temporale, the region of the superior temporal lobe posterior to the primary auditory cortex (Nolte, 2009), to be less asymmetric in children with SLI (Gauger, Lombardino, & Leonard, 1997). This finding is again at odds with the position of the PDH that brain abnormalities should be found in basal ganglia and inferior frontal regions. An alternative is that the reduced asymmetry in these regions results from a compensatory shift from procedural to declarative memory processing for grammatical processing (Ullman & Pierpont, 2005), but this speculation has not been directly tested. Other findings, also concerning the reduced lateralization of brain regions for individuals with SLI, have identified less leftward asymmetry for components of Broca’s area, a finding more easily reconciled with the PDH (de Guibert et al., 2011). Broca’s area has been clearly linked to the procedural network, so a reduced cortical volume in this area is consistent with a procedural deficit.
Taken together, there is limited support for the PDH at the biological level. Much of the evidence is based on studies of individuals not representative of the larger population with SLI. Studies comparing children more representative of the SLI population to their typical language peers find some differences for structures thought to support procedural memory function. On the other hand, there are also findings of neuroanatomical differences in children with SLI that are difficult to reconcile with the PDH, and which may be more consistent with SLI as a general information processing deficit.

**Cognitive evidence for the Procedural Deficit Hypothesis**

Ullman and Pierpont (2005) reviewed findings on working memory deficits in SLI. They proposed that procedural memory circuits may play a particularly important role in the rehearsal functions involved in working memory. They also suggested that procedural circuits are involved in rapid processing of sequential information. Working memory and processing speed deficits are well documented in the cognitive profile of SLI. The novel component of the PDH is the assertion that SLI involves procedural memory deficits and intact declarative memory. Recent research assessing the cognitive ability predictions of the PDH has focused primarily on procedural, and to a lesser extent, declarative memory tasks.

A number of studies have found that children and adolescents with SLI perform significantly more poorly on visual procedural memory tasks than do typical peers (Kemeny & Lukacs, 2010; Lum, Gelgic, & Conti-Ramsden, 2010; Tomblin et al., 2007). A group of typical adults and children age 11 with and without SLI performed the Weather Prediction Task, a measure of probabilistic classification thought to tap procedural learning (Kemeny & Lukacs, 2010; Knowlton et al., 1996). Children with SLI did not exceed chance on the first 50 trials of
the task, and performed more poorly than typical peers. Adults performed better than typical children. Children with SLI did not show evidence of learning until trials 100-150.

Adolescents and children with SLI have also performed more poorly on serial response time (SRT) tasks that tap procedural memory (Lum et al., 2010; Tomblin et al., 2007). The SRT task required participants to press a button corresponding to the location of a creature appearing in different locations on a computer screen. Some sequences were random, and others followed a pattern. The group with SLI was slower to show a response time decline than the typical group during the patterned blocks. The group with SLI showed learning, but improved at a slower rate. Learning rate was correlated with grammar weakness, but not with vocabulary weakness. The authors concluded that adolescents with SLI differed in their ability to implicitly learn sequential patterns, and that the procedural memory system was likely implicated (Tomblin et al., 2007).

Poorer SRT task performance on patterned blocks has been replicated with 8 year-old children with SLI (Lum et al., 2010).

Adults with SLI have also shown weaker implicit learning of sequential patterns (Plante et al., 2002). The task for the study focused on learning an artificial grammar. A brief exposure to an artificial language was followed by probe sentences. Participants were to judge whether the probe sentences were “good” sentences—allowed by the artificial grammar. Typical adults were better than chance, and significantly better than the group with SLI, which did not exceed chance performance. This study demonstrated that difficulty with implicit learning could continue into adulthood, and that such difficulties could bear on language abilities, particularly those related to sentence processing.

Other recent evidence on procedural and declarative memory performance has not been consistent with the PDH. Some studies have reported findings contrary to the predictions of the PDH (Gabriel, Maillart, Guillaume, Stefaniak, & Meulemans, 2011; Lum et al., 2010). One study developed an SRT task with a probabilistic element. As in other SRT tasks, participants indicated
as quickly as possible where a figure appeared on a computer screen. When the figure appeared, participants touched the screen to indicate the location. The sequences had a pattern, but elements of the pattern occurred with less than 100% probability. The authors contended that this better reflected the nature of sequential pattern learning in language: sequential patterns in language are present, but are not fully consistent (Gabriel et al., 2011). There was no group difference in the performance change between the patterned blocks and the block with a random sequence. Children with SLI learned the sequential pattern as well as the typical children.

The few studies of declarative memory performance in children with SLI have had mixed results (Lum et al., 2010). In one study, eight-year-old children, with and without SLI, performed both visual and verbal declarative memory tasks. The verbal task involved recall of unrelated word pairs. After controlling for vocabulary, non-verbal intelligence, and non-word repetition performance, the children with SLI performed more poorly on the verbal declarative memory task than the typical language children. The effect size was large ($r^2 = .36$). On the visual declarative memory task, the same children tried to learn the association between a target pattern and its location on a grid. After controlling for non-verbal intelligence, there was no difference in performance between the group with SLI and the group with typical language, and the effect size was small. Without controlling for non-verbal intelligence, the effect size for the performance difference was medium ($d = .56$). Findings of verbal declarative memory deficits, even after adjusting for a range of related cognitive differences, raise questions about the predictions of the PDH that declarative memory is normal in individuals with SLI. Clearly, the evidence is limited and requires additional studies with a range of participants and declarative memory tasks.

A final difficulty for the predictions of the PDH with respect to cognitive abilities is that not all implicit learning abilities are tied to grammatical ability. Children and early adolescents with and without SLI were exposed to a continuous auditory stream made up of artificial words (Evans et al., 2009). The goal was to learn to identify “words” based on only the transitional
probabilities. The likelihood of one sound following another was higher within words than between words. This is a task that infants have performed at above chance levels, and is thought to be an important skill for word learning (Aslin, Saffran, & Newport, 1998). Children with SLI performed more poorly than children with typical language. Furthermore, children with SLI required considerably more exposure to the training condition to perform above chance than did typical children. An implication of the study is that word learning is not purely reliant on declarative memory—there are implicit, procedural learning skills at work as well.

Research assessing the cognitive ability predictions of the PDH has focused on procedural memory. Much, but not all, of the recent research has shown that children, adolescents and adults with SLI tend to lag typical peers in procedural learning tasks (Kemeny & Lukacs, 2010; Lum et al., 2010; Plante et al., 2002; Tomblin et al., 2007). The scant evidence on declarative memory performance, and comparisons of procedural and declarative task performance in the same individuals with SLI have had mixed results for the PDH (Lum et al., 2010). Children with SLI may also have impaired declarative memory, particularly for verbal material. Their performance may be better for visual declarative memory tasks, but only after controlling for general non-verbal ability. Finally, children with SLI perform more poorly in implicit learning tasks related to word learning (Evans et al., 2009), a fact not accounted for by the PDH. Comparative of performance on declarative and procedural tasks in adults with SLI has not been studied.

**Behavioral evidence for the PDH: Argumenthood and sentence processing**

A fundamental prediction of the PDH is that tasks involving rule-governed, grammatical processing should be impaired in SLI, whereas tasks relying on lexical knowledge should be normal (Ullman & Pierpont, 2005). For sentence processing, lexical knowledge includes
knowledge of argument structure. Argument structure is part of the lexicon, the repository of arbitrary associations of words and referents, not derivable by any rule of language (Chomsky, 1970; Ullman, 2001). The arguments of a word are implied by its meaning. They include components of the sentence, subjects and objects, that provide the context for the word (Radford, 2004). The verb hit, for example, implies someone to do the hitting, and an object to be hit. Receptive tasks involving such lexical knowledge are predicted to be performed normally by individuals with SLI. Such tasks rely on declarative memory without engagement of the elements of the procedural system involved in word retrieval, according to the PDH. Word retrieval engages frontal-basal ganglia circuits associated with procedural memory (Ullman, 2006).

Combining elements of sentences without the lexical support of argument structure should be more difficult for individuals with SLI (Ullman & Pierpont, 2005). Adjuncts may be such elements. Adjuncts are peripheral to the meaning of the word they modify, and so are not part of the lexical entry of the modified word (Quirk, Greenbaum, Leech, & Svartvik, 1985; Radford, 2004). Adjuncts are optional modifiers the further specify an action or event. For the verb hit, in the morning is an adjunct phrase that specifies the timing of the action, but is not central to the meaning of the verb. In the absence of this lexicalized meaning relation, deriving the meaning of an adjunct constituent in a sentence involves combining the adjunct phrase with the modified constituent—a merge operation in the terminology of generative syntactic theory (Radford, 2004). This distinction of arguments as central to the meaning of a word, and adjuncts as peripheral to the meaning is one made across multiple theories of grammar (Quirk et al., 1985; Radford, 2004; Thompson, 1997). The view of the PDH is that combinatorial operations in language, such as adjunct processing, are supported by the procedural system, and are therefore expected to present difficulties to individuals with SLI.

Ullman and Pierpont (2005) reviewed studies of tasks that required combining elements of language. Many of these tasks were more difficult for children with SLI than for their typical
language peers. Tasks included formation of questions, understanding passive sentences, and appropriately marking tense and agreement. Difficulties with adjunct production for children with SLI are also among these results. Difficulties with adjuncts were viewed by Ullman and Pierpont as consistent with a procedural memory deficit. The typical-range performance on tasks requiring knowledge of argument structure was viewed as consistent with the presence of normal declarative memory ability in children with SLI.

The evidence that individuals with SLI process arguments normally and have difficulties with adjuncts, as predicted by the PDH, is limited. Several studies support the view that argument processing is typical in children with SLI (Grela & Leonard, 2000; Thordardottir & Ellis Weismer, 2002). In this context, typical means no more impaired than younger children matched to the children with SLI at the developmental level of their sentence productions. In a study of language samples, children with SLI omitted no more obliged arguments than did mean length of utterance (MLU) matched controls (Thordardottir & Ellis Weismer, 2002). A study of auxiliary verb errors in story completions also found that more complex argument structure did not interact with group status in predicting the error rate for children with SLI and MLU matched controls (Grela & Leonard, 2000). The children with SLI, however, made a numerically higher proportion of errors in more complex argument structure conditions than did their MLU matched peers. Children with SLI in another study were no more affected by argument structure complexity in their production of past tense than were MLU-matched and age-matched peers (Owen, 2010).

Few studies have examined effects of adjunct status on the language production of individuals with SLI. Two studies have examined differences in adjunct production in language samples from children with SLI and typical controls (Fletcher & Garman, 1988; Johnston & Kamhi, 1984). In both studies, children with SLI were shown to be less likely to produce adjunct structures than their typical peers. In one study, the primary difference was a tendency to omit
information specifying the timing of an action when it was needed in the discourse context (Fletcher & Garman, 1988). In the second study, children with SLI—compared to MLU matched controls—were found to omit adjunct phrases that would provide information on time, place or manner (Johnston & Kamhi, 1984). No studies of individuals with SLI have examined the effects of adjunct status on language production in more controlled conditions such as sentence imitation. One study examined errors made on a sentence imitation task, but this study involved children with varied psychiatric conditions (Kotsopoulos, 1997). Most participants in the study were language impaired, but would be excluded from most studies of SLI. The study did find that content omission errors were more likely to occur on adjunct elements as opposed to argument elements of sentences.

The literature on children with SLI largely supports the prediction of the PDH that their production of arguments is more similar to that of typical children than is their production of adjuncts. None of the studies, however, employed a well-controlled production task, such as sentence imitation, to test the hypothesis. The hypothesis has also not been evaluated for adults with SLI. It is possible that developmental changes and experience would reduce the adjunct disadvantages observed in children. The PDH makes the clearest prediction on the processing of arguments in receptive language tasks. The prediction is that individuals with SLI should not differ from those with typical language in argument processing for a comprehension task (Ullman & Pierpont, 2005). This prediction has not been tested, nor has the effect of adjunct status in sentence comprehension been evaluated for adults with SLI. Argument versus adjunct effects have been evaluated in sentence comprehension tasks in typical adults, however. I will next review findings on how argument status affects sentence processing difficulty for typical language adults.
Argument preference in sentence processing

Many studies have explored the moment-to-moment processing of sentences in typical adults (Mitchell, 1995, 2004). Studies have primarily considered the time participants take to process components of a sentence. Longer response times are taken as evidence for greater difficulty processing a component of the sentence. For example, a sentence may be presented in writing, and the measure is how long the participant’s gaze rests on a sentence region. Or, a sentence may be presented auditorily, and the measure is how long the participant listens to the region of interest before moving to the next.

Competing theories of sentence processing have argued that processing is influenced primarily by the sentence’s syntactic structure on the one hand, or by a range of interacting influences, particularly lexical influences, on the other (van Gompel & Pickering, 2007). Structural theories contend that initial sentence processing is a rapid, automatic process guided by the degree of complexity in the syntactic structure of the sentence (Frazier, 1987). Other influences, such as lexical characteristics, play a role only in a later stage re-analysis of the sentence. Minimal Attachment, a structural theory proposed by Frazier (1987), predicts a preference for attaching new material encountered in a sentence to the existing sentence structure in the simplest way possible, using the fewest nodes in the hypothesized tree structure. Interactive theories contend that other influences may override the preference for structural simplicity (MacDonald, Pearlmutter, & Seidenberg, 1994). Lexical factors, sentence plausibility, or frequency characteristics, for example, may have a greater influence on processing time than syntactic simplicity.
Argument effects in sentence processing

The effect of argument status on sentence processing in typical adults has been evaluated from both structural and interactive perspectives (Boland & Blodgett, 2006; Clifton, Speer, & Abney, 1991; Schutze & Gibson, 1999). Clifton et al. (1991, p. 255) evaluated prepositional phrases acting as arguments or adjuncts in sentences such as the following:

A. Verb attachment, argument
   a. The saleswoman tried to interest the man in a wallet during the storewide sale at Steigers.

B. Verb attachment, adjunct
   a. The man expressed his interest in a hurry during the storewide sale at Steigers.

C. Noun attachment, argument
   a. The man expressed his interest in a wallet during the storewide sale at Steigers.

D. Noun attachment, adjunct
   a. The salesman tried to interest the man in his fifties during the storewide sale at Steigers.

A preference for simpler syntactic structure favored conditions A and B (Clifton et al., 1991). When the attachment of in a wallet or in a hurry is to a verb, the authors’ analysis indicated that the structure of the sentence contained fewer nodes, and so the prepositional phrase should be processed faster than in conditions C or D, where attachment is to a noun. The noun attachment resulted in more nodes in the syntactic structure. A preference for arguments over adjuncts favored conditions A and C. The two theories had contrasting predictions for conditions B and C, noun argument versus verb adjunct. The experiments used two different reading
methods, self-paced reading and eye tracking. In the self-paced reading experiment, typical
college students read verb-attached prepositional phrases (conditions A and B) faster, but
argument conditions resulted in faster reading times in the region following the prepositional
phrase. In the eye tracking study, first pass reading times indicated that verb-attached readings
were faster, whereas total reading times were faster in argument conditions. The authors
concluded that structural preferences played the primary role in the earliest phase of sentence
processing, but that argument preference influenced the overall comprehension process.

A subsequent study concluded that plausibility and length differences across conditions
may have influenced the results of Clifton et al. (1991) (Schutze & Gibson, 1999). In their study,
Schutze and Gibson tested the structural versus argument preference theories using materials
better controlled for plausibility and length. They included noun argument and verb adjunct
conditions only, since these conditions (as in B and C above) had contrasting predictions under
the Minimal Attachment view versus the interactive, argument preference view. As shown in
Figure 1, the verb adjunct condition had fewer nodes than the noun argument condition, and so
the prepositional phrase or subsequent regions should be processed faster according to structural
theories (Minimal Attachment). The noun argument condition maximizes argument relations. As
a result, the noun argument prepositional phrases, or subsequent regions, should be processed
faster according to an argument preference theory. Argument preference theory suggested that a
lexical factor, whether the constituent is an argument or an adjunct, influences sentence
processing time. By this theory, arguments are proposed to be preferred to non-arguments. In
self-paced reading experiments with typical adult participants, they found faster reading times in
the noun argument condition as compared to the verb adjunct condition for the region just
following the prepositional phrase, supporting the argument preference theory over Minimal
Attachment. They did not report effect sizes.
Schutze and Gibson (1999) concluded that argument relations influence sentence processing at the earliest stages of comprehension. They went on to propose that such findings could be the result of either a preference for the semantic property of argumenthood, or the fact that arguments simply co-occur with their related verbs or nouns with greater frequency than do adjuncts. The frequency of co-occurrence could be influencing reading times in both the argument and adjunct conditions (MacDonald et al., 1994). Their analysis showed that although noun arguments co-occurred with their related noun more frequently than did the verb adjuncts, there was no linear correlation with reading times. They proposed a series of tests for argumenthood showing that arguments behave differently in sentences than adjuncts, and suggested that the distinction between arguments and adjuncts goes beyond frequency effects.

Boland and Blodgett (2006) also studied argument effects for sentences in which prepositional phrases acted as argument or adjuncts of nouns or verbs. They found that typical

**Figure 1**

Sentence structures for noun argument (left) and verb adjunct (right) conditions showing that the noun argument condition has more nodes. Source: “Argumenthood and English Prepositional Phrase Attachment,” by Carson T. Schutze & Edward Gibson, 1999, *Journal of Memory and Language*, *40*, p. 413. Copyright 1999 by Academic Press.
adults processed arguments faster than adjuncts. Their eye-tracking study found that arguments were processed faster in both the first-pass reading times and total reading times. The effect was sometimes found in the prepositional phrase region, sometimes in the subsequent region of their sentences. Effect sizes were not reported. A goal of the Boland and Blodgett study was to test predictions of the Argument Structure Hypothesis and the Pure Frequency Hypothesis.

The Argument Structure Hypothesis (ASH) suggests that argumenthood is a semantic property, and that the arguments required by a word are stored as part of the word’s lexical entry (Boland & Blodgett, 2006). Adjuncts are not inherent to the meanings of the verbs or nouns they modify, and are not stored as part of the lexical entry of the modified word. The ASH predicts that when participants encounter a word, the potential arguments associated with the word are activated in the brain. This activation eases the processing of argument phrases encountered later in the sentence. Since adjuncts are not stored as part of the lexical entry, encountering a word does not activate any potential adjunct phrases. This means that processing of adjunct phrases will take longer than argument phrases, and that integrating adjunct phrases into the meaning of the sentence must rely on rule-based syntactic processing (Boland & Boehm-Jernigan, 1998).

The Pure Frequency Hypothesis (PFH) also predicts that argument phrases will be processed more easily than adjunct phrases (Boland & Blodgett, 2006). Unlike the ASH, the PFH holds that both arguments and adjuncts are stored in the mental lexicon. In fact, there is no fundamental semantic or syntactic distinction between arguments and adjuncts in the PFH. Adjuncts simply co-occur with the words they modify less frequently than do arguments (MacDonald et al., 1994).

Boland and Blodgett (2006) reasoned that the ASH and the PFH could be distinguished by analyzing the correlations between reading times and co-occurrence frequencies for the words and their arguments or adjuncts. Both the ASH and PFH predict negative correlations between processing times and co-occurrence frequencies for argument conditions. A higher co-occurrence
frequency should result in a shorter processing time. Correlations for adjunct conditions should follow the same pattern according to the PFH. For the ASH, processing of adjunct phrases follows syntactic rules so no frequency effects are predicted. Boland and Blodgett (2006) found, however, that the variance of the argument co-occurrence frequencies was larger than the variance of adjunct co-occurrence frequencies in their study. As a result, their finding that adjunct co-occurrence frequency did not correlate with reading times may have been due to the restricted range of adjunct co-occurrence frequencies in their materials rather than to a processing difference for adjuncts versus arguments.

Despite the limitations, Boland and Blodgett (2006) presented their correlation analyses and found statistically significant negative correlations in the verb argument conditions, but not the noun argument conditions. Both the ASH and PFH would have predicted significant negative correlations in both verb and noun argument conditions. No significant correlations were found between reading times and co-occurrence frequencies in adjunct conditions.

Across multiple studies, the evidence suggests that argument status eases sentence processing for typical adults when other factors are well controlled. Whether the argument effect results from a categorical distinction between arguments and adjuncts or simply from differences in co-occurrence frequencies remains unclear. What also remains unclear is the effect size of the argument advantage in sentence processing. The argument advantage has been evaluated in the context of other factors that also affect sentence processing, plausibility and verb bias. When sentences are both more plausible, and prepositional phrases are arguments, there is a clear easing of sentence processing (Speer & Clifton, 1998). But more plausible conditions are also easier to process when prepositional phrases are adjuncts. When verbs more often are followed by direct objects (transitive biased verbs), argument prepositional phrases are faster to read than adjuncts. However, when verbs more often are not followed by a direct object (intransitive biased verbs), there was no difference between argument and adjunct prepositional phrase processing
(Kennison, 2002). It appears that argument effects may be smaller than other lexical or plausibility influences on sentence processing for typical adults.

**Argument effects and compensation in adults with specific language impairment**

While sentence processing in typical adults is facilitated by argument status, it remains unclear how argument status affects sentence processing for adults with SLI. The literature on the effects of argument status on individuals with SLI has focused on children, and on children’s language production. Argument effects in online sentence comprehension have not been measured in adults with SLI. One prediction consistent with the Procedural Deficit Hypothesis (PDH) is that adults with SLI will be affected to a greater degree by adjunct status than will typical language adults. This assumes that adjuncts are processed by syntactic mechanisms, as is proposed by the Argument Structure Hypothesis (Boland & Boehm-Jernigan, 1998). If online processing of argument phrases is eased by a normal lexical/declarative memory system, and adjunct phrases are processed primarily by a deficient syntactic/procedural memory system, then a larger degree of slowing for adjunct phrases as compared to argument phrases should be expected for adults with SLI. If this is not the case, and the processing of adjuncts and arguments in adults with SLI is similar to that of adults with typical language, then adults with SLI may compensate by using their declarative memory systems to process adjunct phrases (Thomas, 2005; Ullman & Pierpont, 2005).

Evidence of compensatory processing should be expected from two sources (Thomas, 2005). First, use of the declarative memory system should result in frequency effects (Ullman, 2001). For argument and adjunct processing, correlations should be expected between the processing times for adjunct phrases and co-occurrence frequencies for adjunct prepositional phrases and the words they modify. Frequency effects for adjuncts would also be consistent with
the prediction of the Pure Frequency Hypothesis for typical language adults. Support for the Argument Structure Hypothesis in typical adults in the presence of frequency effects for adjunct processing for adults with SLI would provide the clearest evidence for the use of compensatory processes predicted by Ullman and Pierpont (2005).

A second prediction is that if adults with SLI achieve normal performance by relying on compensation from their declarative memory, then their declarative memory systems should be normal (Thomas, 2005). This is consistent with Ullman and Pierpont’s (2005) prediction that the declarative memory systems of individuals with SLI could be normal or superior to typical language individuals, based on the extensive use of the declarative system for compensatory processing. To validate this logic, a test of declarative memory ability independent of the sentence processing task is required. Examples of such tasks would be arbitrary word association recall, or a visual paired-associate recall task relying on declarative memory. Adults with SLI should exhibit normal declarative memory ability, according to these predictions.

To recap, evidence for argument processing advantages in sentence processing suggests that encountering a word in a sentence activates the argument structure of the word (Boland & Blodgett, 2006). This activation facilitates the processing of arguments of the word encountered later in the sentence. Assuming that information about adjuncts is not stored as part of the lexical entry of the word, as predicted by the Argument Structure Hypothesis, then integrating adjuncts into the rest of the sentence will rely on syntactic rules (Boland & Boehm-Jernigan, 1998). Based on the PDH (Ullman & Pierpont, 2005), if adults with SLI use normal declarative systems for lexical processing and a deficient procedural systems for syntactic processing, then there should be a difference between their argument and adjunct processing times. This difference is predicted to be larger for adults with SLI than for adults with typical language. If there is no group difference in argument and adjunct processing times, the PDH would suggest that the normal performance of the group with SLI relied on compensatory engagement of their normal
declarative memory systems (Thomas, 2005; Ullman & Pierpont, 2005). Frequency effects for adjunct processing not found for the typical group, and normal performance on declarative memory tasks would provide evidence for such compensation.

If an argument advantage is found, but correlational analyses support the Pure Frequency Hypothesis, this would suggest that both argument and adjunct information is part of the lexical entry (Boland & Blodgett, 2006; MacDonald et al., 1994). Under this scenario, the predictions of the PDH are unclear. The PDH assumes a division of language processing into rule-based (syntactic) and association-memory based (lexical) elements (Ullman, 2001). The predictions of the PDH may not hold for a language system that does not honor this distinction. Any prediction of performance for adults with SLI would need to consider the results in light of theoretical perspectives that assume a unitary language processing system.

Evaluating the PDH by observing effects of argumenthood in online sentence processing requires a research method that eliminates the effects of potential differences between adults with and without SLI that are unrelated to sentence processing. One such difference is word reading ability. There is considerable overlap between SLI and reading disorders (Catts, Fey, Tomblin, & Zhang, 2002). The methods to study argument effects in sentence processing in typical language adults have required reading. For adults with SLI, a method that does not require reading would be a better approach. Self-paced listening is one such approach.

**Self-paced listening**

The auditory moving window, or self-paced listening task was designed as an auditory analogue to the self-paced reading task (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996). The purpose was to measure processing times as sentences unfold. The task involves listening to sentences word-by-word. The participant hears a word to begin a sentence, and then presses a
button to advance to the next word. The process continues until the sentence ends. A computer tracks the listening times from word to word. Sentence structures that are more challenging to process are expected to require more listening time. The time elapsed from the button press to initiate a sentence region to the button press to initiate the following sentence region is an index of that region’s processing difficulty. The self-paced listening technique has been shown to produce results consistent with self-paced reading methods with typical adults (Ferreira, Henderson, et al., 1996) and typical children (Booth, MacWhinney, & Harasaki, 2000). It has also been used effectively with adults who have neurogenic language disorders (Caplan & Waters, 2003). There is one reported use of the self-paced listening technique to investigate sentence processing in children with SLI, but the results have not been published (Marshall, Marinis, & van der Lely, 2007).

The use of an auditory rather than a reading task to measure online sentence processing introduces potential effects of prosody (Ferreira, Anes, & Horine, 1996). An advantage of the technique is that it can be used to assess the effects of prosody, but in the present study prosody is not under consideration. To eliminate influences of prosody and coarticulation, one research team recorded sentence materials word by word from a randomly ordered list (Booth et al., 2000). The sentences were then re-assembled from separate sound files at the time of presentation. This method was expected to control any potential differences in prosody across conditions.

A further issue is how to account for the varied durations of sound files making up regions of interest. If, for example, the direct object noun is a region of interest, there will be some variation in the time required to play the sound files of the direct objects across sentences. One approach to address this variation is to subtract the duration of the sound files for a region of interest from the total listening time for the region (Ferreira, Henderson, et al., 1996). A second approach is to account for the duration by regression modeling (Booth et al., 2000). In the regression approach, the duration of the sound segments is used as a predictor of each
participant’s time to process the region. The predicted duration from the regression is then subtracted from the actual listening duration to arrive at an adjusted listening time.

Self-paced listening is a means of evaluating the effects of argument status on sentence comprehension for adults with SLI. Argument status may also play a role in speech production as well. Differences in adjunct processing for individuals with SLI have been found primarily in speech production tasks (Fletcher & Garman, 1988; L. B. Leonard, 1998). Sentence imitation provides a means of studying the effects of argument and adjunct status in speech production while controlling for potentially confounding factors.

Sentence imitation

An advantage of sentence imitation as a method for studying language production is that it reflects that language systems of the participants while also being highly controlled. Sentence imitation reveals the language production systems of participants because it is reconstructive (Lust, Flynn, & Foley, 1996; Potter & Lombardi, 1998; Slobin & Welch, 1971). In studies of typical adults, productions elicited by sentence imitation are influenced by previously heard or previously produced lexical items and sentence structures. For example, when a sentence could convey the same meaning with two different sentence structures, participants increase use of the structure they have recently heard, even if it differs from the sentence they were asked to repeat (Lombardi & Potter, 1992; Potter & Lombardi, 1998). In children, sentence imitation has been shown to reflect their level of language development (Slobin & Welch, 1971). These findings indicate that sentences are reconstructed using the participant’s language system in sentence imitation, and do not reflect rote repetition of an auditory memory.

Sentence imitation can be decomposed into the following likely steps. First, the target sentence is heard and comprehended, and the conceptual gist of the sentence is held in short-term
memory (Bley-Vroman & Chaudron, 1994; Potter & Lombardi, 1998). The participant then engages their speech production system, and produces a sentence based on the conceptual message of the target sentence. That message is expressed by selecting appropriate lexical items, assembling them into a grammatical string, adding appropriate inflections, and phonologically encoding the result for articulation (Bley-Vroman & Chaudron, 1994; Bock & Levelt, 1994). The lexical items and sentence structures selected are influenced by the priming of words and sentence structures heard in the target sentence (Potter & Lombardi, 1998). For this model to hold and to reveal differences in the language systems of the participants, the sentences must exceed the ability of participants to simply replay the auditory trace of the sentence (more than 7-8 words), but must also not be so long that the sentences are well beyond the processing capacity of participants (Bley-Vroman & Chaudron, 1994; Lust et al., 1996).

Working memory and sentence imitation

Given the requirement to hold the gist of the target sentence in memory while reconstructing the sentence for production, it is clear that working memory plays a role in sentence imitation performance. For children and adolescents with SLI, measures of phonological short-term memory (nonword repetition) and working memory (digits backward) correlate with sentence imitation performance (Alloway, Gathercole, Willis, & Adams, 2004; Riches, Loucase, Baird, Charman, & Simonoff, 2010).

The number of words recalled in sentence form is considerably greater than the number of words recalled from unrelated word lists (Allen & Baddeley, 2009; Baddeley, Hitch, & Allen, 2009; Jefferies, Ralph, & Baddeley, 2004). The evidence suggests that the meaning and syntactic relations among words in sentences allow participants to construct larger units, or memory “chunks,” for storage in working memory. The number of chunks may not differ between word
lists and sentences, but the size of each chunk is larger in the case of sentences (Gilchrist, Cowan, & Naveh-Benjamin, 2008). The binding of words into these memory chunks involves interactions of short-term memory with long-term memory of language. In a series of experiments, this binding process has been shown to be an automatic process, not requiring significant central executive processing resources, in the terminology of the Baddeley and colleagues models of working memory (Allen & Baddeley, 2009; Baddeley et al., 2009; Jefferies et al., 2004). This suggests that cognitive processing speed, or the processing components of working memory tasks, may not be important in the comprehension and encoding phase of the sentence imitation task. It is unclear whether the subsequent formulation and production phases are demanding of processing resources.

**Syntactic structure effects in sentence imitation**

The role of long-term memory for language in the binding of memory chunks, as well as the reconstructive nature of the task suggests that sentence imitation reflects the language abilities of the participants. Complex sentence structures have been shown to be more difficult for children with SLI, and these difficulties are reflected in their poorer performance imitating sentences containing these structures (Panagos & Prelock, 1982; Redmond, 2005; Redmond et al., 2011; Riches et al., 2010). For example, sentences with non-canonical word orders and sentences with relative clause structures are repeated with more errors by children with SLI as compared to peers with typical language. In the Riches et al. (2010) study, sentence structure interacted with group assignment in the sentence imitation task: the group with SLI had a greater performance decrement for complex sentences as compared to simpler sentences than did the group with typical language. The PDH predicts that production of adjunct phrases will be particularly difficult for adults with SLI as compared to argument phrases (Ullman & Pierpont, 2005), but
interactions of SLI status and argument structure in sentence imitation performance have not been evaluated.

**Processing speed and sentence imitation**

The generalized slowing theory of SLI also has predictions relevant for sentence imitation performance (Bishop, 1994; L. B. Leonard, 1998). The theory considers the effects of reduced processing speed in light of a model of speech production (Bock & Levelt, 1994; Levelt, 1989). As outlined above, speech production is viewed as a sequence of processes. Beginning with a conceptual message, in the case of sentence imitation, the gist of the target sentence, the process continues with lexical selection. Selection of the lexical items includes not only the meanings, but also the grammatical roles those lexical items can play in sentences and what argument structure is required. The lexical items are then assigned a function in the sentence, such as subject, or direct object. The lexical items are then sequenced into phrasal structures, and then function words and inflections are added to the sequences. These inflected and sequenced structures are then submitted to phonological encoding and articulation.

The potential effect of slower processing on sentence imitation derives from the fact that speech production has timing demands. If these timing demands are not being met in the sequence of processes required for speech production, then errors reflective of a rushed or incomplete process will be the result (L. B. Leonard, 1998). For example, since the addition of inflections is a later step in the process of phrasal assembly than lexical selection, the correct word stem may be produced for the context, but without an obligatory inflection. Leonard (1998) and Bishop (1994) argued that the pattern of errors observed in the speech production of children with SLI fit the predictions of the generalized slowing hypothesis. If this hypothesis is correct, then individual differences in processing speed should explain differences in sentence imitation.
accuracy. The uncertainty is what level of difficulty for the speech production system is required to observe the effects of these processing limitations in adults with SLI.

**Research questions and experiment summaries**

This study focused on two primary issues across three experiments. The first issue was whether the predictions of the PDH are valid for adults with SLI. The predictions of the PDH were evaluated by observing the effects of adjunct versus argument status on the language performance of adults with SLI. Argument effects are specifically addressed by the PDH but have not been evaluated elsewhere in well controlled comprehension or production tasks. The study also tested the prediction of a procedural memory deficit for adults with SLI in the presence of normal declarative memory.

The second issue for this study was the limited information on what factors explain sentence imitation performance for adults with SLI. The study evaluated three factors: the working memory capacity and processing speed of the participants, and the argument status of constituents in the sentences.

The questions for this study were as follows:

1. For evaluation of the PDH:
   a. Do adults with SLI have particular difficulty processing adjuncts, with relatively typical performance processing arguments? In comprehension and production tasks, finding a group by argument status interaction in performance measures will support the predictions of the PDH.
   b. If processing of adjuncts appears to be typical in a self-paced listening task, is there evidence of compensatory involvement of the lexical/declarative memory system for the group with SLI? Frequency effects for adjunct
processing for the group with SLI, where there are no frequency effects for adjunct processing in the group with typical language, would be evidence of compensatory processing consistent with the PDH.

c. Is the procedural memory system of adults with SLI deficient, whereas the declarative memory system is normal or less impaired than the procedural system? Relatively better, or normal, performance on a visual declarative memory task and relatively weak performance versus typical language peers on a visual procedural memory task would support the hypothesis of the PDH.

2. For evaluation of factors explaining sentence imitation performance:

a. Do working memory, processing speed, and argument status each contribute to sentence imitation performance in the presence of the others?

b. To what degree does each measure explain sentence imitation performance?

c. If a working memory composite measure is a significant predictor of sentence imitation, does a scope of attention measure (Running Span) predict sentence imitation performance as well as a storage and processing measure (Competing Language Processing Task)?

To address these questions, three experiments were conducted. The first experiment is summarized in Figure 2 and is the subject of Chapter 3. The PDH was evaluated in a sentence comprehension task using the self-paced listening method. Participants listened to sentences, and their listening times for selected phrases of the sentences were measured. The phrases of interest acted as either arguments adjuncts. The PDH predicted that the group with SLI would have difficulty with adjunct phrase processing, since adjunct processing relies on syntactic rules (Boland & Boehm-Jernigan, 1998; Ullman & Pierpont, 2005). Evidence supporting this
prediction would be a significant group by argument status interaction in the model predicting listening times. Alternatively, the group with SLI may compensate and use declarative memory to support adjunct phrase processing. Evidence to support this prediction would be a significant correlation of adjunct phrase frequency and adjunct listening times for the group with SLI, where none is found for the group with typical language.

The second experiment is summarized in Figure 3 and is the subject of Chapter 4. In this experiment, participants imitated sentences. This task provided a second test of the predictions of the PDH. Here, the prediction that adjunct phrases would be particularly difficult for adults with SLI (Ullman & Pierpont, 2005) was evaluated in a language production task. Again, evidence supporting the PDH would be a group by argument status interaction in a regression model predicting sentence imitation accuracy. The second purpose of the task was to evaluate working memory capacity, speed of processing, and sentence structure (arguments versus adjuncts) as

---

**Figure 2**

The purpose, predictions and evidence required to support the predictions for the sentence comprehension task.

a Procedural Deficit Hypothesis

---
factors explaining sentence imitation accuracy. To support the hypothesis that each of these factors plays a role in sentence imitation, each would be a significant predictor in the regression model of sentence imitation accuracy. Finally, two views on the structure of working memory were evaluated. The hypothesis was that a scope of attention measure, Running Span (Cowan et al., 2005), would equal a storage and processing measure, the Competing Language Processing Task (Gaulin & Campbell, 1994), as a predictor of sentence imitation accuracy. Evidence to support this hypothesis would be a finding that Running Span was a significant predictor of sentence imitation, and that the effect size for the Running Span measure would be equivalent to the effect size of the Competing Language Processing Task.

Figure 3
The purpose, predictions and evidence required to support the predictions for the sentence imitation task.

\( ^a \) Procedural Deficit Hypothesis \( ^b \) Competing Language Processing Task
The third and final experiment is represented in Figure 4 and is the subject of Chapter 5. This experiment consisted of two tasks designed to directly measure declarative and procedural memory in primarily visual tasks. The Paired Associate Recognition Task (PART) (Ragland et al., 1995) was designed to measure declarative memory. The PDH predicts that individuals with SLI will have typical range declarative memory, or that their declarative memory would be a relative strength as compared to their procedural memory (Ullman & Pierpont, 2005). Evidence to support this prediction would be either a finding of no significant group difference for the PART, or a smaller effect size for the group difference as compared to the procedural memory task. The Weather Prediction Task (WPT) (Knowlton et al., 1996) is a measure of procedural memory. The PDH predicts that individuals with SLI have deficits in procedural memory. Evidence supporting this prediction would be significantly poorer performance for the group with SLI on the WPT, and a larger effect size for the group difference as compared to that found for the PART.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Purpose</th>
<th>Predictions</th>
<th>Evidence to Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired Associate Recognition Task</td>
<td>Test Predictions of PDH: Declarative Memory</td>
<td>SLI: Typical Range or Less Depressed Performance</td>
<td>Non-Significant Group Difference or Smaller Effect</td>
</tr>
<tr>
<td>Weather Prediction Task</td>
<td>Test Predictions of PDH: Procedural Memory</td>
<td>SLI: More Depressed Performance</td>
<td>Significant Group Difference and Larger Effect</td>
</tr>
</tbody>
</table>

**Figure 4**

The purpose, predictions and evidence required to support the predictions for the visual measures of declarative and procedural memory.

\(^a\) Procedural Deficit Hypothesis
All three experiments included the same participants. Details of participant selection will be presented in the next chapter. Chapters 3 to 5 present the findings of each experiment. A general discussion of the overall study findings and implications will follow in Chapter 6.
Chapter 3

Experiment 1: Self-Paced Listening

Participants

Forty-six young adults (33 females), age 18-27 years, participated in the study. Twenty-four of the participants were recruited at post-secondary schools in central Pennsylvania and 22 from a list of adults who were part of a longitudinal study of developmental language impairment conducted by researchers at the University of Iowa. Details on the initial identification of the Iowa participants can be found in Tomblin, Zhang, Buckwalter and O’Brien (2003). Adults were invited to participate if they spoke English as their first language, and reported no history of autism spectrum disorders, cerebral palsy, serious neurological injury, intellectual disability or hearing loss.

After providing written informed consent, all participants passed a pure tone hearing screening at 25 dB HL at 500, 1000, 2000, and 4000 Hz. Six participants tested in rooms with loud air conditioning systems passed the hearing screening at 30 dB HL at 500 Hz, and at 25 dB at all other frequencies. To measure participants’ performance intelligence quotient (PIQ), three subtests of the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III) (Wechsler, 1997) were administered: Picture Completion, Digit Symbol Coding, and Matrix Reasoning. Scores from these subtests were used to estimate PIQ using the method developed by Sattler and Ryan (1999). This method of PIQ screening was selected in order to minimize the time of administration while maintaining high reliability (.93) for the measure. Participants with PIQs of 80 or above (on a scale with mean of 100, SD of 15) were included in the data analysis for the study.
From those participants who met the screening criteria, two groups were formed according to the participant’s language ability. Characteristics of the groups are presented in Table 1. Participants were classified as language impaired by their performance on a series of language measures that have been shown to discriminate between typical and impaired language in young adults (Fidler et al., 2010). The tasks were written spelling, spoken word definitions, and following directions presented in increasingly complex sentences. The written spelling task consisted of 15 words from Fidler, Plante and Vance (2010). The examiner spoke each word, used the word in a sentence, and then repeated the word. Participants wrote the word on a response form. The same example sentences were used for all participants, and words were scored as correct or incorrect. The spelling score was the number of words spelled correctly. The word definitions task was the Word Definitions subtest from the Clinical Evaluation of Language Fundamentals, 4th edition (CELF4) (Semel et al., 2003). Participants were presented with a word and were asked to provide a spoken definition of the word. Participant responses were audio recorded. The task was presented and scored according to the directions of the CELF4. The standard score was determined according to the adult norms for the test (for age 18-21 years). To measure reliability, a randomly selected 11% of Word Definitions responses were scored by a second rater who did not have knowledge of the language group status of the participants. The second rater was an undergraduate research assistant trained on the procedures of the Word Definitions subtest. Point-to-point agreement between the first and second raters was 88%. The following directions task was the Modified Token Test (de Renzi & Faglioni, 1978; Morice & McNicol, 1985). Participants indicated their understanding of spoken directions by pointing to or manipulating a set of wooden items with common shapes and colors. There were 44 items on the test, and the score was the number correct.

Participant scores were entered into the following discriminant function from Fidler, Plante and Vance (2010):
6.6626 + (-0.2288 * Spelling) + (-0.1475 * Word Definitions) + (-0.0893 * Token) = 

Group Membership Value

Positive Group Membership Values indicated language impairment; negative values indicated typical language. Participants with positive scores on this function and meeting other screening criteria were classified as specifically language impaired. Participants with negative histories of language difficulties and negative results from the discriminant function were classified as typical language. Participants reporting a history of language-related difficulties who had negative discriminant function results were excluded. Participants reporting histories of depression and anxiety disorders were included in both SLI and typical groups. Eleven participants were enrolled in the study but excluded from the analyses. Four of these participants scored below the PIQ criterion for the study, 2 did not pass the hearing screening, and 1 had history of language difficulties but scored in the negative range on the language classification discriminant function. One participant was excluded due to technical problems during administration of experimental tasks, another because they were medicated in a way that affected results during the experimental session, and two for not meeting the comprehension criteria for the self-paced listening task. Details of the self-paced listening task are described below.

The final language groups each had 23 participants, of whom 14 were female in the SLI group, and 19 were female in the typical language group. For the group with SLI, 9 were from Pennsylvania and 14 were from Iowa. In the typical group, 15 were from Pennsylvania and 8 were from Iowa. The language groups did not differ in age ($t(44) = 1.47, p = .149$), but did differ on years of education ($t(44) = 4.62, p < .001$) and PIQ ($t(44) = 6.05, p < .001$). Nineteen of the SLI group participants had a history of language difficulties, such as a prior diagnosis of language impairment, reading difficulties, or special education related to language difficulties; none of the typical language group reported a history of language difficulties.
Table 1
Mean (SD) demographic, performance intelligence and language measures for participants.

<table>
<thead>
<tr>
<th>Measure</th>
<th>SLI</th>
<th>Typical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>22.3 (2.0)</td>
<td>21.5 (1.8)</td>
</tr>
<tr>
<td>Years of Education</td>
<td>13.2 (1.1)</td>
<td>14.5 (0.9)</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>97.8 (7.7)</td>
<td>113.7 (10)</td>
</tr>
<tr>
<td>Token Test % Correct</td>
<td>70.5 (16.3)</td>
<td>91.3 (5.1)</td>
</tr>
<tr>
<td>Spelling (number correct of 15)</td>
<td>4.0 (2.3)</td>
<td>11.4 (1.9)</td>
</tr>
<tr>
<td>Word Definitions, Standard Score</td>
<td>8.0 (3.3)</td>
<td>13.1 (1.5)</td>
</tr>
</tbody>
</table>

Materials for the self-paced listening task

Sixteen sets of sentences were developed to assess differences in the difficulty of processing adjuncts as compared to arguments. Processing difficulty was indexed by listening time. The materials were based on a prior self-paced reading study focused on the same question (Boland & Blodgett, 2006). Given the high co-morbidity of reading disorders among individuals with SLI, the materials were presented auditorily using the self-paced listening paradigm (Ferreira, Henderson, et al., 1996). There were four conditions in each set of sentences. Two conditions varied on whether a prepositional phrase was an argument or adjunct of the main verb in the sentence, and two conditions varied on whether the prepositional phrase was an argument or adjunct of the direct object noun. The entire final set of materials is presented in Appendix A. Examples of each condition follow:
The police detective…

a. …proposed | an investigation | to the captain | right away. (verb argument)

b. …proposed | an investigation | of the company | right away. (noun argument)

c. …proposed | an investigation | for a day | right away. (verb adjunct)

d. …proposed | an investigation | through the city | right away. (noun adjunct)

The italicized portions of the sentences are the prepositional phrases and constitute the critical region of each sentence. In order to rule out listening time differences not related to argument status, the sentences were balanced across conditions on word frequency, word familiarity, and sentence plausibility. To determine if argument versus adjunct status is a categorical difference or primarily a frequency difference, the co-occurrence frequencies of the verbs or direct object nouns with prepositional phrases similar to those found in the materials were assessed. In Boland and Blodgett’s (2006) original materials, the variance of co-occurrence frequencies differed across argument and adjunct conditions. The materials developed for this study were modified so that the variation in co-occurrence frequency in argument and adjunct conditions did not statistically differ. Development of materials meeting this condition involved as series of corpus analyses and norming studies.

**Corpus analyses**

A language corpus representative of the language environments of the participants was selected for this study (McEnery, Xiao, & Tono, 2006). That corpus was the Corpus of Contemporary American English (COCA) (Davies, 2009). The spoken language corpora in the COCA were used, and these contained more than 79 million words based on unscripted conversation from popular television and radio programs in the United States.
The primary analyses focused on how often a prepositional phrase like that found in the self-paced listening materials followed either the verb or direct object noun used in the materials. These procedures largely followed those of Boland and Blodgett (2006). First, the lemma frequency of the relevant verb or direct object noun was assessed. A lemma is the canonical form of the word and all its variants. For example, *propose* is a canonical form, and encompasses the stored information about the word’s meaning and the roles it plays in sentences whether it takes the form *propose, proposed, or proposes* (Levelt, 1989). The lemma frequency was chosen given evidence that information regarding sentence structure relevant to sentence processing times is related to the lemma, rather than any specific word form (Traxler & Tooley, 2007). For example, a direct object is expected to follow any form of *propose*.

The co-occurrence frequency of the prepositional phrase with the related verb or noun was then assessed. First, the number of occurrences of the relevant preposition within 9 words of the verb or direct object noun was determined. In the case of example *a* above, the preposition *to* followed *propose* and all its verb forms in 386 cases. Not all of these prepositional phrases played the same role as in example *a*, that of the destination or recipient of the proposal. So the instances of the phrases headed by the same preposition and in which the prepositional phrase played the same role as the experimental sentences were identified. In cases where more than 200 sentences were available, the proportion of those phrases playing the same role was assessed in the first 200 cases, and that proportion was applied to the total number of co-occurrences. In example *a*, the prepositional phrase played the same role in 75% of the first 200 cases. As a result, in 291 (75% * 386) of the 386 cases, the verb lemma *propose* was followed by a prepositional phrase headed by the same preposition (*to*). Furthermore those prepositional phrases played the same role (recipient) as in the experimental sentence. This number of occurrences became the numerator of the conditional probability calculation, and the verb or noun lemma frequency was the denominator. *Propose* had 3748 occurrences in the spoken portion of
the COCA. In example a, the probability of a prepositional phrase headed by to and with the same sense as in the stimulus sentence following the verb lemma propose was .078 (291/3748 = .078).

For the final revised set of self-paced listening materials, the probabilities of prepositional phrases co-occurring with the main verb or direct object in the argument conditions were larger than the probabilities in the adjunct conditions, with the argument condition mean (SD) of .139 (.066), and the adjunct condition mean (SD) of .035 (.059). The variance of the argument and adjunct co-occurrence probabilities did not differ based on Levene’s test for equality of variances ($F = 1.72, p = .194$). This equivalence of the variances across the argument and adjunct conditions enabled valid comparisons of correlations of frequency and listening times for the two conditions. The lemma frequency of the main verbs and direct object nouns did not differ across four conditions for the task ($F(3, 60) = .073, p = .974$). The same lemma frequency also did not differ across the argument and adjunct conditions ($p = .96$).

**Norming studies**

Sentence plausibility and word familiarity were also controlled across conditions for the self-paced listening materials. Typical students, age 18-25 years, rated the sentence plausibility and word familiarity of the materials. Participants were students in introductory psychology courses at The Pennsylvania State University who received course credit for their participation. All were native speakers of English, and all provided written informed consent prior to beginning the study.

Plausibility ratings used a 7-point scale, and the directions for participants followed prior studies of sentence processing (Boland & Blodgett, 2006; Schutze & Gibson, 1999). An initial plausibility study resulted in self-paced listening materials that differed in plausibility rating
across conditions. After revising the materials, a second study was conducted, and those results are reported here. Participants were asked to rate how likely to occur were the events described in the sentences. Unlike prior studies (Boland & Blodgett, 2006; Schutze & Gibson, 1999), which converted the sentences to a passive form for their norming studies, the sentences were presented as they appeared in the main study. After rating two trial sentences to verify their understanding of the directions, 14 participants rated self-paced listening sentences presented in 3 different randomized orders. Inter-rater agreement was assessed by calculating the intra-class correlation, which was .82 for the plausibility ratings. The final set of materials did not differ on plausibility rating across conditions ($F(3, 60) = .69, p = .562$) nor across argument and adjunct conditions ($p = .15$).

Word familiarity ratings followed the approach of prior studies by asking participants how often they had seen, heard or used the target words (de Groot, Dannenburg, & van Hell, 1994). Ratings again were on a 1-7 scale. There were three different word familiarity studies conducted, each with 14 participants as described above. Materials for the word familiarity studies consisted of both self-paced listening and sentence imitation materials. The first two studies were conducted in parallel. Ten percent of the words across the first two studies overlapped, and the familiarity ratings for these words were highly correlated across studies ($r = .94$). Intra-class correlations were .90 for these studies. A third word familiarity study was conducted in combination with the second plausibility study outlined above. For that study, the intra-class correlation was .89, and for the set of words common to the 3rd study and prior studies, the correlation of ratings across studies was high ($r = .75$). For the final set of self-paced listening materials, mean word familiarity ratings drawn from all studies did not differ across conditions for main verbs ($F(1,62) = .00, p = 1.0$), direct object nouns ($F(1, 62) = .442, p = .525$), nor for nouns acting as the objects of the critical prepositional phrases ($F(1, 62) = 2.549, p = .115$).
Confirmation of argument status

The final set of self-paced listening materials was reviewed to ensure that the prepositional phrases met the conditions of argument status indicated by their experimental condition. The original materials had all passed the linguistic tests for argument status formulated by Schutze and Gibson (1999). Therefore, any revisions to those materials were reviewed to ensure they met the same set of tests.

For prepositional phrases to appear in argument conditions, they had to pass one or more of the six tests. For phrases to appear in adjunct conditions, they had to pass none of the tests. These tests rely on the intuitions of native speakers for whether a sentence is grammatical or ungrammatical. For example, argumenthood was evaluated in the following sentence:

*The police detective proposed an investigation of the company, right away.*

In this case, the prepositional phrase *of the company* was judged to be an argument of the noun *investigation*. One test of argument status is iterativity (Schutze & Gibson, 1999). This test indicates that a series of adjuncts may modify a noun, whereas arguments may not be iterated in the same way. Here, it is not grammatical to say *an investigation of the company of the industry*. One may, however, grammatically say *an investigation through the city along the river*. The ungrammaticality of the first case suggests that *of the company* is an argument, whereas the acceptability of second case indicates that *through the city* is an adjunct.

Procedures

The final stimulus sentences for the self-paced listening task were recorded word-by-word by a male native speaker of standard American English. Word lists were developed for recording without regard to the order of the words in the sentence materials. This resulted in
recordings that did not have the typical prosodic features of these words appearing in sentences. Recordings were made in a sound-attenuated room using a Marantz PMD660 digital recorder. All sound files were trimmed for presentation to eliminate periods of silence before or after the stimulus word. Sound files were normalized to a common intensity level using Praat software (Boersma & Weenink, 2006).

The sound files were presented by computer using a script developed in E-Prime 2.0 (Psychology Software Tools, 2009). The materials were pseudo-randomly assigned to four different lists so that each list contained only one member of each of the experimental sentence quadruplets, or 16 experimental sentences. Lists were balanced by condition. Thirty-two filler sentences were added to each list, for a total of 48 sentences per list. The filler items were of varied syntactic structures, all designed to reduce the likelihood of the participant identifying a pattern in the experimental items and responding strategically. For each list, 18 of the 48 sentences were followed by comprehension questions. The questions were simple yes/no questions, recorded as complete sentences. The comprehension questions were designed to ensure that the participant was listening in a way that let them understand the sentences. Each of the 4 lists was presented in two pseudo-randomized orders. Randomization of the list orders was accomplished using the Mix program (van Casteren & Davis, 2006). Sentences were presented so that no experimental sentence immediately followed a comprehension question.

The following instructions for the task were read aloud to each participant:

You will listen to sentences, then answer questions on some of them. You will hear just a piece of the sentence at one time. You will press the SPACE bar to hear the next piece [the investigator pointed to the SPACE bar]. You will see a plus as long as there is more of the sentence. At the end of the sentence, you will hear a beep. Press SPACE again to hear the question, or for the next sentence. To answer the questions, press 1 for YES and 2 for NO [“yes” and “no” labels were affixed to the 1 and 2 keys of the computer]. Go to the next piece of the sentence as fast as you can. But make sure you give the right answers to the questions. The computer will keep track of your answers.
The participants were then presented with six practice sentences, four of which were followed by comprehension questions. Once participants demonstrated an understanding of the task, they proceeded to the experimental items. After the instructions, the task was presented under headphones at a comfortable loudness.

Each participant’s performance was evaluated for comprehension question accuracy. An accuracy of 78% (14 of 18 correct responses) was required to exceed chance (50%) and to be included in the analysis. Two participants did not reach this criterion, and were excluded from all study analyses on this basis.

Difference times were obtained for two regions, the prepositional phrase region, and the final region of the sentence. For the sentence, *The administrator announced many cuts to the reporter confirming rumors*, the prepositional phrase region is from the offset of *cuts* to the offset of *reporter*; the final region is from the offset of *reporter* to the offset of *rumors*.

The difference time (DT) was calculated as outlined in Ferreira and colleagues (1996) by first measuring the inter-response time (IRT) for each region. The IRT was the duration from the space bar push advancing the sentence from the prior region to the space bar push advancing the sentence to the following region, or to the concluding tone for the sentence. The duration of the sound files for each region were measured using Praat, and these durations were subtracted from the IRT. The E-Prime program also measured the computer’s latency for initiating the sound files, and these latencies were also subtracted to arrive at the DT. In summary, the DT was the IRT minus the sum of the duration of the sound files for the region and the latency of sound file presentation.

Outliers for each of the DT measures were identified by eliminating responses more than 2.5 standard deviations beyond the mean DT for the each participant by region. DT’s less than -150 ms and more than 2000 ms were also eliminated. A DT of 0 allowed the participant to listen to the entire region, so a DT greater than 2000 ms was judged to be a response during which the
participant was distracted from the comprehension task. Conversely, a DT of less than -150 ms did not permit the participant to hear enough of the region to understand the phrase, and so indicated that the participant skipped this region. This outlier removal standard resulted in a loss of 4.4% of data from the prepositional phrase region, and 4.2% of data from the end region. This degree of loss was the equivalent of a loss of 2 participants, which still allowed the analyses to meet the goals for statistical power, a .80 chance of detecting a large effect.

In several prior studies of online sentence processing, a second approach has been used to account for the variation in stimulus length. The length of a stimulus item (the word or sentence region) is used as a predictor of the reading time of the participants (Boland & Blodgett, 2006; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Schutze & Gibson, 1999). The actual reading time is then subtracted from the predicted reading time to arrive at an adjusted reading time. The regressions to predict reading times are done by participant. This technique has been extended to the self-paced listening task by using the duration of the audio presentation for the region of interest as the predictor of inter-response times (Booth et al., 2000). To enhance the comparability of the results for this study to those of Schutze and Gibson (1999) and Boland and Blodgett (2006), these adjusted listening times were also calculated by participant as an alternate dependent variable. By participant regressions were performed after elimination of outliers.

Analytical approach

To detect effects of condition, group, and group by condition interactions, a series of mixed effects linear regressions were performed for the prepositional phrase and the end regions of the experimental sentences. Since both participants and items are samples from larger populations of individuals and linguistic items, sentence processing data has often been analyzed using two separate ANOVAs, one with means by participant as the dependent variable, and a
second with means by item (Baayen, Davidson, & Bates, 2008). Recently, researchers have suggested that mixed effects regression may be superior to the use of separate ANOVAs, since such models better tolerate missing data and provide the ability to model random effects for both participants and items in one analysis. The regression approach also allows for better control of the family-wise error rate (Baayen et al., 2008; Locker, Hoffman, & Bovaird, 2007).

For this study, mixed effects regressions were performed with both difference times and adjusted listening times as dependent variables. Two models were evaluated. The first evaluated the Argument Structure Hypothesis of Boland and Blodgett (2006). The hypothesis was that arguments are generally a part of the lexical entry of the associated head (verb or noun). As a result, encountering a head word is proposed to activate the potential arguments of the word, resulting in faster processing of phrases acting as arguments. Fixed effects were therefore evaluated for argument status (at 2 levels, verb and noun arguments, verb and noun adjuncts), group (SLI and typical), and a group by argument status interaction. A second analysis was performed to evaluate the Argument Preference Strategy of Schutze and Gibson (1999). This hypothesis was that maximizing argument relations would be preferred during sentence processing over minimizing structural complexity (number of nodes in the syntactic tree). This hypothesis specifically tested noun argument versus verb adjunct processing, because the predictions of the Argument Preference Strategy contrast with the structural simplicity (Minimal Attachment) hypothesis for these conditions. Minimal Attachment would predict faster processing for verb adjuncts, whereas the Argument Preference Strategy predicts faster processing for noun arguments. Therefore the fixed effects of argument (2 levels, noun argument and verb adjunct) and group (SLI and typical) were evaluated for this analysis. Both linear mixed effects models were constructed using the linear mixed effects package lme4 for the R computing environment (Bates & Maechler, 2010) and followed the approach for analysis of psycholinguistic data outlined by Baayen (2008).
Mixed effects regressions evaluated the whether the fixed factors of group and argument status (or noun argument compared to verb adjunct) contributed the fit of the model. Because there is no consensus on how to evaluate the degrees of freedom in these models, statistical significance is evaluated by two standards (Baayen et al., 2008). First, $t$ statistics are computed, and those with absolute values greater than two are considered statistically significant (Baayen, 2008). Second, the contributions of independent variables to model fit is evaluated by comparing models including the predictor of interest to reduced models excluding the predictor. The contribution of the predictor to model fit is evaluated by the likelihood ratio test (Arnon & Snider, 2010; Baayen, 2008).

Correlations were also computed for difference times and adjusted listening times with the co-occurrence probabilities of the argument and adjunct phrases.

Results and discussion

The 46 participants included in the analyses for the self-paced listening task each performed better than chance on the comprehension questions embedded in the task. The mean ($SD$) for the SLI group was 90% (6%) correct, and the mean ($SD$) for the TL group was 94% (7%) correct. The difference times and adjusted listening times by group and condition are summarized for the prepositional phrase (PP) region and the end region in Table 2. The typical language group had shorter absolute difference times, but with considerable variation around the means. Adjusted listening times are zero insofar as the processing time is simply the result of the duration of the sound files for the region of interest. When listening times are shorter than this prediction, the adjusted listening times are negative; when they are longer than this prediction, adjusted listening times are positive. Review of the adjusted listening times suggests that argument phrases were processed more rapidly than adjunct phrases in the prepositional phrase
region, although the contrast is more apparent in the noun argument versus verb adjunct contrast than in other conditions.

**Table 2**
Mean (SD) difference times and adjusted listening times by condition, group, and sentence region.

<table>
<thead>
<tr>
<th>Condition</th>
<th>PP Region</th>
<th>End Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SLI</td>
<td>Typical</td>
</tr>
<tr>
<td><strong>Difference Times</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb Argument</td>
<td>598 (314)</td>
<td>517 (340)</td>
</tr>
<tr>
<td>Noun Argument</td>
<td>540 (257)</td>
<td>441 (291)</td>
</tr>
<tr>
<td>Verb Adjunct</td>
<td>597 (365)</td>
<td>543 (272)</td>
</tr>
<tr>
<td>Noun Adjunct</td>
<td>603 (401)</td>
<td>462 (335)</td>
</tr>
<tr>
<td><strong>Adjusted Listening Times</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb Argument</td>
<td>21 (232)</td>
<td>34 (231)</td>
</tr>
<tr>
<td>Noun Argument</td>
<td>-45 (234)</td>
<td>-61 (225)</td>
</tr>
<tr>
<td>Verb Adjunct</td>
<td>25 (252)</td>
<td>45 (205)</td>
</tr>
<tr>
<td>Noun Adjunct</td>
<td>-2 (293)</td>
<td>-14 (211)</td>
</tr>
</tbody>
</table>
The mean difference times by argument condition across sentence regions are shown in Figure 5. While participants appeared to have taken longer to process the prepositional phrase as compared to other regions, there is little effect of argument status. Figure 6, which focuses on the noun argument versus verb adjunct contrast, suggests a larger performance difference by condition. For both participant groups, noun arguments appear to be processed faster than verb adjuncts. There does not appear to be a more pronounced effect of noun argument versus verb adjunct for the group with SLI; in fact, the group with SLI appears to show less of a performance difference for the noun argument versus verb adjunct contrast.

Figure 5
Mean difference time in milliseconds by sentence region, by argument versus adjunct conditions.
Figure 6
Mean difference times in milliseconds by sentence region for noun argument versus verb adjunct conditions. Group with typical language top panel, group with SLI bottom panel.
Models of argument status and group effects

To better approximate normal distributions, difference times were transformed by taking the square root of the difference time plus 150. The difference times were shifted in order to eliminate negative values prior to the transformation. The transformed difference times for the prepositional phrase region were the dependent variables in a series of mixed effects linear regression models. Initial models determined whether random slopes by participant by item were justified, in addition to random intercepts by item and participant. Random slopes by item by participant were not justified by the likelihood ratio test ($\chi^2 (2) = 3.04, p = .218$). Random intercepts for both items and participants were justified, and were included in the subsequent models ($p's < .05$). A model including fixed effects of group, argument status and an argument status by group interaction suggested that none of the fixed effects significantly contributed to the model fit, with all $t$-values less than 2.00. Comparison of full and reduced models confirmed that the group by argument status interaction did not contribute to the model fit ($\chi^2 (1) = .44, p = .51$). Likelihood ratio tests also confirmed that group ($\chi^2 (1) = 1.95, p = .16$) and argument status ($\chi^2 (1) = .71, p = .40$) were not significant predictors. The estimated coefficients for group ($b = 1.5, SE = 1.3$), and argument status ($b = -.63, SE = .58$) were in the expected direction: being in the SLI group added to the predicted response time for the segment, and an item being classed as an argument reduced the predicted response time, although neither factor was statistically significant.

A linear mixed effects model predicting transformed difference times for the end region also found no reliable effect on model fit for the fixed effects of group, argument status, and the group by argument status interaction (all $t$-values < 2.00). Models were also evaluated using adjusted listening times for the prepositional phrase and end regions. Both models supported the same conclusions as those based on the difference times. In both cases, the fixed effects of group,
argument status, and the group by argument status interaction did not contribute to model fits (all $t$-values < 2.00). The model results are consistent with the descriptive and graphical analysis of the data: the effect size of argument status, based on condition means and standard deviations, was small, $d = .07$. Furthermore, there was considerable variation in the data around the conditional means, which made this very small effect difficult to detect in the models.

**Models of the noun argument and verb adjunct contrast**

A second analysis of the self-paced listening data considered the contrast in processing times for the noun argument versus verb adjunct conditions. This contrast provided evidence to distinguish the Argument Preference Strategy from Minimal Attachment (preference for simpler syntactic structure) in sentence processing (Schutze & Gibson, 1999). The descriptive statistics and graphical presentation of the data indicated that this contrast in conditions had a larger effect size ($d = .26$) than the overall argument versus adjunct contrast. To evaluate the effect of argument status, only the data from conditions 2 and 3, noun arguments and verb adjuncts, were considered in the following models.

The first dependent variable for the noun argument versus verb adjunct models was the square root transformed difference times for the prepositional phrase region. After evaluating random effects, random intercepts were included for participants and items in the final models. Fixed effects were entered for group, argument status, and the group by argument status interaction. Argument status in this model significantly contributed to model fit ($b = -1.84, SE = .61, t = -3.0, \chi^2(1) = 8.21, p = .004$). The coefficient was negative, suggesting that, as predicted, arguments were processed faster in the prepositional phrase region. The group by argument status interaction did not contribute to model fit ($t = 1.75, \chi^2(1) = 3.05, p = .08$). The estimated coefficient for the interaction was positive ($b = 1.85, SE = 1.06$), the opposite of the
direction predicted by the Procedural Deficit Hypothesis. Given group (SLI = 1, typical = 0) and argument status (argument = 1, adjunct = 0) coding in the model, for participants with SLI processing arguments, this term suggests that the response time to argument constituents for the group with SLI was slightly longer than for the typical language group. This implies a narrower response time difference between argument and adjunct conditions for the group with SLI, consistent with Figure 6. The prediction of the PDH was for a larger difference. The group term did not contribute to model fit (\( b = .83, SE = 1.29, t = 1.5, \chi^2(1) = 2.35, p = .13 \)).

For the end region, a model predicting transformed difference times found that none of the fixed effect predictors--group, argument status, or the group by argument status interaction--improved the model’s fit over an intercept-only model (all \( t \)-values less than 2.00).

Turning to the adjusted listening times, a mixed effects model estimated adjusted listening times in the prepositional phrase region with group, argument status, and group by argument status interaction terms as fixed effects and random intercepts for participants and items. The results were similar to the model using transformed difference times in the prepositional phrase region. Argument status made a contribution to model fit based on the \( t \)-value of -2.14, \( b = -117.5, SE = 54.99, \chi^2(2) = 3.55, p = .060 \), and the coefficient again indicated a processing time advantage for arguments. Neither the group by argument status interaction \( (b = 41.9, SE = 41.3, t = 1.02, \chi^2(1) = 1.04, p = .308) \) nor the group term alone \( (b = -26.7, SE = 38.8, t = 1.40, \chi^2(2) = 1.18, p = .57) \) contributed to model fit.

For the end of sentence region, the mixed effects model predicting adjusted listening times with group, argument, the interaction of group and argument indicated that none of the fixed effects predictors contributed to model fit (all \( t \)-values less than 2.00).
Correlations of co-occurrence probability and response times

With evidence for an effect of argument status in the secondary analysis, the relationship between the time to process the prepositional phrase regions and the probabilities of co-occurrence for the prepositional phrases and the nouns and verbs to which they attached were analyzed. For the typical language groups, the Argument Structure Hypothesis predicted a frequency effect in argument but not in adjunct conditions. The Pure Frequency Hypothesis predicted frequency effects for both argument and adjunct conditions (Boland & Blodgett, 2006).

Inspection of the distributions for the co-occurrence probabilities of the prepositional phrases following their associated nouns or verbs indicated a considerable departure from normality, particularly for the adjunct condition. Transformations did not result in satisfactory distributions for parametric analyses for adjuncts. Further, graphs of the relation between the probabilities and the adjusted listening times indicated non-linear relationships. Scatterplots of the adjusted listening times in the prepositional phrase region with the probabilities of co-occurrence are shown in Figure 7. These plots include curves estimating the relationship between the variables.
The scatterplots suggested that the relationship between the adjusted listening times and the co-occurrence probabilities followed a curvilinear relationship. A mixed effects regression including random intercepts for participants and items, and fixed effects of argument status and the co-occurrence probability indicated that a quadratic term, the square of the co-occurrence probability, was a significant predictor of adjusted listening times ($t$-value = 3.03, $\chi^2 (1) = 8.939$, $p = .003$). This supported the visual evidence for a curvilinear relationship. The distributions of the probabilities of co-occurrence for the adjunct phrases did not meet assumptions of normality,
linearity and homogeneous variance required for Pearson correlation analyses. A non-parametric correlation analysis based on ranks (Spearman’s rho) was therefore conducted to explore these relationships in detail. Given the curvilinear association, the relationship between adjusted listening times and co-occurrence probabilities was assessed separately on two ranges of probability. The inflection point of the estimated curvilinear relation shown in the scatterplots was estimated to fall at a co-occurrence probability of .12. The correlation was assessed for cases with co-occurrence probabilities above and below that inflection point. The PDH predicts that normal-range performance may be achieved by the group with SLI by a compensatory use of associative memory in cases where typical language individuals would use a rule-governed approach. As result, correlations of co-occurrence probabilities and adjusted listening times were assessed separately by group.

The results of this process are presented in Table 3. For the group with typical language, the correlation of adjunct phrase adjusted listening times with co-occurrence probabilities was not significant, whereas for the group with SLI the correlation was significant, and in the expected direction. More frequently co-occurring adjunct phrases were processed faster for the group with SLI, but this pattern did not hold for the group with typical language. This pattern suggests that the group with SLI is processing adjunct phrases differently from the group with typical language, even while achieving similar processing times.

To further validate this difference in adjunct phrase processing, I evaluated whether the correlation coefficient for the group with SLI was larger than the coefficient for the group with typical language. Two ways have been proposed to test differences between Spearman correlation coefficients, both using Fisher’s Z-transformation (Myers & Sirois, 2006). The tests were conducted using an automated tool for testing differences in correlation coefficients based on the Z-transformation (Lowry, 2012). The first approach is to estimate the Pearson correlation coefficients and enter them into the formulas for the Z-transformation and standard error
estimates. Using this approach, the correlation of adjunct listening times and co-occurrence frequencies was larger for the group with SLI ($r = .111$ (TL), $r = .364$ (SLI), $z = -1.7$, $p = .045$, one-tailed). A more conservative approach is to enter the Spearman correlation coefficients into the formulas to derive the test statistic and standard error (Myer & Sirois, 2006). By this method, the test evaluating whether the correlation coefficient for group with SLI was larger than that of the group with typical language approached significance ($z = 1.46$, $p = .072$, one-tailed).

The pattern for argument phrase processing in Table 3 is more complex. Whereas the expected negative correlation for argument phrases is evident for both groups at co-occurrence frequencies falling below the inflection point of the curve, only the group with typical language showed a significant positive correlation above the inflection point. This positive correlation was not anticipated by any of the sentence processing hypotheses predicting faster listening times for more frequent argument structures (Boland and Blodgett, 2006) and is an important factor in the analysis given that many cases fell in the range above the inflection point.

Correlations using difference times did not reach statistical significance.
Table 3
Correlations (Spearman's Rho) of prepositional phrase region adjusted listening times and the co-occurrence probabilities of the argument or adjunct phrases.

<table>
<thead>
<tr>
<th></th>
<th>Group with Typical Language</th>
<th>Group with SLI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noun argument adjusted listening times</td>
<td>p (co-occurrence)</td>
</tr>
<tr>
<td></td>
<td>p (co-occurrence)</td>
<td></td>
</tr>
<tr>
<td>&lt; Inflection Point (N = 22)</td>
<td>-.492*</td>
<td>-.529**</td>
</tr>
<tr>
<td>&gt; Inflection Point (N = 65)</td>
<td>.257*</td>
<td>.201</td>
</tr>
<tr>
<td>Verb adjunct adjusted listening times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; Inflection Point (N = 79)</td>
<td>-.163</td>
<td>-.377**</td>
</tr>
<tr>
<td>&gt; Inflection Point (N = 5)</td>
<td>na 1</td>
<td>na 1</td>
</tr>
</tbody>
</table>

* p < .05, **p < .01

1 No variation in probability of co-occurrence in this range between the adjunct phrase and the verb.
Discussion

The results of the self-paced listening experiment indicated that general effects of argument status on sentence processing were small and difficult to detect. Prior studies have not reported effect sizes, but studies comparing argument preference effects to other factors on sentence processing have shown that argument effects are small in comparison to effects of sentence plausibility and verb bias (Kennison, 2002). Verb bias is the pattern of sentence structures in which a given verb appears. For example, some verbs are more likely to be followed by direct objects than sentential complements, so the verbs are said to be direct object biased (Garnsey et al., 1997).

Several methodological differences may account for the failure to find a general effect of argumenthood across all conditions in this study even though such effects were found in the study by Boland and Blodgett (2006). First, co-occurrence probabilities of the prepositional phrases with their associated nouns or verbs were adjusted in this study to obtain equal variances across argument and adjunct conditions. This process resulted in replacement of some sentences from the Boland and Blodgett study with sentences with lower probabilities of co-occurrence in argument conditions, and higher probabilities of co-occurrence in adjunct conditions. These changes may have narrowed the processing time differences between arguments and adjuncts.

A second methodological difference for this study was in the norming study to obtain plausibility ratings. For the present study, sentences as presented in the main study were rated for plausibility. In the Boland and Blodgett (2006) study, sentences were rated for naturalness in passive form. For example, whereas the sentence for this study was presented as “The leader withheld his comment about the candidate without warning,” in the Boland and Blodgett study, this sentence would have been presented for rating as “His comment about the candidate was withheld by the leader.” It is unclear what effect this change had on plausibility ratings, but the
justification offered by Boland and Blodgett was to avoid an ambiguity effects. It seems likely that any sentences with ambiguity may well have been rated as less natural. If this approach resulted in more implausible sentences being included in adjunct conditions in the Boland and Blodgett study, this may also account for their findings of argument effects.

Finally, the Boland and Blodgett (2006) participants were 32 typical language college students. The present study had a smaller typical group, and the participants in general were more heterogeneous. These participant differences likely resulted in more variability in the results of the present study, making it more difficult to detect a small effect of argument status. Furthermore, the trend for the group with SLI was for a narrower difference between noun argument and verb adjunct conditions in the secondary analysis, suggesting that inclusion of participants with SLI decreased the likelihood of finding a significant argument effect.

Despite the absence of general argument effects, the secondary analysis replicated the Schutze and Gibson (1999) findings for an argument preference over a preference for simpler syntactic structure in typical adults. The present study extends the argument preference findings to the auditory modality, and also shows that adults with SLI show an argument processing advantage. The effect size of the difference between the noun argument and verb adjunct conditions was also small, but the findings of an argument preference were consistent across both the difference time and the adjusted listening time analyses.

None of the analyses found a group by argument interaction, as predicted by the PDH. The non-significant group by argument interaction effect trended toward a narrower difference for argument versus adjunct processing for the group with SLI, contrary to the prediction of the PDH. Adults with SLI also performed similarly to their peers with typical language on the task in general. This surprising result suggests that neither the presumed procedural deficit of adults with SLI, nor the other deficits of adults with SLI, such as processing speed or working memory
limitations, had a material effect on online sentence processing time for the sentence structures used in this study.

The co-occurrence frequencies in this study were designed to have equal variance across conditions, allowing for analyses of the correlations between co-occurrence frequencies and processing times across conditions. The results suggest that the relation between co-occurrence frequencies and processing times may be non-linear. It may be that co-occurrence frequency eases processing within some ranges, but outside that range other factors in the sentence interact with frequency to make the relations of frequency and processing time more complex.

A second finding from the correlational analyses is the different pattern of results for the group with SLI as compared to the group with typical language. The findings for the group with typical language tentatively support the Argument Structure Hypothesis (Boland & Blodgett, 2006). Significant frequency effects were found for arguments but not for adjuncts. These findings suggest that processing of arguments and adjuncts differs. The Argument Structure Hypothesis posits that only argument structure information is part of the lexical entry for the verbs or nouns attached to the prepositional phrases. Information on adjuncts is not part of the lexical entry, and so frequency does not affect processing. Processing of adjuncts is by application of a general syntactic rule.

The negative correlation in one range of co-occurrence frequencies and positive correlation in another range were anticipated by neither the Argument Structure Hypothesis nor the Pure Frequency Hypothesis. This curvilinear relationship urges caution in adopting the Argument Structure Hypothesis, and suggests that a factor not captured by this study may interact with co-occurrence frequency to affect processing times in certain ranges of frequency. Plausibility is one candidate for such an interacting factor. Although plausibility ratings were not different across conditions, the three noun argument sentences with the lowest plausibility ratings (all below four on a seven point scale) all had co-occurrence frequencies over .12. It is possible
that lower plausibility slowed the processing of these sentences, despite their high co-occurrence frequencies.

The correlation findings for the group with SLI differed from the group with typical language. The group with SLI showed frequency effects for both argument and adjunct processing, and the greater magnitude of the adjunct correlation coefficient for the group with SLI as compared to the group with typical language approached statistical significance. This supports the view of the PDH that individuals with SLI may compensate for syntactic (procedural) deficits by increasing reliance on associative processes supported by declarative memory (Ullman & Pierpont, 2005). The PDH suggests that adults with SLI have stronger associative memory connections for more frequently co-occurring adjunct structures and these stronger associations affect their processing of such structures.

If adults with SLI have compensated for their procedural memory deficits by utilizing declarative memory systems for processing adjuncts, the PDH may be supported by these findings. Based on the suggestions of Thomas (2005), this supposes that the declarative memory systems of the group with SLI are normal. Normal range performance relying on compensatory mechanisms suggests utilization of an intact system to accomplish the compensation. As a result, whether the self-paced listening results support the PDH or bring it into question depends on the findings of the visual declarative and procedural memory measures. A further test of the PDH is whether adults with SLI have particular difficulties with adjuncts as compared to arguments in a speech production.
Chapter 4

Experiment 2: Sentence Imitation

The sentence imitation task was presented to the same participants who participated in the self-paced listening task. The goals of this task were to assess the predictions of the PDH for a sentence production task involving contrasting argument and adjunct conditions. The production task may be more challenging, revealing a vulnerability to adjunct processing that was not evident in the self-paced listening task. Further, measures of working memory and speed of processing were included to assess the roles of these factors on sentence imitation accuracy.

Materials

Forty-eight sentences were developed for the sentence imitation task. There were four conditions, each of which varied by whether the sentence was constructed with primarily argument constituents or adjunct constituents, and whether the sentence was short or long. Short conditions were 16 syllables in length and long conditions were 25 syllables in length. Length in syllables was determined by an automated syllable counter (McVeigh, nd) to ensure consistency across conditions. Short argument conditions were a mean (SD) of 11 (.7) words, and short adjunct conditions were 11 (1.2) words, and these conditions did not differ in length in words ($t(22) = .649, p = .523$). Long argument sentences were a mean (SD) length of 16 (1.8) words, and long adjunct conditions were 17 (1.9) words, and long conditions also did not differ in length in words ($t(22) = 1.30, p = .206$). The following are examples from each condition:

a. *The president* urged *the army to act* and blocked *their concerns.* (argument/short)
b. *The amateur* chopped *the stump with an axe* while *the professionals* estimated *his ability.* (argument/long)

c. *The dog* struggled in the pet motel and jumped at the signal. (adjunct/short)

d. *The lady* confessed in a whisper after the trial while *the household* grieved in the afternoon at church. (adjunct/long)

Italicized portions of the sentences are argument phrases. Classification of constituents as arguments or adjuncts was based on the six tests of argument status proposed by Schutze and Gibson (1999). The complete set of sentence imitation materials is provided in Appendix B.

Development of the sentences began by gathering verbs suited to each condition. Verbs with a bias for appearing in transitive constructions were used for argument conditions, and verbs with a bias for appearing in intransitive constructions were used for adjunct conditions. A transitive bias indicates that the verb is more likely to appear in sentences where there is a noun phrase object following the verb, as in *They bought rustled cattle from the outlaw.* An intransitive bias indicates that the verb is more likely to appear in sentences without an object following the verb, as in *The teacher remembered in class.*

**Corpus analyses**

Transitivity bias for verbs was obtained by combining data from two published corpus studies, and the classification of sentence frames to transitive or intransitive categories followed the approach used in an earlier study by Connine, Ferreira, Jones Clifton and Frazier (1984). Roland, Dick and Elman (2007) provided data on the frequency with which a set of 203 verbs appear in various transitive and intransitive sentence frames across both spoken and written corpora. Gahl, Jurafsky and Roland (2004) conducted a separate corpus analysis which provided transitivity ratings for 281 verbs. For their study, 200 sentences containing the target verbs were
reviewed for the sentence construction in which the verb appeared. Following the approach of Gahl and colleagues (2004), verbs appearing in transitive constructions in more than two-thirds of cases were labeled high transitive verbs for the present study. Those appearing in transitive constructions in one-third or fewer cases were labeled intransitive biased verbs. Gahl and colleagues found that including passive constructions shifted the transitivity bias of 34% of verbs. Because passive constructions have been shown to be more common in written corpora than in spoken language experience (Roland et al., 2007), passive constructions were excluded from calculations of verb transitivity bias for development of the sentence imitation materials. The combined data resulted in 278 candidate verbs for use in the sentence imitation task, 141 of which were transitive biased and 137 of which were intransitive biased. The spoken frequency of the verbs resulting from this process was determined from the Corpus of Contemporary American English (Davies, 2009), and was appended to the verb lists. Sixty transitive biased verbs that closely matched a set of 60 intransitive biased verbs on spoken frequency and length were determined using the Match program (van Casteren & Davis, 2007). Match is a computer program that finds near-optimally matching items across multiple factors. The resulting matched verbs did not differ on spoken frequency ($p = .98$).

For all conditions, sentence subjects were animate nouns. Candidate animate nouns, inanimate nouns, and modifiers for sentence development were extracted with spoken frequency, age-of-acquisition, and familiarity information from the MRC Psycholinguistics dataset (Clark, 1997). For each word class, candidate words for the sentence imitation task were randomly assigned to four lists, one for each sentence imitation condition. These lists were then submitted to the Match program to find sets of words well matched on length, frequency, and familiarity. These lists for each of the word categories then became the candidate lists for generation of sentence imitation materials.
For the final set of sentence imitation materials, spoken word frequency was determined from the Corpus of Contemporary American English for each of the main word classes (animate nouns, verbs, inanimate nouns used as objects, and modifiers). The spoken frequency did not differ across sentence imitation conditions ($F(3, 305) = .880, p = .452$).

**Norming studies**

Sentence imitation materials were controlled for the familiarity of the words and the plausibility of the sentences across conditions. Familiarity and plausibility ratings were obtained through the same series of norming studies described above for the self-paced listening materials. Balanced plausibility ratings for the sentence imitations materials were found in the first norming study in which there were 14 participants. Mean plausibility ratings for the sentence imitation materials ranged from 3.4 to 4.0 on a 7-point scale. Plausibility did not differ across the four conditions ($F(3, 44) = 1.51, p = .225$). Word familiarity ratings for the animate subject nouns, verbs, inanimate object nouns and modifiers also did not differ across conditions for the final materials ($F(3, 303) = .650, p = .584$).

**Procedures**

The final sentences were audio recorded by a male speaker of American English in a sound-attenuated room. The recordings were made using a Marantz PMD660 digital recorder. Complete sentences were spoken at a normal conversational rate. The sound files for each sentence were then edited in Praat (Boersma & Weenink, 2006) to ensure each sentence was presented at a common intensity.
The 48 sentences were pseudo-randomly assigned to two sets of 24 items each, balanced by condition. Three different random orders were created for each of the 2 sets of sentences. Each participant completed all 48 sentences presented in 2 lists, with each list presented in a separate experimental session. Experimental sessions were separated by a break of ½ hour at a minimum. List orders were counterbalanced across participants. The task was presented by computer using an E-Prime (Psychology Software Tools, 2009) script. Participants were told that they would hear sentences, and that they were to repeat the sentences exactly as they heard them in the recordings. Given that some sentences were long, and would be hard to remember, the participants were asked to repeat as much of the sentence as they could. There were two practice sentences to ensure that all participants understood the task.

Experimental sentences were presented auditorily at a comfortable loudness for the participant. Participants were given as much time as they required to repeat as much of the sentence as they could recall. All participant responses were audio recorded, again using the Marantz PMD660. Words recalled for each sentence were noted during the experimental session by the investigator. The participant’s recall accuracy was later verified and corrected using the audio recordings.

Scoring of the sentence imitation task used two methods. The first method assigned a score from 0 to 4 to each sentence, using procedures from the Clinical Evaluation of Language Fundamentals, 4th edition (Semel et al., 2003). A score of 4 was assigned for verbatim recall. Errors were counted by word, with one omission, addition or substitution resulting in a score of 3. Two errors resulted in a score of 2, three errors a score of 1, and four or more errors resulted in a score of 0. After the pilot study, many sentences resulted in 0 scores. To avoid floor effects, a percentage of words recalled scoring method was added. For this method, based on Allen and Baddeley (2009), the number of words recalled in correct serial order was recorded and that number was divided by the total number of words in the sentence. The result (multiplied by 100)
was the percentage of words recalled correctly in correct serial order. The participant was given credit for the recall of the correct initial and final words in the sentence. Further, any word recalled with at least one other correct adjacent word in the correct serial position was credited.

A research assistant was trained in the procedures for scoring the sentence imitation task. This second rater was blind to the language group classification of the participants. The research assistant scored 11% of the sentence imitation task items using the percentage of words recalled correctly method. The items were from randomly selected participants. The inter-rater reliability was assessed using the intra-class correlation (ICC) statistic, which estimates the variation of a measure due to true score variation after partitioning the variation due to rater differences (Howell, 2007). The ICC was .98, suggesting that very little of the sentence imitation score variation was due to rater variation.

**Working memory tasks**

The Competing Language Processing Task (CLPT) (Gaulin & Campbell, 1994) and the Running Span task (Cowan et al., 2005) were administered as measures of working memory capacity. The CLPT was designed for children. During pilot testing, no typical adult participants reached ceiling performance on the CLPT, so it was presented as designed by Gaulin and Campbell (1994). Participants listened to 42 simple sentences arranged in sets from one (Level 1) to six (Level 6) sentences. For each sentence, they immediately rendered a truth value judgment by saying “yes” or “no.” For example, on hearing *pumpkins are purple*, they were to say *no*. After hearing a set of sentences, they are asked to recall the final word of each sentence in the set. The sentences were all three words in subject-verb-object, subject-verb-modifier, or subject-auxiliary-main verb constructions.
To introduce the task, the investigator read instructions from Ellis Weismer, Evans and Hesketh (1999, p. 1253):

The person on the tape will tell you what to do, but I’ll tell you a little bit first. You’ll hear groups of sentences. After each sentence, tell me yes if it is true, or no if it is false. The person on the tape will ask you to tell me the last word in each sentence in the group. As we go, it gets harder because there are more and more sentences in the groups, but just do your best. We’ll do some practice ones so you get an idea of what the sentences are like.

The CLPT was presented from a computer audio file under headphones. The audio presentation followed the instructions from Gaulin and Campbell (1994). The recording was made by a female speaker of American English, using a head-mounted microphone and a Marantz PMD650 minidisk recorder. The recording allowed 4 s for the participants to provide the truth value judgments, and from 7 s (Level 1) to 19 s (Level 6) to recall the sentence final words. Following prior research on this task, participants were allowed only the pause time in the recording to make their responses (Mainela-Arnold & Evans, 2005). Participants were encouraged to adjust the presentation to a comfortable loudness level. The CLPT begins with 4 practice sentences during which the investigator provided feedback to ensure the participant understood the task. Participant truth value judgments and words recalled were noted during the session by the investigator, and audio recorded for later verification.

The scores for the CLPT were the percentage of correct truth value judgments, and the percentage of sentence-final words correctly recalled. Words recalled were accepted as correct with morphological variations, such as wheel for wheels.

The Running Span task followed the approach outlined in Cowan and colleagues (2005), which was based on a task originally developed by Cohen and Heath (R. L. Cohen & Heath, 1990). In this task, the participant heard lists of digits. The lists varied from 12 to 20 items in length. After each list, the participant was asked to recall the last five, six or seven digits from
the end of the list in the correct serial order. List lengths varied randomly, so participants were not able to anticipate when they would be asked to recall the last 5, 6 or 7 digits.

Digits one to nine were recorded individually by a male speaker of American English using a Marantz PMD660 digital audio recorder. Each sound file was then edited and compressed to be 250 ms in duration using Praat (Boersma & Weenink, 2006). The task was presented by computer using an E-Prime 2.0 (Psychology Software Tools, 2009) script. The script presented the task in a practice block and 3 experimental blocks. The practice block consisted of audio presentation of two lists of digits. After each list, the participant was asked to write the last five digits in the list in correct serial order. A paper form provided 5 boxes for practice responses. The participant was asked to fill in all boxes on the form. For any digits they could not recall, they were instructed to enter 0. Zero was not among the digits presented in the lists. The investigator provided corrections and repeated instructions until the participant demonstrated that they understood the task.

The experimental blocks began with the participant being asked to write the final seven digits from each list in correct serial order. The computer script presented nine lists of 12 to 20 digits in length for each block. The list lengths were presented in random order, and the digits making up each list were presented by randomly selecting audio files of digits without replacement, then repeating the random selection process until the designated list length was complete. The second block presented nine further lists in the same manner, but the participant was asked to write the final 6 digits from the list. In the third block, the participant was asked to write the final 5 digits.

Each trial began with Ready? on the screen, followed by the audio list, and then an array of response boxes. The participant handwrote their responses on a paper form, and the investigator entered the final participant response into the response boxes on the computer screen. The participant was asked to verify that their response and the response entered into the computer
were identical. Pilot participants had indicated that they were more comfortable writing their responses than entering them into the computer.

The data file created by the computer script listed the final five, six or seven digits presented by the computer, and the participant response for each trial. To score the task, these digit strings were compared. The participant responses were verified by checking the original paper form responses. The number of digits recalled in the correct serial position was the score for each trial. The mean for each block was calculated. The highest of these block means was the final score for the running span task.

**Processing speed tasks**

A truth-value judgment task and a word monitoring task were administered to measure participants’ speed of language processing. The truth-value judgment task was replicated from the response-time study conducted by Miller, Kail, Leonard and Tomblin (2001). In this task, participants were presented with a picture on a computer screen. After 2 s, participants heard a sentence presented from a sound file while the picture remained on the screen. If the meaning of the sentence matched the picture, the participant was to press a green button on a response box. If the meaning of the sentence did not match the picture, the participant was to press a red button.

The pictures for the task were the black and white line drawings used in the Miller et al. (2001) study. The stimuli were presented on a laptop computer using an E-Prime 2.0 (Psychology Software Tools, 2009) script and a Psychology Software Tools serial response box. These pictures were associated with 36 sentences, with equal numbers of three structures: simple active (*The girl is chasing the boy*), simple passive (*The baby is being fed by the girl*), or compound subject (*The boy and the horse are washing the cow*). The sentences matched the pictures for half of the trials, and did not match for the other half. The sentences and sound files
were also those from the Miller et al. (2001) study. The sentences were audio recorded as complete sentences with normal rate and prosody by a male speaker of standard American English.

The word monitoring task was adapted from the task developed for L. B. Leonard, Miller and Finneran’s (2009) study of adolescents with SLI. Leonard and colleagues presented sentences that were both grammatical and ungrammatical. For the current study, only grammatical sentences from the Leonard and colleagues materials were used. In this task, the participant heard a target word. They then heard a sentence. They were to press a button on a response box as soon as they heard the target word in the sentence. For example, participants heard the target word dinner, followed by the sentence The hiker at Yosemite always cooks dinner over a campfire.

The word monitoring task included 34 target word and sentence pairs presented at a normal speaking rate. The sentences ranged in length from 8 to 12 words, with the target word occurring after 5 to 7 words. The target words all exceeded a written frequency threshold of 2 occurrences in the Kucera and Francis (1982) corpus (L. B. Leonard et al., 2009). There were nine target word and sentence pairs interspersed with these trials in which the target word did not appear in the sentence. These trials were included to encourage the participants to listen carefully to the sentences before responding. Thirty-four additional trials were presented at a slowed rate, but these trials were not analyzed for this study.

Stimuli for the word monitoring task were recorded by a male native speaker of standard American English. Recordings were made in a sound attenuated room using a Marantz PMD650 digital recorder. All sound files were normalized to a common intensity. Experimental target words and sentences were presented in two pseudo-random orders. Both orders began with two items in the normal speaking rate condition. The remaining items were presented with no more
than five items of the same rate condition in sequence. At least one in nine items was a trial in which the target word did not appear in the sentence.

**Procedures**

For both speed of processing tasks, the response box was placed to the side of the computer aligned with the participant’s dominant hand. The investigator read the following instructions for the truth-value judgment task, based on those from Miller et al. (2001):

In this task, we’ll use the computer, but you’ll only need to worry about these 2 buttons. In this, and other tasks that we’ll do, the main thing to remember is to answer as fast as you can without making mistakes.

This time, here’s what’s going to happen. A picture will come up on the screen, and then you’ll hear a sentence. If the sentence is talking about the same thing that’s happening in the picture, push the green button. If the sentence is talking about something different from the picture, push the red button. We’ll do some practice ones first.

The task was presented at a comfortable loudness under headphones. After the instructions, the participant made judgments on six practice trials. The investigator indicated whether the participant was correctly following the instructions for the practice trials. After the practice trials, the participant had the option to repeat the practice trials or move on to the experimental trials. For each trial, pressing the response box button terminated that trial and the experiment moved to the next trial. Experimental items were presented in a single pseudo-random order, with no more than 3 trials of the same type (matching or not matching) or sentence structure.

The primary measure for the truth value judgment task was the mean response time in milliseconds for valid trials. Trials with incorrect judgments were excluded from the analysis. Trials with response times greater than twice the participant’s mean were also excluded.
Response time was measured from the onset of the sentence sound file to the participant’s button push.

For the word monitoring task, the investigator read the following instructions:

You will hear a word. Then you will hear a sentence. Listen for the same word in the sentence. As soon as you hear it, press the green button on the response box. If you don’t hear the word in the sentence, don’t do anything.

The participant then put on headphones, and the task continued with six practice trials during which the investigator reinforced compliance with the task instructions. Each experimental trial began with 2 s of silence, followed by the presentation of the target word. After 500 ms the sentence sound file was initiated. The trial ended with the participant response or after a 2 s wait time.

The primary measure for the word detection task was the response time from the onset of the target word to the button press. The response time recorded for each trial measured the duration from the initiation of the sentence sound file to the button press. For each sentence sound file, the duration from the beginning of the sound file to the initiation of the target word’s first phoneme was measured using Praat (Boersma & Weenink, 2006). This duration to the target word was subtracted from response time recorded in the E-Prime data file to arrive at the response time for detecting the target word. Response times with negative values and those greater than twice the participant’s mean response time were eliminated before calculating the final mean response time for all valid trials.

**Analytical approach**

To ascertain which factors contributed to sentence imitation performance, a mixed effects regression model was constructed to predict the percentage of words recalled correctly. Examination of the distributions of the item-level responses indicated that for many sentences in
short conditions, the typical group repeated 100% of words. As a result, item-level responses did not meet the assumptions for normality required for linear regression modeling. Inspection of scores averaged across conditions for each subject revealed distributions that were suitable for regression modeling.

The dependent variables for the analyses were therefore the means by subject by sentence condition, or four mean scores per participant. Since there were repeated measures by participant, the regression modeling included random subject effects, which accounted for the lack of independence of the sentence scores, much as a repeated measures ANOVA or ANCOVA would account for such clustering (Milliken & Johnson, 2009). Since the data had already been collapsed across items, no random effects by item were included. The mixed effects regression modeling was carried out in the R computing environment using package lme4 (Bates & Maechler, 2010). The modeling proceeded first with modeling and testing the random effects structure. This was done by evaluating the contribution of random intercept and random slope terms by subject using the likelihood ratio test (Baayen, 2008). Random subject terms which significantly contributed to model fit were retained in the model. Fixed effects were then evaluated sets by minimizing the information criteria values (AIC, BIC) for the models. Individual fixed effects were evaluated using the $t$-values (absolute value of 2 or more indicating significance) and likelihood ratio tests.

An additional question for this study was the magnitude of effect for argument status, working memory, processing speed, and group status. An $R^2$ equivalent for mixed effects regression models has been proposed (Kramer, 2005; Magee, 1990). This statistic is based on the change in log-likelihood for a model that includes the predictor of interest as compared to a reduced model that excludes that predictor. Like the conventional $R^2$, the statistic ranges from 0 to 1, and is an indicator of the portion of variance in the dependent variable accounted for by the predictor. For this study, calculation of the $R^2_{LR}$ (R-squared likelihood ratio) compared nested
models with the same random effects structures. A second means of evaluating the practical significance of the predictors for the sentence imitation task was to evaluate the predicted change in sentence imitation score given a change in the predictor of interest, with all other variables held constant. This was done using the final regression model intercept and coefficients. For example, the effect of argument status was evaluated by calculating the difference between the predicted score for sentences in the argument, short condition to predicted scores for the adjunct, short condition. This difference was evaluated for each participant group in turn, and assumed study-average working memory capacity. The result was a measure of change in sentence imitation percentage score for the isolated effect of argument condition.

**Results: Working memory, speed of processing and sentence imitation**

For sentence imitation, there were two primary questions. First, the performance on the argument as compared to the adjunct conditions provides another test of the Procedural Deficit Hypothesis. If the PDH holds, then we would expect to observe an argument status by group interaction, with the lack of argument support resulting in a greater decrement in performance for the group with SLI as compared to the group with typical language. The second focus of the study was to evaluate factors that may explain sentence imitation performance, particularly argument status, participant working memory, and participant language processing speed.

**Working memory and processing speed**

A summary of performance on the working memory and language processing speed tasks is presented by group in Table 4. In the Competing Language Processing Task (CLPT), both groups were highly accurate in judging the truth of the sentences. Final CLPT Recall
performance and mean digits recalled in Running Span both resulted in markedly different group performance.

For the language processing speed tasks, group performance differences were more modest. There was no group difference in the accuracy of matching sentences meanings to pictures in the Truth Value Judgment task. The absolute response times of the group with SLI trended longer than those of the group with typical language on both tasks, but differences were more pronounced for the word detection task.

**Table 4**

Mean (SD) performance, tests of group mean differences, and effect sizes for the working memory and language processing speed tasks.

<table>
<thead>
<tr>
<th>Task</th>
<th>SLI M (SD)</th>
<th>Typical M (SD)</th>
<th>t (44)</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Working Memory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLPT Truth Value Accuracy</td>
<td>99% (2%)</td>
<td>99% (1%)</td>
<td>.65</td>
<td>.52</td>
<td>0</td>
</tr>
<tr>
<td>CLPT Recall</td>
<td>70% (11%)</td>
<td>86% (10%)</td>
<td>5.5</td>
<td>.00</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Running Span</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Digits Recalled</td>
<td>3.18 (.60)</td>
<td>4.19 (.73)</td>
<td>5.1</td>
<td>.00</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Language Processing Speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truth Value Judgment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>96% (4%)</td>
<td>97% (2%)</td>
<td>1.1</td>
<td>.29</td>
<td>.3</td>
</tr>
<tr>
<td>Mean RT (ms)</td>
<td>2366 (286)</td>
<td>2199 (304)</td>
<td>1.9</td>
<td>.06</td>
<td>.6</td>
</tr>
<tr>
<td>Word Detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean RT (ms)</td>
<td>422 (125)</td>
<td>341 (60)</td>
<td>2.8</td>
<td>.01</td>
<td>.8</td>
</tr>
</tbody>
</table>
Sentence imitation performance

Sentence imitation performance by both the ordinal, 0-4 scoring approach, and the percentage of words recalled in correct serial order approach are presented in Table 5. In the short conditions, scores for the adjunct conditions are below those of the argument conditions, as expected. Also as expected, scores for the group with SLI trended lower for all conditions. The maximum score for items in nearly all conditions for both groups was 100%. The only exception was the argument, long condition for the group with SLI, where the maximum score was 94%. An unexpected result was the trend for adjunct scores to exceed argument scores in the long conditions. This trend is reflected Figure 8.

Inspection of the ordinal scores revealed floor effects, particularly for the group with SLI in long sentence conditions. As a result, all analyses were based on the percentage correct scores. With that approach, many item scores in the typical group for the short conditions reached 100%. To better meet assumptions of a normally distributed dependent variable for linear regression modeling, scores were averaged by participant for each condition. The dependent measure for the regression models of the sentence imitation task were the four condition means for each participant.
Table 5

Mean (SD) performance by condition and group for the sentence imitation task. Scores for both the ordinal (0-4) and percentage correct scoring methods.

<table>
<thead>
<tr>
<th>Condition</th>
<th>SLI</th>
<th>Typical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ordinal</td>
<td>% Correct</td>
</tr>
<tr>
<td>Argument – Short</td>
<td>2.4 (1.5)</td>
<td>83.3 (20.7)</td>
</tr>
<tr>
<td>Adjunct – Short</td>
<td>1.6 (1.5)</td>
<td>70.8 (24.6)</td>
</tr>
<tr>
<td>Argument – Long</td>
<td>.1 (.4)</td>
<td>38.7 (20.1)</td>
</tr>
<tr>
<td>Adjunct – Long</td>
<td>.1 (.4)</td>
<td>40.9 (22.7)</td>
</tr>
</tbody>
</table>

*Group difference effect sizes for percent correct measures of sentence imitation accuracy.
Correlation of predictors to sentence imitation accuracy

Correlations of the grand mean accuracy for sentence imitation with proposed predictors of sentence imitation accuracy are summarized in Table 6. The strongest correlations are between the measures of working memory and sentence imitation accuracy. The working memory measures also correlate with each other, but they do not appear to be completely
overlapping measures. Of the language processing speed measures, word detection response
times correlated with sentence imitation accuracy.

Table 6
Correlations (Pearson) of sentence imitation grand mean percentage correct with scores of
working memory and processing tasks.

<table>
<thead>
<tr>
<th></th>
<th>SI Mean</th>
<th>CLPT Recall&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Run Span&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Word Det&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Truth Val RT&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI Mean</td>
<td>1</td>
<td>.64**</td>
<td>.67**</td>
<td>-.35*</td>
<td>-.16</td>
</tr>
<tr>
<td>CLPT Recall</td>
<td>1</td>
<td>.63**</td>
<td>-.31*</td>
<td>-.05</td>
<td></td>
</tr>
<tr>
<td>Run Span</td>
<td>1</td>
<td></td>
<td>-.35*</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Word Det</td>
<td>1</td>
<td></td>
<td></td>
<td>.33*</td>
<td></td>
</tr>
<tr>
<td>Truth Val RT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* <sup>a</sup> Competing Language Processing Task percentage of words recalled.  
<sup>b</sup> Highest mean digits recalled for a block.  
<sup>c</sup> Word detection response time.  
<sup>d</sup> Truth value judgment task mean response time.

Models of factors affecting sentence imitation accuracy

A mixed effects linear regression model was constructed to understand how sentence
argument and length conditions and individual differences in working memory and processing
speed affected sentence imitation accuracy. First, random effects were evaluated for their fit to
the data. A random slope by participant by sentence length improved model fit over random
intercepts by participant alone ($\chi^2 (2) = 149.9, p < .001$), whereas a random slope by participant
by argument condition did not further improve the fit ($\chi^2 (3) = 3.5, p = .315$).
Modeling proceeded with evaluation of fixed effects and their interactions. To reduce correlations among the predictors, the predictors argument, length, and group were centered and contrast coded using -1, 1 (Baayen, 2008). For example, argument conditions were coded 1, adjunct conditions as -1. The CLPT Recall and Running Span scores were converted to standard scores and summed to create a composite score for working memory. The same approach was taken with Truth Value Judgment and Word Detection scores to create a processing speed composite measure.

The fixed effects of argument, group and length were added to the random intercepts and random slopes by length by subject factors. All possible interactions of fixed effects were included in this model. To this base model, the addition of the processing speed composite was evaluated. The processing speed composite did not improve model fit ($t$-value = -.46, $\chi^2 (1) = .21, p = .645$). Since the Word Detection measure resulted in larger group differences and better correlated with overall sentence imitation performance, the mean response time from the Word Detection task was added to the model. It also did not improve model fit, ($t$-value = -1.24, $\chi^2 (1) = 1.6, p = .206$).

The composite working memory score was next evaluated in the model already including argument, group, and length factors. The working memory composite significantly improved model fit, ($t$-value = 3.19, $\chi^2 (1) = 7.34, p = .007$). The model with fixed effects of argument, length, group, and working memory was then refined by evaluating the interaction terms. Those making significant contributions to model fit were retained in the model. The best fitting model is summarized in Table 7. Examination of diagnostic plots, including the distribution of model residuals versus fitted values and the normal probability plot of residuals indicated that the model fit the data well and met the assumptions of linear regression models.
Table 7
Final mixed effects regression model predicting sentence imitation accuracy. The model also included random intercepts by subject, and random slopes by length by participant.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>b</th>
<th>SE</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.67</td>
<td>.013</td>
<td>50.31</td>
</tr>
<tr>
<td>Group</td>
<td>-.037</td>
<td>.017</td>
<td>-2.12</td>
</tr>
<tr>
<td>Argument</td>
<td>.016</td>
<td>.004</td>
<td>3.68</td>
</tr>
<tr>
<td>Length</td>
<td>-.169</td>
<td>.006</td>
<td>-27.27</td>
</tr>
<tr>
<td>Working Memory Composite</td>
<td>.041</td>
<td>.003</td>
<td>4.05</td>
</tr>
<tr>
<td>Length x Working Memory</td>
<td>.015</td>
<td>.003</td>
<td>4.39</td>
</tr>
<tr>
<td>Length x Argument</td>
<td>-.033</td>
<td>.004</td>
<td>-7.73</td>
</tr>
<tr>
<td>Argument x Working Memory</td>
<td>-.007</td>
<td>.002</td>
<td>-3.05</td>
</tr>
</tbody>
</table>

The final model indicated that the group with SLI was significantly less accurate than the typical group ($\chi^2 (1) = 4.42, p = .036$). The effects of argument, length, and working memory, however, depended on the interactions of these effects. For example, the effects of working memory and argument varied by whether the sentences were in the short or long conditions. Further, the effect of working memory also depended on the argument condition of the sentences.

An interaction of particular interest for testing the hypotheses of the PDH was the argument by group interaction. The argument by group interaction was evaluated in the context of the model summarized in Table 7. Inclusion of this interaction did not significantly improve model fit ($t$-value = -.97, $\chi^2 (1) = .93, p = .334$). Interactions maintained in the model were justified by both associated $t$-values and by likelihood ratio tests. The length x working memory
composite interaction ($\chi^2 (1) = 16.73, p < .001$), length by argument interaction ($\chi^2 (1) = 47.27, p < .001$) and the argument by working memory interaction ($\chi^2 (1) = 9.17, p = .002$) all contributed to model fit. All other possible interactions of the individual predictors were evaluated, and none reached statistical significance (all $p$'s > .05, all $t$-values less than 2).

For significant predictors in the model, the effect size and practical effect of each was evaluated by calculating an $R^2_{LR}$ for the model and each factor, and by observing how contrasting conditions affected the model’s prediction for sentence imitation accuracy. The total $R^2_{LR}$ for the fixed effects in the model (after accounting for the random effects) was .70. The change in $R^2_{LR}$ associated with group was .02. This is the effect of group absent the effects of working memory differences. In practical terms, the change in predicted sentence imitation accuracy solely associated with group was 7.5 percentage points. This was estimated using the terms in the model outlined in Table 7 for the group with SLI versus the group with typical language for the short, argument condition, holding working memory constant at the study average.

The effect of argument status on sentence imitation accuracy was reflected in the interactions involving the argument term, as well as the main effect of argument. The total $R^2_{LR}$ for all argument terms was .28. Of this, the change in $R^2_{LR}$ for the working memory by argument interaction was .05, and for the length by argument interaction was .23, leaving a negligible amount for argument alone. The model predicts a nearly 10 percentage point advantage for argument condition over adjunct condition sentences for short conditions, all else held equal. In contrast, for long sentences, the argument condition is associated with a 3.5 percentage point disadvantage versus adjunct conditions. These argument effects are nearly identical by group, reflecting the absence of a group by argument interaction.

The length factor accounted for more than 30% of variation in sentence imitation performance ($R^2_{LR} = .31$). For adjunct conditions, the isolated effect of length was a 27 percentage point drop in accuracy from short to long conditions. For argument conditions, that
gap increased to a 40 percentage point drop in accuracy from short to long conditions. Clearly, sentence length explained a great deal of sentence imitation accuracy.

Working memory and interaction terms including working memory accounted for an estimated 17% of variation in sentence imitation accuracy, \( R^2_{LR} = .17 \). Of this, the argument by working memory interaction accounted for 5%, the length by working memory interaction 9%, leaving 3% associated with working memory alone. The working memory composite was the sum of the standardized CLPT Recall and Running Span scores. The average working memory composite score for the group with SLI was -1.24 and was 1.24 for the group with typical language. A working memory 1 composite score unit above the study mean was associated with a 5 percentage point increase in sentence imitation score for long, argument conditions. A working memory measured at 1 composite score unit below the study average resulted in a 5 percentage point drop in sentence imitation score. This decrement increased to 6 percentage points for adjunct long conditions. In short conditions, the 1 composite unit change in working memory has a predicted 3 percentage point change in sentence imitation score.

Sentence imitation performance was the result of a complex set of factors and their interactions. The largest overall effect on scores was length, accounting for 27 to 40 percentage point difference across conditions. Argument effects were quite varied, improving scores by 10 points in short conditions, but decreasing scores in long conditions. Next in magnitude was group. Classification as SLI was associated with a more than 7 point drop in score. For working memory, a 1 unit increase or decrease in composite working memory was associated with a 3 to 6 point change in score, with a larger magnitude difference found for long conditions.
Comparing the Competing Language Processing Task and Running Span

Since the composite working memory measure was a significant predictor of sentence imitation accuracy, a further question was whether Running Span would explain sentence imitation performance as well as CLPT Recall. The first question was whether each would significantly contribute to model fit as substitutes for the composite working memory measure in the model outlined in Table 7. Both did when entered separately. Running Span significantly contributed to model fit (t-value = 2.44, $\chi^2 (1) = 4.88, p = .028$), as did CLPT Recall (t-value = 2.20, $\chi^2 (1) = 4.05, p = .044$). When both measures were simultaneously entered into the model, neither was significant (t-values less than 2).

The degree to which Running Span and CLPT Recall explained sentence imitation variance was examined. With Running Span was entered in place of the working memory composite in the final model, the model explained 68% of sentence imitation score variation ($R^2_{LR} = .68$), and the contribution of Running Span and the interactions involving Running Span was 12% ($R^2_{LR} = .12$). When CLPT Recall was entered in place of the working memory composite in the final model, the model explained 69% of sentence imitation score variation ($R^2_{LR} = .69$), and the contribution of CLPT Recall and the interactions involving CLPT Recall was 14% ($R^2_{LR} = .14$). These data indicate that Running Span and CLPT Recall were very similar in their ability to explain variation in sentence imitation performance.

A final analysis examined Running Span and CLPT Recall as the only fixed effect predictors in the model with the same random effects structure as the model presented above. In the presence of Running Span, the CLPT Recall measure added to model fit (t-value = 2.78, $\chi^2 (1) = 6.713, p = .009$). Conversely, in the presence of CLPT Recall, Running Span added to model fit (t-value = 2.88, $\chi^2 (1) = 7.016, p = .008$). So whereas CLPT Recall and Running Span were
roughly equivalent predictors of sentence imitation performance in the context of other relevant predictors, they did not measure completely overlapping abilities.

Discussion

Unlike the self-paced listening task, there were clear group differences in performance on the sentence imitation task. Group differences were also evident in the working memory measures. The model predicting sentence imitation performance included two terms that contribute to explanations of sentence imitation group differences in performance: working memory and group. The group factor represented the vocabulary knowledge, sentence comprehension and spelling abilities assessed as part of the language ability classification. This factor made contributions to the model of sentence imitation performance beyond working memory.

Predictions of the Procedural Deficit Hypothesis

One of the goals of the sentence imitation task was to evaluate the prediction of the PDH that the group with SLI would have particular difficulties with adjunct processing, and would show typical performance for argument processing (Ullman & Pierpont, 2005). The clearest support of this prediction would have been a significant group by argument status interaction, with the group with SLI showing differentially poorer accuracy in adjunct conditions. There was no evidence of a group by argument interaction in the regression models.

The differences in group means were larger in adjunct conditions as compared to the argument conditions, as shown by the effect sizes. These differences appear to have been associated with group differences in working memory. There was a significant working memory
by argument interaction. The coefficient for this interaction term was negative. Given the coding of argument status and the working memory composite, this suggests that adjunct conditions (coded -1) increased the difficulty of the task for participants with below-study-average working memory (scores below 0). According to the PDH, the working memory deficits found among individuals with SLI are the result of procedural memory deficits (Ullman & Pierpont, 2005).

Given the working memory by argument status interaction, the results could be interpreted as supporting the predictions of the PDH. The argument conditions assisted participants with poor working memory by activating arguments prior to encountering them in the sentence, providing an enhancement to performance from the lexical/declarative system. Adjunct conditions did not provide the lexical activation enhancement. In other words, argument conditions mitigated the effects of working memory limitations, whereas adjunct conditions exacerbated working memory limitations. The fact that the group by argument interaction was not significant, even in preliminary models that did not yet include a working memory factor, makes this position more difficult to support. Some individuals with typical language fell below the study mean for working memory, and some individuals with SLI fell above the study mean for working memory. It seems unlikely that the working memory measure would better align with the presence of a procedural deficit than the group classification measure.

The final way in which the results of the sentence imitation task could align with predictions of the PDH is the argument that the group with SLI compensated for their procedural deficit using an intact declarative system. As in the case of the self-paced listening findings, this argument relies on an intact measure of declarative memory (Thomas, 2005). What differs in the sentence imitation data is the presence of clear group performance differences. The group and working memory factors account for significant differences in sentence imitation accuracy as compared to the group with typical language. The working memory effects are consistent with the PDH, but the absence of particular difficulties with adjunct conditions for the group with SLI
remains at odds with the PDH. For the PDH to be supported there must be an argument for selective compensation from declarative memory. In other words, the declarative system was engaged by the group with SLI to process adjuncts as effectively as did the group with typical language. Compensation of the declarative system was not, however, available to normalize performance differences resulting from working memory limitations. The more parsimonious account is that working memory and other group differences in language knowledge account for group differences, not a mix of compensated and uncompensated procedural deficits.

**Working memory and language processing speed**

For the working memory tasks, the results of this study replicate prior findings that adults with SLI have more limited working memory (Isaki et al., 2008), particularly when measured with tasks involving both processing and storage elements as was the case with the Competing Language Processing Task (CLPT). The CLPT was designed for children to provide an auditory measure of reading span as had been used in the framework of Just and Carpenter’s (1992) theory of working memory (Gaulin & Campbell, 1994). It is striking that the task produced large group differences for adults, suggesting that it remains a sensitive measure of working memory at the adult stage of development.

Running Span has not been measured previously in adults with SLI. The fact that the Running Span measure resulted in a large group difference, of the same magnitude as the CLPT, indicates that the storage element of working memory, conceived as the focus of attention (Cowan et al., 2005) is a particularly important element of working memory capacity differences in this population. An advantage of the Running Span task from a psychometric perspective was that no participant reached ceiling level performance on the task.
Of the language processing speed measures, the offline Truth Value Judgment task resulted in the smallest group difference of the four processing speed and working memory tasks. The word detection task resulted in somewhat larger group differences. Group differences on very similar tasks have been found in children with SLI at age 8-11 years (Montgomery, 2005), suggesting that differences in processing speed related to language impairment are not fully resolved in adulthood. Findings of group differences in processing speed is consistent with prior findings of processing speed differences between adults with and without a history of language difficulties (Miller & Poll, 2009). The word detection task results correlated with mean sentence imitation performance, suggesting a relationship between language processing speed and sentence production accuracy, although the correlation was not as strong as that found between language processing speed and sentence comprehension in the Miller and Poll study ($r = -.55$ for all participants).

Factors related to sentence imitation accuracy

The factors evaluated in this study accounted for 70% of sentence imitation performance variation. The effect of length accounted for the greatest portion. Scores for shorter sentences, which averaged 11 words, as compared to longer sentences, which averaged 16 words, differed by 40 percentage points in argument conditions. This difference was 27 points in adjunct conditions. Longer sentences resulted in larger group score differences. Larger effect sizes are associated with better group discrimination for diagnostic purposes, so longer sentences are likely to contribute more to the diagnostic accuracy of sentence imitation.

The length effect interacted with argument status. This tempers the diagnostic value of longer sentences, since it was clear that the effect of sentence structure, in this case argument status, was more evident in short conditions. Argument conditions had a 10 percentage point
score advantage over adjunct conditions for short sentences. This advantage reversed to a 3.5 point disadvantage for arguments in long sentences. The opposite trend was true for working memory. The effect of a greater working memory capacity was magnified in long conditions as compared to short conditions. These results validate the finding that there is an ideal length range in which sentence imitation reflects the syntactic systems of participants, likely below 11 words (Bley-Vroman & Chaudron, 1994). At lengths beyond this, sentence imitation appears to be a greater indicator of individual differences in working memory.

In short conditions, argument status had a considerable impact on sentence imitation accuracy. The effect suggests that, as predicted by models of the speech production process, argument constituents may take priority over adjuncts in the production process and are therefore more likely to be produced (Levelt, 1989). The results are also consistent with the notion that encountering lexical items in a sentence activates the arguments of those lexical items, easing their recall and production, in an extension of views of argument preference found for comprehension tasks (Boland & Boehm-Jernigan, 1998).

The argument disadvantage in long conditions was not expected. It is possible that long adjunct sentences required holding fewer chunks of information, or enabled the holding of larger chunks of information for recall than did long argument sentences. The mapping of sentence constituents to memory chunks is not well understood (Bley-Vroman & Chaudron, 1994). Some researcher have proposed the clause as the memory unit, or chunk, for sentence material (Gilchrist et al., 2008). This approach may be a credible way to test hypotheses about differences in chunk size versus number of chunks for working memory capacity. It does not clarify how clausal units with variation in the number of arguments or number of propositions may affect working memory capacity. Specifying why the argument advantage is overridden by memory-related factors in long sentences will require refinement of our understanding of how units of information from sentences are formed for maintenance in working memory.
Group classification explained a further 7-8 percentage points in sentence imitation score, after accounting for working memory, length and argument status. This suggests that the abilities assessed by the classification measures, such as vocabulary knowledge and sentence comprehension ability, also contribute to sentence imitation performance. There is evidence that the strength of word knowledge or lexical representations have an effect on recall of verbal material for children with SLI (Gathercole & Baddeley, 1990; Mainela-Arnold, Evans, & Coady, 2010). Difficulty comprehending sentences in the Modified Token Test may signal greater difficulties arriving at the gist of the sentences to hold in working memory, subsequently affecting sentence imitation accuracy.

Language processing speed was not a significant predictor of sentence imitation accuracy after accounting for argument, length, and working memory factors. The sentences in the sentence imitation task for this study had simple structures. They were in canonical word order and did not have embedded clauses. Longer sentences were constructed by adding elaboration and by creating compound sentences, sequences of clauses at the same level (Quirk et al., 1985). It is possible that more complex structures, involving more embedding, would require more cognitive operations for comprehension and recall, and would result in detectable effects of individual differences in language processing speed on sentence imitation. Additional cognitive operations may be required if complex sentences have additional planning units in the speech production process (Levelt, 1989).

In contrast to language processing speed, the working memory composite was a significant predictor of sentence imitation accuracy. This finding is consistent with prior findings of associations between working memory capacity and sentence imitation performance in children with SLI (Alloway et al., 2004). A new finding is that a scope of attention measure, Running Span, is very similar to a storage and processing measure, the CLPT, in predicting sentence imitation performance. This finding suggest that the storage element of working
memory, or the capacity of the focus of attention (Cowan et al., 2005), is crucial to sentence imitation performance, and is a crucial difference between the group with SLI and the group with typical language.

The multi-factor model of sentence imitation developed in this study was successful in identifying important, simultaneous effects of length, argument structure, working memory, and group classification. The analyses indicated that sentence imitation performance is complex, and not only depends on each of these factors independently, but on interactions of length, sentence structure, and working memory capacity. The model confirms the widespread finding that working memory limitations are an important element of SLI. The PDH suggests that working memory limitations are part of the procedural memory deficit that explains SLI. A second prediction of the PDH, however, that adjunct production would present particular difficulties for individuals with SLI as compared to argument production was not supported. In the less likely scenario that the absence of differential adjunct effects was the result of compensation, this would imply that the declarative memory abilities of the group with SLI should be normal (Thomas, 2005). The abilities of the adults with and without SLI on non-verbal measures of declarative and procedural memory are the focus of the final set of experimental tasks in this study.
Chapter 5

Experiment 3: Visual Procedural and Declarative Memory Tasks

The PDH links language processing to the procedural and declarative memory systems. Two well established measures of these memory systems do not rely heavily on language processing. If the PDH holds, then deficits in language processing associated with deficits in procedural memory should be accompanied by deficits in non-verbal procedural memory tasks. If language processing findings suggest that adults with SLI were compensating for their procedural deficits by relying on declarative memory, then declarative memory should be a relative strength, or should be normal.

In this study, procedural memory was assessed with the Weather Prediction Task (WPT) (Reber et al., 1996) and declarative memory was assessed with the Paired Associate Recognition Task (PART) (Ragland et al., 1995). Both tasks focused on learning visual patterns.

The Weather Prediction Task

The Weather Prediction Task is a probabilistic classification task. Participants predicted “rain” or “sun” using cards as cues. For each trial, the participant saw cards in four possible positions on a computer screen. They then had 5 s to press one of two buttons on a response box to indicate rain or sun. Their prediction was followed by a feedback screen showing the weather outcome and a happy or sad face to indicate whether their prediction was right or wrong. The same screen was accompanied by a synthesized sound to reinforce the feedback. The sound associated with a correct prediction had a rising frequency contour, and the sound associated with an incorrect prediction had a falling frequency, similar to sounds in arcade games.
The task consisted of 50 trials. Each of the 14 different card combinations presented in the task was associated with a probability of rain or sun ranging from 0 to 1. The probabilities of sun for the card combinations were those used in Reber, Knowlton and Squire (1996), shown in Table 8. Cards were presented in four positions on the computer screen. For example, if cards were present in the first and third positions (left to right), the probability of sun was .75. When a card was presented in only the fourth position, the probability of sun was .14. The probabilities for sun and rain were balanced across the 50 trials. Ten of the trials had probabilities of sun of either 1 or 0. Eight of the trials had probabilities of sun of .50.
Table 8

Card patterns, associated outcomes, and probabilities of sun for each card combination presented in the Weather Prediction Task.

<table>
<thead>
<tr>
<th>Cards Presented</th>
<th>Repetitions (Trials)</th>
<th>Sun (Outcomes)</th>
<th>Rain (Outcomes)</th>
<th>p(sun)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1100</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1110</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1.00</td>
</tr>
<tr>
<td>1000</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>0.86</td>
</tr>
<tr>
<td>1010</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>0100</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0.60</td>
</tr>
<tr>
<td>1001</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>0110</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>1011</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>1101</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>0010</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0.40</td>
</tr>
<tr>
<td>0101</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>0.25</td>
</tr>
<tr>
<td>0001</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>0.14</td>
</tr>
<tr>
<td>0111</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.00</td>
</tr>
<tr>
<td>0011</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The Weather Prediction Task was presented by laptop computer using a script developed in E-Prime 2.0 (Psychology Software Tools, 2009). The 14 card combinations and associated outcomes were entered into a list controlling the 50 trials. The planned probability of card combinations and outcomes was enabled by the number of repetitions for the combination and outcomes entered into the list. For example, if cards presented in positions 1 and 3 were to be associated with sun with a probability of .75, three of the trials for this card combination in the list resulted in sun, and one trial resulted in rain. The 50 trials were presented in a pseudo-randomized order for each participant. The randomization of card combinations and outcomes was constrained by a rule that did not allow the same card combination and outcome to be presented in consecutive trials.

**Procedures and scoring**

For both the WPT and the PART, the response box was positioned on the side of the computer closer to the participant’s dominant hand. The investigator read the instructions from the computer screen. In the WPT, participants were told that they would learn how to predict sun or rain using a deck of four cards. For each set of cards, they were to press either the sun or rain button on the response box. They would then see the weather and if their prediction was correct. After an example of the task sequence, they were told that at first using the cards would feel like guessing. Gradually, however, they would get better at using the cards to predict sun or rain. They were also told that they would need to make their predictions within five seconds.

After the investigator answered any questions, the task continued. For each trial, a card combination was presented for up to 5 s. If the participant did not press a button for sun or rain during this time, a “no response” screen was presented. When the participant did make a
prediction within 5 s, the feedback screen was shown for 1500 ms, followed by a 500 ms interval before presentation of the next trial. All participants completed 50 trials.

The outcome measure for the Weather Prediction Task was the percent correct by block. The 50 trials in the task were divided into 5 blocks for scoring. For each block, the participant’s prediction was compared to the probability of sun or rain for the card combination presented in that trial. If the participant’s prediction aligned with the probability of the card combination, it was scored as correct. For example, if a card combination was associated with sun at a probability of .60, and the participant predicted sun, that trial was counted as correct. If a card combination was associated with sun with a probability of .40 and the participant predicted sun, that trial was scored as incorrect. The eight trials with card combinations associated with sun at a probability of .50 were eliminated before calculating the block percent correct. Trials with no response were also dropped prior to calculating the participant score.

All but two participants recruited from the Iowa longitudinal study had participated in another study involving the WPT five to six months prior to the present study. The E-Prime script for this study was developed jointly with the investigators directing the earlier study, and procedures were identical for both studies. For participants who had completed the task in the earlier study, results from the earlier study were used in the analyses for this sentence processing study. This avoided practice effects in the WPT performance of these participants.

The Paired Associate Recognition Task

The measure of declarative memory for this study was the Paired Associate Recognition Test (PART), replicated from Ragland, Gur, Deutsch, Censits and Gur (1995). In this task, participants were shown four pairs of cards in each block. One of four key cards was shown at the top of the computer screen, and a target card was shown below the key card in each trial.
After seeing the four card pairs, there was a two minute delay. Participants were then asked to indicate which cards had been shown together. During the probe, the target cards were presented in the same order as they had been presented in the learning phase. With each target card, the four possible key cards were present on the screen, and were numbered from one to four. The participant was to press the number of the key card that had been shown with the target card. There were a total of five blocks, each consisting of the presentation of four card pairs, the delay, and the memory probe.

Cards from the Wisconsin Card Sorting Test (Grant & Berg, 1948) were digitally scanned and converted to color bitmaps for use in computer presentation of the PART. Four key cards were selected. Then, six target cards with no shared features (color, shape, or number) were randomly selected for association with each key card. For each of the five experimental blocks a different randomly selected order of key card presentation was used, and a unique set of the target cards associated with the key cards was randomly selected for the block. The task was presented by laptop computer using a script developed in E-Prime 2.0 (Psychology Software Tools, 2009) and a Psychology Software Tools Serial Response Box.

**Procedures and scoring**

The instructions for the PART followed those of Ragland and colleagues (1995). The participant was told that they would be shown pairs of objects, and were asked to remember them. They were then shown examples of the key and target cards, and completed three practice trials. For the practice trials, one pair of cards was shown followed by a probe asking the participant to indicate by number which of the four key cards was shown with the target card. Then the computer indicated whether their response was correct or incorrect.
In the experimental trials, each of the 4 card pairs was displayed for 5 s. The learning phase was followed by a 2 minute delay. A red circle was slowly formed during the delay to show progress through the delay period. Then the participant indicated which of the 4 key cards was shown with each target card. Participants responded at their own pace during the recall probe. The next block was initiated with instructions indicating that the task would continue as before but with a new set of cards. The participant then initiated the learning phase of the next block.

The score for the PART was the percentage of cards recalled correctly. The numbers of each of the key cards presented in the learning trials were recorded by the computer, as were the participant’s responses to the probe phase. These numbers were compared to determine the number of correct responses across the 20 total trials. Correct responses were divided by total responses and multiplied by 100 for the total percent correct.

**Results: Visual procedural and declarative memory tasks**

The Weather Prediction Task (WPT) and the Paired Associate Recognition Task (PART) were designed to assess procedural memory, in the case of the WPT, and declarative memory, in the case of the PART. Both tasks rely primarily on learning and recall of visual patterns. If SLI is fundamentally a procedural deficit, the implication is that tasks measuring procedural memory will result in larger performance differences than those tapping declarative memory when groups with SLI are compared to groups with typical language.
The Weather Prediction Task

The dependent variable for the WPT was the proportion of responses aligned with the probability of the cue set. Chance performance was .50. Performance by block and by group is presented in Figure 9. Both groups began with performance near chance, but improved considerably by the fifth block. The largest performance difference between groups was in block four.

Figure 9
Mean proportion of predictions aligned with probabilities by block and group for the Weather Prediction Task.
The proportion correct for the WPT was arcsine transformed to better approximate normality. A 2 (group) x 5 (block) repeated measures ANOVA revealed a main effect for block ($F(4, 176) = 4.98, p = .001$, partial $\eta^2 = .102$). Neither the main effect of group ($F(1, 44) = 2.55, p = .118$, partial $\eta^2 = .055$) nor the group by block interaction ($F(4, 176) = 1.24, p = .294$, partial $\eta^2 = .028$) were significant. The improvement of both groups resulted in performance at block five that exceeded chance, with typical group mean (SD) accuracy at .76 (.20) ($t(22) = 6.4, p < .001$) and group with SLI mean (SD) accuracy at .70 (.18) ($t(22) = 5.12, p < .001$).

An ANOVA analyzing block four performance alone tested the intuition from Figure 9 that group performance diverged at that block. The effect of group trended significant ($F(1, 44) = 4.60, p = .038$, partial $\eta^2 = .095$) although a Bonferroni correction for multiple contrasts on the five blocks would require a $p$-value at or below .01 to reach conventional levels of significance.

The correlation between WPT block four performance and PIQ was not significantly different from zero, $r = .20, p = .178$. Block four performance did significantly correlate with the output of the language group classification discriminant function, $r = -.29, p = .05$. For the language discriminant function, a more positive result was associated with more impaired language. The correlation indicates that as language ability was more impaired, performance on block 4 was lower.

**The Paired Associate Recognition Task**

The dependent variable for analysis of the PART was the percentage of trials where the participant identified the correct matching card from a field of four, so chance performance on the task was 25%. The mean group performance was well above chance for both groups, with the SLI group mean (SD) of 64% (16%) correct, and the typical group mean (SD) of 77% (18%) correct. The dependent variable was transformed by multiplying the arcsine of the proportion
correct by two, following Ragland et al. (1995). A one-way ANOVA confirmed that the mean for the group with SLI was less than that of the typical group ($F(1,44) = 6.992, p = .011$, partial $\eta^2 = .137$).

Participant PIQ correlated with their transformed PART scores, $r = .46, p = .001$. The output of the discriminant function used for language group classification also correlated with the transformed PART score, $r = -.389, p = .008$.

**Stratification by Performance Intelligence Quotient**

The groups differed by PIQ as well as language ability in this study. For the PART and the WPT, it is possible that the difference in PIQ masked effects due to language ability alone, making it difficult to detect a normal-range declarative memory and deficient procedural memory as predicted by the PDH. The correlation of language ability and PIQ make disaggregating these effects difficult, if not impossible given the intertwined nature of language and cognitive ability. One approach to observing group differences attributable to language ability alone is to select a subset of each group so that the new, smaller groups do not differ on PIQ. This process resulted in a new group of 12 participants with typical language and mean (SD) PIQ of 106 (8), and new group of 15 participants with SLI and mean (SD) PIQ of 102 (6). These new groups did not differ on PIQ ($t(20) = 1.71, p = .10$).

These new groups were compared on WPT and PART results. For the WPT, performance on block 4 was for the typical group, a mean (SD) of .96 (.35); for the group with SLI, a mean (SD) of .71 (.32). The difference did not reach statistical significance ($F(1,25) = 3.74, p = .064$, partial $\eta^2 = .13$). On the PART, the mean recall of the typical group was .75 (.18), and for the group with SLI was .66 (.16). The difference in group performance also did not reach statistical significance ($F(1, 25) = 2.73, p = .11$, partial $\eta^2 = .10$). These null results must be
interpreted with caution, however, given the low power to detect group differences for groups of 12 and 15 participants. For the WPT, the power to detect an effect of the magnitude found for block 4 performance was .45, and for the PART the power to detect a group difference of the observed magnitude was .26. These power estimates were calculated for 2-tailed tests of mean differences, assuming an alpha of .05. Estimates were calculated using G*Power (Faul, 2009; Faul, Erdfelder, Lang, & Buchner, 2007).

**Discussion**

The group with SLI recalled fewer card associations in the PART than did the group with typical language, and the effect size was large (partial $\eta^2 = .137$), whereas there was no statistically reliable group difference for overall performance on the WPT. The greatest group difference on the WPT at block four had an effect size that was smaller than the group difference on the PART. The results do not support a hypothesis that procedural memory, as indexed by the WPT, is more impaired in the group with SLI than is declarative memory as indexed by the PART. Correlations indicated that PIQ was more associated with PART performance than with WPT performance, and that language ability measures were associated with both tasks.

The secondary analysis of the smaller participant groups matched on PIQ resulted in no reliable group differences on either task, although the power to detect group differences was compromised by small sample sizes. Given the power issue, a focus on effect sizes suggests that for both the PART and the WPT, the group with SLI did not perform as well as the group with typical language. The effect size of the group difference was modestly larger on the WPT.

Considering both the primary and the secondary analyses, it is difficult to argue that the group with SLI had a procedural memory deficit in the presence of normal range declarative memory. Individual differences in PIQ appeared to have a greater role in performance on the
PART, but the correlations indicated that differences in language ability played as large a role score variation in the PART as they did in the WPT.
Chapter 6

General Discussion

This study set out to evaluate the PDH, and to better define factors associated with sentence imitation performance in adults with SLI. As expected, sentence comprehension and imitation are complex tasks that involve not only the simple effects of a series of factors, but also interactions of those factors. Before considering the pattern of evidence in more detail, I will first review potential limitations of the study and how those limitations affect the inferences that may be drawn from the data.

Limitations and differences with prior studies

The self-paced listening study failed to find significant general effects for group or argument status across all four sentence conditions. A potential explanation is that the self-paced listening technique is too insensitive to detect the effects of argument status or interactions of group and argument status. A prior study employed self-paced listening to evaluate differential effects of sentence structure across two groups of young children (age 6), one with low and the other with high working memory (Felser, Marinis, & Clahsen, 2003). This study did find a group by sentence structure interaction. Another study using self-paced listening was also successful in finding effects of sentence structure in children’s sentence processing (Booth et al., 2000). So the self-paced listening technique has successfully detected subtle differences in sentence processing for groups likely to have variable performance.

Self-paced listening was able to detect sentence processing differences in the secondary analysis in this study, the analysis focused on noun argument compared to verb adjunct conditions. The significant effects of argument structure replicated findings for typical language
adults using self-paced reading (Schutze & Gibson, 1999). The argument effect in this secondary 
analysis was small. This evidence suggests that self-paced listening is suitable to assessing 
effects of sentence structure on sentence processing, and the null findings for the general effect of 
argument status in the primary analysis were due to the very small magnitude of processing 
differences.

**Methodological differences with prior studies**

There are several other factors that may bear on the null findings in the primary analysis of the self-paced listening task. One factor is the development of the stimuli for this study versus prior studies of argument effects. Prior studies have shown that argument status effects are smaller than other influences on sentence processing, such as plausibility (Kennison, 2002). The norming studies to assess sentence plausibility differed from prior studies. In the current study, sentences were rated as presented in the experimental task. In prior studies, sentences were presented as passive constructions, such as *Many cuts in the staff were announced by the administrator*, rather than *The administrator announced many cuts in the staff, without warning* (Boland & Blodgett, 2006). The initial plausibility rating study conducted for this study found that argument conditions were more plausible than adjunct conditions. To arrive at a set of materials with equivalent plausibility across conditions, sentences were changed so that argument conditions were rendered less plausible and adjunct conditions more plausible. It is possible that for prior studies, the presentation of sentences in passive form resulted in higher plausibility ratings as compared to presentation in active form. If so, this may have resulted in less plausible sentences being included for adjunct conditions in prior studies, resulting in more pronounced combined effects of argument and plausibility.
A second change from prior studies was the adjustment of co-occurrence probabilities. In Boland and Blodgett’s (2006) study, the variance of argument and adjunct co-occurrence probabilities was not equivalent. For this study, the variance across conditions was equivalent, allowing valid comparative assessment of correlations between co-occurrence probabilities and processing times. To obtain equivalent variances, the argument condition variances were reduced, and the adjunct condition variances were increased. This was achieved by substituting argument prepositional phrases with lower co-occurrence frequencies for some sentences with very high co-occurrence frequencies, and substituting adjunct phrases with higher co-occurrence frequencies. These adjustments likely reduced the frequency-related processing advantage for the argument conditions.

The plausibility and frequency adjustments may have interacted. Four of the sentences in the argument conditions with the highest co-occurrence frequencies were also rated individually among the least plausible (below four on a seven point scale). If plausibility has a stronger influence on sentence processing than argument status, these sentences may explain part of the reduction in argument processing advantage, and may suggest an explanation of the unexpected curvilinear relations of co-occurrence frequency and processing time. Constraint satisfaction models of sentence processing view processing times as resulting from multiple interacting factors, as may have been the case here (MacDonald & Seidenberg, 2006).

Prior studies of argument effects in sentence processing have not reported effect sizes (Boland & Blodgett, 2006; Schutze & Gibson, 1999). It is therefore difficult to determine how the size of the argument effects found in this study compare to these prior studies of typical adults. The participation of adults with SLI likely increased the variability of performance over the more homogeneous college student participants in the prior studies, which may also have contributed to the null findings across the four conditions.
Nonetheless, the methodological changes made for this study were important to the study objectives. The use of self-paced listening rather than reading techniques avoided the confound of reading ability differences. The change to the plausibility rating approach resulted in plausibility ratings with greater face validity. Finally, the change to achieve equivalent variance in co-occurrence probabilities was necessary to conduct a valid comparative analysis of correlations between processing times and co-occurrence frequencies. Given the significant findings replicating and extending Schutze and Gibson’s (1999) findings, the absence of significant argument effects in the primary analysis is more likely a valid null result, reflecting very small argument status effects and similar processing speeds of the groups with and without SLI. The significant argument effects and the absence of group effects in the secondary analysis are also likely to reflect very limited or no group differences in processing. The implications of these findings will be considered further after first assessing participant group differences.

**Participant differences**

A second potential limitation of the study was the fact that the participant groups differed not only on language ability but also on PIQ and educational attainment. Seeking equivalent PIQs across groups is valid if SLI is considered an isolated linguistic impairment, and further presupposes that the linguistic system is isolated from other cognitive processes. This position is taken by some explanations of SLI based on linguistic theory (van der Lely, 2005). Information processing theories, in contrast, take the position that cognitive abilities affect language processing (Joanisse & Seidenberg, 2003; L. B. Leonard, 1998). Evaluation of cognitive factors such as working memory and processing speed must assess groups that meet criteria for language impairment, but are allowed to vary on the cognitive abilities under study.
Furthermore, there is considerable evidence that non-verbal cognitive abilities change over the course of development in SLI (Botting, 2005). Children who may have been more similar to their peers on measures of PIQ at young ages tend to have lower PIQ measures at later stages of development. This suggests that language ability affects the pace of development in non-verbal cognitive ability. For these reasons, research on developmental disabilities that artificially constrains PIQ to achieve a match to typical peers may be simultaneously constraining the variation in abilities that are under study (Dennis et al., 2009). The two studies that have provided the best empirical support for classification of adults with SLI and with typical language have recruited participant groups that differed on PIQ (Fidler et al., 2010; Tomblin et al., 1992). In the present study, all adults included in the study were screened to ensure that their PIQ was well above the level for cognitive impairment. The group differences in PIQ then are representative of the populations of young adults with SLI, and those with typical language.

Differences in educational attainment are also to be expected when comparing populations of adults with and without SLI (Carroll & Dockrell, 2010; Johnson et al., 2010). This study sought participants for both groups from populations with similar educational standing. For example, participants with language impairment and typical language were recruited from post-secondary schools in Pennsylvania. Differences in educational attainment arose from recruitment of participants registered by the Iowa longitudinal study. These participants started from the same school settings, but the children with typical language have more often pursued and completed college degrees than have children with SLI. Fourteen of the 23 members of the group with SLI, and 17 of the 23 members of the group with typical language were in the modal category of “some college.” What differed across groups were the 6 members of the typical group who had completed college, and the 9 members of the group with SLI who had not pursued college. In a population-based longitudinal study of individuals with and without SLI, only 3% of adults diagnosed with SLI in kindergarten had completed college at age 25, as compared to
32% of typical controls (Johnson et al., 2010). The group difference in education in this study is representative of the populations of adults with and without SLI.

**Predictions of the Procedural Deficit Hypothesis**

The PDH had predictions for the outcomes of each of the three experiments in this study. The PDH predicted that there would be little difference in argument processing between individuals with SLI and those with typical language since argument processing is supported by an intact lexical/declarative capability (Ullman & Pierpont, 2005). Adjuncts were predicted to be more problematic for adults with SLI, since they are processed syntactically (Boland & Blodgett, 2006; Boland & Boehm-Jernigan, 1998), and therefore rely on a deficient procedural system. Adjunct processing may appear normal for individuals with SLI if they rely on compensatory processing from the declarative system, but if this is the case, normal declarative capabilities should be evident (Thomas, 2005).

**Evidence consistent with the Procedural Deficit Hypothesis**

The clearest evidence supporting the predictions of the PDH was the correlation findings indicating compensatory processing in the self-paced listening task. There were significant correlations between adjusted listening times and co-occurrence frequencies for both argument and adjunct conditions for the group with SLI. This finding supported the prediction of the PDH that individuals with SLI may depend on associative memory for processing language structures that typically rely on rules (Ullman, 2001; Ullman & Pierpont, 2005). Further support for the claim was the finding of correlations for argument condition processing times and no significant correlation for adjunct processing for the typical language group.
The inference that adults with SLI used compensatory mechanisms to process adjuncts received some support in the sentence imitation task in that there was no group by argument interaction. The inference of compensation is more complicated in the case of sentence imitation. First, there was an argument by working memory interaction: adjuncts were particularly difficult for those with more limited working memory. One possible position is that compensation via associative memory is possible to overcome rule-governed processing deficiencies, but not to overcome working memory capacity limitations. This may be the case if working memory capacity is determined by brain physiology, such as synaptic transmission rate (Fry & Hale, 2000). Changes in transmission rate may depend more on maturation, and less on training or experience.

The argument for compensation depends not only on the PDH specifically, but on a view of typical language processing as utilizing both rule-governed and associative memory processes. If language is processed by a unitary associative memory process, as is predicted by information processing views of language (Joanisse & Seidenberg, 2003; MacDonald et al., 1994), then an argument for compensation is more difficult to make. The pattern of results for the typical group, however, better fit the predictions of the Argument Structure Hypothesis, which views argument and adjunct processing narrowly as components of language processed by the dual associative and rule-governed mechanisms (Boland & Blodgett, 2006). The evidence from the self-paced listening task best supports a view of typical-range processing times for the group with SLI being achieved by reliance on a processing mechanism different from that used by the group with typical language.
Evidence less consistent with the Procedural Deficit Hypothesis

The difficulty for the compensation inference comes from the further requirement for falsifying the PDH imposed by Thomas (2005). He suggested that if normal-range processing was achieved by the group with SLI by compensatory use of declarative memory, then there should be evidence of a normal declarative memory system. Processing similar to typical language individuals logically must rely on a system that is as capable, or in the case of a compensatory system, perhaps more capable than normal. Ullman and Pierpont (2005) suggest that if individuals with SLI rely heavily on declarative memory for compensation, then this extensive use may result in a better than normal declarative system. The evidence from the visual procedural and declarative memory tasks does not support this prediction. The performance of the group with SLI was lower relative to the typical group on the PART, the declarative task, than it was on the WPT, the procedural task. This does not support compensation depending on a normal or super-normal declarative memory as the most likely factor accounting for normal range performance on the self-paced listening task.

The evidence for a deficit in declarative memory for the group with SLI in this study was stronger than the evidence for a procedural deficit. This pattern of findings is not consistent with the predictions of the PDH, and is among the first evidence that adults with SLI may have deficits in both declarative and procedural memory. It is consistent with findings of deficits in declarative memory in children with SLI (Lum et al., 2010). The findings of this study are largely consistent with prior findings of procedural or implicit learning deficits in adults with SLI (Plante et al., 2002), although the evidence for a procedural deficit in this study did not meet conventional levels of statistical significance.

The weaker evidence for a procedural deficit in this study may be due to the nature of the WPT. Although the task has clearly shown dissociations of procedural versus declarative
memory in neuropsychological studies (Knowlton et al., 1996), other researchers have argued that there are both explicit and implicit learning elements in WPT performance (Newell, Lagnado, & Shanks, 2007; Price, 2009). If explicit, declarative-memory supported learning was an element of the WPT performance, it seems that better performance would be supported by a stronger declarative memory, a conclusion at odds with the findings of more impaired performance on the PART in this study. It is possible that tasks measuring procedural memory abilities that are either based on sequential pattern learning (Tomblin et al., 2007) or which evaluate statistical patterns in language (Plante et al., 2002) would result in clearer evidence of a procedural deficit in adults with SLI.

Declarative memory findings, according to some researchers, need to account for individual differences in working memory (Lum & Blesses, in press). In this study, CLPT Recall ($r = .45$) and Running Span ($r = .42$) significantly correlated with the transformed PART scores, but did not correlate with performance on the fourth block of the WPT ($p's > .17$). So it is possible that the deficit in the PART for the group with SLI was due to individual differences in working memory. Given findings of working memory and PIQ differences between language groups, and the covariation of these factors with language abilities, controlling for these factors would likely result in a reduced ability to detect true differences, Type II error. The PART was designed specifically to minimize the role of working memory, and in the original development of the task, it was not found to correlate with working memory measures for typical adults (Ragland et al., 1995). Given these features of the PART, the group differences in PART scores may in fact indicate differences in declarative memory with non-causal covariation with working memory measures. To support this conclusion, other measures of declarative memory with even less potential to involve working memory ability are required.

The lack of group by argument interactions in both the comprehension and sentence imitation tasks is problematic for the PDH. Given evidence for deficits in declarative memory for
adults with SLI, it seems unlikely that compensation reliant on a deficient system accounts for the null interaction findings. The stronger evidence is of an argument advantage in processing for both the language ability groups in both the comprehension and production tasks.

**Refinements to the Procedural Deficit Hypothesis**

There are several possible refinements to the predictions of the PDH that may better fit the overall pattern of results from this study. The large differences between groups on measures of working memory are consistent with a procedural memory system deficit (Ullman & Pierpont, 2005), although it is not clear why working memory would be a particular difficulty if the deficit in SLI is in the overall procedural system.

A refined explanation of SLI suggested by this study is that the procedural-grammatical deficit in adults with SLI may be more selective than is predicted by Ullman and Pierpont’s proposal. There was evidence from both the comprehension and production tasks that adjunct processing is not a particular difficulty for adults with SLI. It is also possible that adjunct processing deficits are present in children with SLI (Fletcher & Garman, 1988; Schuele & Tolbert, 2001), but they are resolved at adulthood. Difficulty with adjunct processing is part of a very broad array of deficits in combinatorial aspects of language predicted by the PHD. It is likely that other grammatical difficulties continue into adulthood for those with SLI, such as those predicted by the Agreement Tense Omission Model (Poll et al., 2010) and perhaps the Computational Grammatical Complexity account (Marshall et al., 2007). These deficits would also be consistent with the PDH. A refined view of the PDH must account for how some deficits in the combinatorial aspects of language may be present in adults with SLI whereas others are not. It may be that only certain kinds of combinatorial language tasks remain vulnerable to procedural deficits at adulthood, and others have been successfully compensated.
Complementing the view that the deficits in SLI at adulthood are more selective than a general procedural memory deficit, it may also be that the deficits in language and memory are not confined to the procedural-grammatical system. Deficits in visual declarative memory were evident in this study. Other studies of adults with SLI have yet to test the presence of visual declarative memory deficits, but some studies of children with SLI have found deficits on tasks of verbal declarative memory (Lum & Blesses, in press), even when controlling for vocabulary knowledge (Lum et al., 2010). Other studies have found that implicit procedural learning deficits may have an impact on the lexical knowledge of children with SLI, also contributing to the notion that SLI is likely not a selective procedural-grammatical deficit (Evans et al., 2009). One study reported normal-range visual declarative memory for children with SLI, but this result was obtained only after controlling for PIQ (Lum et al., 2010). Here again, is it unclear how group differences representative of individuals with SLI should affect or not affect measures of declarative memory, particularly given that SLI is widely shown to be the result of both linguistic and cognitive differences.

Support for other accounts of specific language impairment

The results of this study did not directly follow from the predictions of the PDH, but the findings may better align with another theoretical perspective on SLI. The clearest support from the study results is for a processing limitations account, particularly for the limited working memory capacity account. For both the CLPT Recall and Running Span measures, group differences were large. Furthermore, a considerable portion of the group difference in sentence imitation performance was associated with individual differences in working memory.

The findings of working memory deficits in adults with SLI are consistent with prior findings that more complex working memory tasks elicit larger group differences than simple
short-term memory tasks in this population (Isaki et al., 2008; Poll et al., 2010). In adolescents with SLI, working memory deficits appear to be quite general, spanning both visual and verbal material (Evans et al., 2011). A link to the PDH is that the neural network supporting working memory appears to overlap considerably with the procedural memory network described by Ullman and Pierpont (2005) (Awh et al., 1996).

More constrained working memory is thought to have an effect on both language learning and language processing for children with SLI (Ellis Weismer, 1996; Gathercole & Baddeley, 1990). The ability to hold information in mind for processing and long-term memory formation is critical to word learning, as well as grammatical learning. The effect on sentence imitation in this study follows from the analysis of the sentence imitation task: the meaning of the target sentence must be held in mind as the sentence is regenerated (Bley-Vroman & Chaudron, 1994; Potter & Lombardi, 1998). The association between working memory and sentence imitation was more pronounced for longer sentences as the demands placed on working memory were greater.

The effect of working memory capacity on sentence imitation may also follow from a general account of speech production (Levelt, 1989). According to Levelt’s account, there are demands on working memory at the stage of conceptualizing the message the speaker wants to convey. This message is selected from many potential alternatives in natural speech situations, but from a predetermined set of messages in sentence imitation. Speakers with SLI may compensate for their capacity constraints to an extent by slowing their rate of speech production (Tomblin et al., 1992). Or, they may limit the number of propositions they are preparing to convey at any given moment as compared to typical language individuals. The challenge of sentence imitation may be, in part, the fact that the speaker does not choose the number of propositions per utterance. The better performance of participants on long adjunct conditions may follow from the fact that adjunct conditions had fewer propositions to convey, but this was not directly measured in this study.
Refinements to the limited working memory account of specific language impairment

An unexpected finding was the lack of group differences in processing times for the self-paced listening task. Some accounts of working memory capacity suggest that working memory plays an important role in online sentence processing for typical adults (Just & Carpenter, 1992; King & Just, 1991). Recent research has also argued that working memory capacity plays a role in both simple and complex sentence comprehension for children with SLI (Montgomery & Evans, 2009). One distinction that may be important to understanding the role of working memory in sentence comprehension is whether the task is online or offline. An online task is one such as the self-paced listening task in this study where processing speed measures reflect the degree of difficulty with integrating the incoming sentence material with the previous material. An offline task is one such as the Truth Value Judgment task in this study, where the meaning of the sentences is used to make a judgment about whether that meaning matched a picture, a task that takes place after the interpretation of the sentence.

Interpretive versus post-interpretive working memory

Findings for a role of working memory in sentence comprehension have primarily utilized offline tasks. For example, the Montgomery and Evans (2009) study found a relation between working memory measures and the ability of children to select a picture reflecting the meaning of the sentence they just heard. In contrast, a study that employed an online task, cross-modal priming (Marinis & van der Lely, 2007), found no relation between working memory capacity and processing times during the course of sentence interpretation. Montgomery (2000) observed that working memory may play a role in offline tasks using the results of sentence comprehension for children with SLI, but may not play a role in the automatic processes of online
sentence interpretation. This view of differences in working memory effects for online, interpretative processing versus offline, post-interpretive processing has been proposed by Caplan and Waters (1999).

Caplan and Waters (1999) proposed that working memory resources may differ within individuals for initial, automatic language interpretation processes as opposed to more conscious, controlled post-interpretive processes. They argued that adults have had a huge amount of experience interpreting syntactic structures. The result is a specialized capacity for online sentence processing that is distinct from the capacity supporting the conscious, controlled post-interpretive processes.

The Caplan and Waters (1999) account is consistent with the findings of this study. Working memory capacity was found to differ between participant groups, and the capacity measured by working memory tasks was associated with sentence imitation accuracy, arguably a post-interpretive task. These clear working memory effects were not apparent in the self-paced listening task. Relations between working memory and listening comprehension times were not directly assessed, but the absence of a group difference is consistent with a null or small effect of working memory on online sentence comprehension processes. An alternative account is that the materials used in the self-paced listening task were not demanding of working memory resources, so no differences in listening times were related to differences in working memory capacity. Other studies using more demanding structures, however, have also found that working memory capacity was not correlated with online sentence processing. This provides further support for the interpretive versus post-interpretive distinction. In one such study, processing of long distance filler-gap dependencies was not correlated with a working memory measure (Marinis & van der Lely, 2007). This proposal requires direct testing in adults with SLI. If adults with SLI exhibit effects of working memory limitations only for post-interpretive processing, this would be an important refinement to existing working memory accounts of SLI.
Scope of attention measures

A further potential refinement to the limited working memory account of SLI comes from the finding in this study that Running Span resulted in large group differences, and was equivalent to CLPT Recall in explaining variation in sentence imitation accuracy. Cowan and colleagues (2005) have contended that storage and processing measures of working memory, such as the CLPT, are difficult to interpret because the tasks involve concurrent demands for storage and processing. Scope of attention measures, such as Running Span, were designed to measure a simpler construct: the capacity of the focus of attention. It appears that the storage element of working memory is critical to sentence imitation given the equivalent ability of the scope of attention measure to predict sentence imitation accuracy.

Some perspectives on SLI as a deficit in working memory capacity have also focused on storage for verbal information, particularly the capacity of the phonological loop component of Baddeley’s (2000) working memory model. Children and adults with SLI have been shown to have poorer ability to repeat nonwords, a measure of the phonological loop (Clegg et al., 2005; Graf Estes et al., 2007; Poll et al., 2010). The limited capacity of the phonological loop interferes with both word learning and grammatical learning, according to the theory (Gathercole & Baddeley, 1990). Consistent with a focus on storage as the most fundamental deficiency for children with SLI, measures of phonological loop capacity have been shown to be impaired relative to language matched controls and after controlling for differences in the central executive processing component of the Baddeley model (Briscoe & Rankin, 2009).

A finding of a more limited scope of attention for adults with SLI compared to typical language peers differs from prior accounts of limitations in the storage component of working memory, however. The phonological loop component incorporates both storage capacity and rehearsal (Baddeley, 2000). A critical requirement for scope of attention measures is the
suppression of rehearsal (Cowan et al., 2005). Rehearsal and storage may be supported by different neural mechanisms (Awh et al., 1996). The conclusion is that for adults with SLI, the difficulties they have with sentence imitation, and potentially other language tasks, stem from a deficient storage capacity for information to be held for ongoing cognitive operations, independent of their ability to refresh that information by rehearsal.

The outcome of scope of attention measures is the number of information chunks that can be held in the active attention (Cowan et al., 2005). Cowan has argued that for typical adults, that capacity is four chunks (Cowan, 2001), a finding quite consistent with the mean number of digits recalled in the Running Span task by the typical language group in this study (4.18). The mean digits recalled for the group with SLI was about 75% of the mean for the typical group. Cowan (2001) suggested that individual differences in capacity may be related to differences in the ratio of the frequencies of neuronal firing cycles.

A model of short-term memory that supports the role of frequency ratios in brain activity suggests that each memory is held in a set of neurons (Lisman & Idiart, 1995). These sets of neurons fire at a relatively high frequency (the gamma range). The number of memories that can be held in active attention is the number of these firings that take place within a cycle of the theta range, a lower frequency band of brain activity. It may be possible to estimate the physiological constraints on scope of attention by measuring activity in the gamma and theta frequency ranges via electroencephalography (EEG) (Jensen & Lisman, 2001). Recent research on children with and without a family history of language-learning difficulties indicates that working memory and sentence imitation ability at 5 years are correlated with individual differences in the intensity of brain activity in the gamma frequency range (Gou, Choudhury, & Benasich, 2011).

This proposal relating working memory and sentence imitation to the frequency characteristics of neuronal firing has not been tested for adults with SLI. Should it be supported, it would suggest that SLI in adulthood can be traced to a fundamental and general cognitive
processing difference. With respect to performance on sentence imitation tasks, these findings on capacity limits for the scope of attention lead to questions on the size of information chunks bound in sentence imitation tasks.

**Understanding sentence imitation**

One of the main results of the sentence imitation task was the importance of working memory and particularly the interactions of working memory with sentence length and argument condition. The length by argument status interaction indicated that the effect of argument status had the opposite effect for long sentences as compared to short sentences. Whereas argument status eased recall of constituents in short conditions, it hindered recall in long conditions. The implication is that for long sentences, sentences consisting of more adjuncts either permitted recall of more chunks of information, or permitted binding sentence information into larger chunks.

Given that there were common participant groups across long conditions and that long sentences did not differ in length, it appears that sentences with adjunct constituents were bound into larger chunks than sentences loaded with arguments. Additional arguments may result in additional propositions, which were perhaps maintained as separate chunks. Adjuncts may be bound to argument-based propositions, enabling the formation of larger chunks. There is little consensus on the chunk size for speech production planning units, and prior research suggests that developing compelling evidence for a consistent chunk size may be difficult (Levelt, 1989). Research on sentence repetition has explored the trade-off of chunk size and chunk number by assuming that a short clause constitutes a memory chunk (Gilchrist et al., 2008). Further research on sentence imitation and adults with SLI could address whether different sentence materials
enable binding of larger chunks as compared to others, supporting refinement of sentence imitation tasks.

The meaning of sentence imitation varies by sentence length

The results of this study suggest that performance on shorter sentences reveals different information than performance on long sentences. The effects of working memory were more pronounced in long conditions, and the anticipated argument advantage reversed to a small argument disadvantage in long conditions. These findings suggest that sentences in long conditions were beyond the processing capacity of most participants. In such conditions, sentence imitation reveals more about group or individual differences in working memory than do shorter conditions, consistent with the view of the task proposed by Bley-Vroman and Chaundron (1994). These authors suggested that when sentences exceed processing limitations, the sensitivity of the task to difference in the grammatical systems of participants is reduced. This is consistent with the reduction and reversal of argument effects in long as compared to short conditions. For the purpose of distinguishing adults with SLI, however, long sentences still appear to have value. Since limitations in working memory capacity are closely associated with SLI, performance on long sentences may be an indicator of such differences. Furthermore, group did not interact with length, suggesting that group differences beyond working memory capacity influence accuracy for both long and short sentences. The working memory and group effects appear to be the main contributors to the larger effect sizes observed for group differences in long conditions.

Shorter sentences, nearer to or within the working memory capacities of most participants, better reveal differences related to sentence structure (Bley-Vroman & Chaundron, 1994). In the present study, working memory differences were less important, and response to
the argument structure contrast was more important in performance in short conditions. In fact, the argument advantage mitigated the effect of a more limited working memory capacity. Conversely, adjunct status increased the effect of limited working memory. The neural activation of arguments implied by the meaning of verbs encountered in the sentence imitation task may have eased the working memory demands of the task (Boland & Boehm-Jernigan, 1998). Alternatively, sentence imitation accuracy may have been enhanced for short argument conditions because arguments are a higher priority in the sequence of the speech production process than adjuncts (Levelt, 1989).

Other sentence structures may reveal differences that are characteristic of adults with SLI. Candidate structures include those that add complexity to the sentence in the form of long-distance dependencies as predicted by the Computational Grammatical Complexity (CGC) hypothesis (Marinis & van der Lely, 2007; van der Lely, 2005). For sentence imitation, particular difficulties have been observed in the performance of adolescents with SLI for sentences with object relative clauses, a structure that introduces a non-local dependency (Riches et al., 2010). The sentences in the Riches et al. study averaged approximately 11 words in length, very similar to the short condition in the current study.

Together these findings suggest that a sentence imitation task may reveal more information about participants by having both shorter and longer sentences. The shorter condition reveals whether the participant has particular difficulties with given syntactic structures, whereas the longer conditions reveal whether participants have more limited working memory capacity. For adolescents and adults, sentences of approximately 11 words appear to be a length appropriate for revealing differences in the grammatical knowledge or grammatical processing abilities of participants, although it may be possible to assess working memory capacity and tailor a sentence length to participant capacity (Jefferies et al., 2004). Sentences of 16 words are beyond the processing capacity of many participants, but were appropriate in this study and that
of Allen and Baddeley (2009) to avoid ceiling effects for typical adult participants. Sentences at this length were discriminating between groups because of working memory capacity effects as well as other differences captured by the group classification process.

An implication of this study and other recent studies of sentence imitation is that to obtain information from sentence imitation, particularly longer stimuli, the scoring approach must reveal differences even when participants make multiple errors (Allen & Baddeley, 2009; Riches et al., 2010). Many sentence imitation measures on norm-referenced tests of language are scored by either a correct/incorrect method, or by a method that results in a 0 score after three or four errors (Hammill et al., 1994; Semel et al., 2003). The application of this latter method to this study resulted in floor effects. The use of the percentage of words recalled provided information that discriminated between adults with and without language impairment.

**Argument preference effects in sentence comprehension**

In addition to tests of the PDH, this study also explored theories of online sentence processing. At the more general level, the study evaluated claims that sentence processing is guided by lexical factors on the one hand (the Argument Preference Strategy) (Schutze & Gibson, 1999) or a preference for simpler syntactic structure (Minimal Attachment) (Frazier, 1987) on the other. The study replicated support for the Argument Preference Strategy across the same noun argument/verb adjunct conditions as those evaluated by Schutze and Gibson (1999). These findings were independent of group differences. Schutze and Gibson (1999) suggested that arguments categorically differ from modifiers (adjuncts) because of the meaning relations between words and their arguments.

At the more specific level, the study evaluated two explanations for the lexical effects of argumenthood. Boland and Blodgett (2006) proposed the Argument Structure Hypothesis, which
predicted that a word’s arguments were part of its lexical entry, whereas adjuncts were not. An alternative explanation for the lexical effect of argumenthood was the Pure Frequency Hypothesis proposed by MacDonald and her colleagues (MacDonald et al., 1994). The frequency-based hypothesis was that arguments co-occur with their associated verbs or nouns more frequently than do adjuncts. Both arguments and adjuncts are part of the lexical entry.

To distinguish the categorical view of argumenthood suggested by the Argument Structure Hypothesis from the frequency-based view suggested by the Pure Frequency Hypothesis (Boland and Blodgett, 2006), prior researchers developed measures of co-occurrence frequency for prepositional phrases and the nouns or verbs to which they were attached (Boland & Blodgett, 2006; Schutze & Gibson, 1999). Prior researchers were not able to present clear evidence favoring either the Argument Structure or Pure Frequency hypotheses. Boland and Blodgett found significant correlations between reading times and co-occurrence frequencies only for verb argument conditions, but not for noun argument conditions. Finding no significant correlation for noun argument conditions was contrary to the Argument Structure Hypothesis. Differences in variance between argument and adjunct frequencies also made it difficult to draw conclusions favoring either theory. Schutze and Gibson (1999) found correlations for neither noun argument nor verb adjunct conditions.

In the current study, significant correlations were found for noun argument conditions after accounting for the curvilinear relationship between frequencies and listening times, but no frequency effects for verb adjunct conditions for the typical language group. It is possible that the prior null findings for frequency effects in noun argument conditions were due to curvilinear relations, since a lack of linear relation may result in null correlation findings even when there is a systematic relationship between variables. What the current study adds to prior findings on typical adult sentence processing is support for the Argument Structure Hypothesis (Boland & Blodgett, 2006) or any view of arguments as categorically different from adjuncts. The evidence
suggests that argumenthood is not reducible to frequency effects, and that the meaning relations between nouns and their arguments influences sentence processing. This is counter both to views that initial sentence parsing is driven by structural preferences (Minimal Attachment) (Frazier, 1987) and to views that the parser simply responds to the frequency differences between arguments and adjuncts (Pure Frequency Hypothesis) (MacDonald et al., 1994).

While the Argument Structure Hypothesis (Boland & Blodgett, 2006) was supported by the data for typical adults, the theory does not account for a curvilinear relation between frequency and listening times for arguments. The Schutze and Gibson (1999) position was that argumenthood is one influence on initial sentence parsing preferences. This is in line with broader interactive views of sentence processing that view parsing preferences as arising from a range of influences, including lexical, pragmatic, frequency, and plausibility factors (MacDonald & Seidenberg, 2006; van Gompel & Pickering, 2007). The data from the present study suggest that influences on argument prepositional phrase processing include co-occurrence frequency, but that influences beyond frequency may arise for frequencies in different ranges.

Whereas interacting influences on sentence processing may explain the curvilinear relation of frequency to processing times in argument conditions, they do not provide an explanation for the absence of frequency effects in adjunct conditions. The dual-route language processing view of the PDH suggests that argument processing is lexical, using associative memory processes, whereas adjunct processing is syntactic, using rule-based processes (Ullman, 2001). In considering other instances where proponents of dual route language processing have used the presence or absence of correlations to argue for a split of association-based and rule-based processing, Ellis (2002) argued that the lack of correlations for rule-governed processing (as in adjunct processing in the current study) could be attributed to the power law of learning. This law suggests that over time, very frequent language structures will be learned to automaticity, resulting in performance at an asymptote. This asymptotic performance makes
frequency effects difficult to find. Ellis’s argument is difficult to make in the case of the present study, where the co-occurrence frequency of adjuncts is lower as a category than arguments. The evidence in the present study better supports a categorical difference in processing for arguments and adjuncts.

**Implications for professional practice**

There are several implications of the findings of this study for adults with SLI and the speech-language pathologists and disability services professionals who work with them. First, the study suggests that slower processing is not always a hindrance to performance on language tasks. Second, the findings confirm that working memory limitations are an important part of the profile of SLI at adulthood. The study adds to our understanding of what elements of working memory are most critical for performance of language tasks. Third, the study evaluated a theory on the causes of SLI, the PDH. The study provided support for an important element of the theory, the role of compensatory processes in developmental language impairment. Finally, the study provided information that can aids professionals who use sentence imitation tasks as part of the language assessment process.

**Slower processing not always a hindrance to language task performance**

Slower processing of language tasks is characteristic of individuals with SLI from childhood to young adulthood (Miller et al., 2001; Miller & Poll, 2009). It is not clear however when being a slower processor hinders language task performance. In adults with SLI, slower processing has been correlated with a poorer ability follow spoken directions (Miller & Poll, 2009). In this study, however, adults with SLI did not differ in how long they listened to critical
parts of sentences in a comprehension task. Slower processing also did not uniquely explain performance on sentence imitation after accounting for working memory and other factors.

There is evidence for children that processing speed and working memory capacity are not always competitive (Towse, Hitch, Horton, & Harvey, 2010). In some tasks, slower processing—taking the time to reconstruct a sentence—may actually help individuals remember more information. Consider a task where participants have to listen to a sentence, decide if the sentence is true, and then remember the last word of the sentence. When the time comes to recall the last word, reconstructing the earlier part of the sentence may provide useful clues to the word they have to recall. For example, if the participant recalls Buses have..., they may more easily recollect that the sentence final word was wheels. It is possible that in sentence imitation there may a similar advantage to slower processing. The implication for clinical practice is that if better language performance is the goal, achieving faster processing times may not always be required to reach that goal.

**Working memory deficits in adults with specific language impairment**

More limited working memory capacity hindered language task performance in this study. Both measures of working memory (the CLPT and Running Span) indicated that the group with SLI had considerably more limited working memory capacity than the group with typical language. This more limited capacity was an important factor in explaining sentence imitation performance.

The findings of this study suggest that working memory limitations are an important factor in the ongoing difficulty that adults with SLI have with language tasks. What this study adds is that this working memory capacity gap did not result in significant group differences for online sentence processing. Working memory limitations may not affect immediate
comprehension during sentence processing, but instead affect how well adults with SLI can maintain information gleaned from online sentence comprehension to perform some additional task.

This study also indicates that a somewhat different conception of working memory, the scope of attention (Cowan et al., 2005), may be a valid way of understanding the nature working memory capacity limitations in this population. The group differences on the Running Span scope of attention measure indicate that the capacity to store information in active attention is a fundamental deficit hindering language task performance in adults with SLI. This conception of working memory differs from others documented in the literature on SLI in that it does not incorporate other skills that may be associated with working memory. For example, the scope of attention is not dependent on task switching skills, speed of subvocal rehearsal, or the ability to efficiently carry out a processing task in parallel with the storage of information (Baddeley, 2000; Just & Carpenter, 1992). This finding suggests that any efforts to remediate or compensate for working memory limitations may best focus on limitations in storage capacity as measured by scope of attention tasks.

Strategies to support adults with SLI would appropriately focus on ways to compensate for this more limited storage capacity. To enhance functioning, strategies that permit adults to create larger memory chunks may help in situations where auditory working memory is required. For example, strategies that provide mnemonics or associations among items to be maintained in memory are likely to be helpful. Alternatively, compensatory strategies that reduce the load on auditory working memory are likely to be helpful. For example, having access to written lecture notes, or having visual supports available may assist learning and functioning in post-secondary and vocational environments.
An explanation of Specific Language Impairment incorporating compensation

Adults with SLI processed argument and adjunct structures at speeds very similar to typical language adults in this study. In sentence imitation, being a member of the group with SLI did not result in particular difficulties with adjunct processing. This suggests that adults process and comprehend adjunct phrases at a level very similar to that of typical language adults, after accounting for more limited working memory.

The implication is that adjunct processing may be an unusual instance of a structure that is delayed in children with SLI, but is processed at typical levels by the adult stage of development. Many of the structures or abilities found to be deficient in children with SLI continue to pose problems for adults with SLI. For example, adults with SLI differ from typical adults in processing tense and agreement (Poll et al., 2010), in rapid temporal processing (Tomblin et al., 1992), and working memory capacity. Adjunct processing appears to be delayed in childhood (Fletcher & Garman, 1988), but processed typically at adulthood.

The PDH predicts that adults with SLI may compensate for their difficulties with syntactic processing by relying more heavily on declarative memory (Ullman & Pierpont, 2005). Few prior explanations of SLI have considered a role for compensatory processes in SLI (Thomas, 2005). This study provides evidence that adults with SLI may successfully use an associative memory process to perform tasks that do not rely on associative memory in typical language adults. This has several important consequences for practitioners. First, a greater reliance on associative memory can allow adults with SLI to successfully reach typical range performance on some language tasks. Language interventions in children sometimes rely on making the rules of syntax or the meaning of language structures explicit (Ebbels, van der Lely, & Dockrell, 2007). This findings of this study and the PDH suggest that this reliance on explicitly gained knowledge and associative memory is an important part of language
development in adults with SLI. Second, the correlational findings from the self-paced listening task indicate that stronger associations are created for more frequently encountered structures. This supports the importance of input frequency in language development in SLI. If stronger associations are formed for more frequent structures, then increasing the exposure frequency for language structures should result in better learning and performance.

From a practical perspective, the role of frequency in language development in individuals with SLI underscores the importance of input frequency in language intervention. When a particular language structure is problematic for an individual with SLI, successful learning and generalization may depend on having sufficient exposure to or practice with the structure. This follows directly from the PDH. If individuals with SLI have difficulty implicitly learning language rules or sequential patterns, the approach to advancing their language abilities may rely on associative memory and explicit training. Successful learning and performance through associative memory is enhanced with more frequent exposure to the target structure.

A final implication for the typical range processing of adjuncts by adults with SLI is that these structures need not be avoided in preparing communications or intervention materials for this population. For professionals developing interventions for individual with SLI, a new or problematic language structure is best introduced in the context of language that is well within the grasp of the person with the disorder (Masterson, 1997). Use of adjunct phrases in sentences causes no particular difficulties for adults with SLI.

Understanding factors affecting sentence imitation performance

For language assessment, this study provides information that may assist professionals in interpreting adult performance on sentence imitation. Consistent with prior findings (Bley-Vroman & Chaudron, 1994), sentences in the short conditions (11 words) were revealing of the
grammatical systems of participants. These conditions clearly demonstrated the argument advantage, and the effect was in the expected direction. Other studies indicate that performance on sentences at this length, presumably within or near the capacity of working memory, accurately indicate the sentence structures that are more difficult for adolescents and adults with SLI (Riches et al., 2010).

Professionals may interpret sentence imitation results differently by sentence length. For sentence lengths centering on 11 words, a consistent pattern of errors on sentences incorporating a given syntactic structure will likely indicate that that structure represents a weakness in language knowledge or processing ability for the participant. For sentences significantly longer than 11 words, such as those of 16 words used in this study, poor performance is likely less indicative of problems with a particular sentence structure. Instead, deficits on these sentences relative to norms are likely to indicate working memory capacity limitations.

Certainly other factors should be considered in the profile of adults with SLI, such as vocabulary knowledge and phonological (spelling) abilities. These factors also enter into sentence imitation performance, as indicated by the general effect of group in sentence imitation.

**Future directions**

This study provided evidence for the predictions of the PDH, and identified factors associated with sentence imitation performance. It suggested a number of refinements to explanations of SLI at the adult stage of development which require additional validation.
Further testing of the procedural deficit hypothesis

This study raised questions on some predictions of the PDH, one of which was the selectivity of the deficit to the procedural system. The validity of the PDH does not require findings of a normal declarative system for individuals with SLI (Ullman & Pierpont, 2005), but it was among the predictions of the theory. The findings of this study were that, for the group with SLI, performance on the PART was more impaired relative to the typical language group than was performance on the WPT. A question raised by recent literature on the PDH is whether performance on declarative tasks should account for group differences in working memory (Lum & Blesses, in press). A separate question is whether the WPT is a pure procedural measure. Future research is needed to clarify whether SLI in adulthood involves a selective procedural deficit by assessing adults on a range of procedural and declarative memory tasks. The declarative tasks should ideally be designed to avoid confounds with working memory. Important to these future studies is presenting the same participant groups with both types of memory tasks.

A second important direction is testing the performance of individuals with SLI on language tasks for which the PDH would predict a dissociation. In this study, the dissociation was for normal argument processing but impaired adjunct processing. A further contrast that may meet these requirements is that of lexical retrieval versus lexical recognition (Ullman & Pierpont, 2005). Lexical retrieval is predicted to be subject to deficits in the frontal-basal ganglia circuits supporting procedural memory, whereas lexical recognition is not. Well controlled tasks of lexical recognition would be predicted to be normal in adults with SLI, whereas tasks requiring word finding are predicted to be impaired.

Also of interest is the view in the PDH that individuals with SLI compensate by use of the declarative system (Ullman & Pierpont, 2005). This study found evidence of compensation in
the processing of adjuncts. The unexpected non-linear relations of listening time and co-
ocurrence frequency, and the findings of less than normal declarative memory, made it unclear
whether the adjunct processing was via compensatory mechanisms, or whether frequency effects
would be found in other circumstances. Exploration of situations where adults with SLI engage
in compensatory processing is of particular interest, since development of functional language
ability may depend on how well adults compensate for language or cognitive abilities that have
reached the limits of developmental growth.

**Extending new findings on working memory in specific language impairment**

The finding that adults with SLI perform quite differently from typical language adults on
a scope of attention measure suggests that the working memory model of Cowan and colleagues
(2005) provides new insight into the nature of SLI at adulthood. To strengthen the conclusion
that the capacity of active attention is an important factor in explaining SLI, these results require
replication, both with the Running Span measure on new participant samples, as well as by
comparing Running Span to other measures designed to measure the scope of attention.

A further extension of this research is to understand whether scope of attention measures
predict variation in other language tasks—beyond sentence imitation—for adults with SLI. Of
clinical and theoretical significance is the relation of scope of attention to tasks functionally
important for adults with SLI, such as reading comprehension, lecture understanding, and
discursive writing. A fundamental question is whether either aspect of working memory capacity
is malleable, chunk capacity or chunk size. Another question is whether the equivalence of
storage and processing and scope of attention tasks holds when tasks have greater processing
demands, rather than simply greater storage demands.
A more basic question for this research is whether patterns in the frequencies of neuronal firing explain the capacity of the focus of attention (Jensen & Lisman, 2001; Lisman & Idiart, 1995). Building on emerging evidence that characteristics of the high frequency (gamma) firing patterns are associated with early language difficulties (Gou et al., 2011), EEG studies of adults with SLI might assess whether resting state firing patterns, or patterns during conditions demanding of working memory differ from adults with typical language.

A second line of working memory research suggested by this study is the potential distinction between interpretive and post-interpretive processing (Caplan & Waters, 1999). This study did not directly address the role of working memory in online sentence processing. The data were consistent with a view that working memory capacity, although quite different between groups, did not result in differences in online processing times. It may be that sentence structures that are specifically manipulated to differentially tax working memory, such as subject relative versus object relative clauses (King & Just, 1991), would reveal a role of working memory for the interpretive phase of sentence processing.

**Processing speed and sentence imitation**

Processing speed measures did not predict sentence imitation accuracy after accounting for working memory and other factors in this study. The sentences had simple syntactic structures. Evidence from studies of processing speed differences between groups with SLI and those with typical language suggest that tasks that involve more cognitive operations reveal greater gaps in performance (Miller et al., 2001; Miller et al., 2006). This follows from the view that response time is slowed by a common factor across different types of cognitive operations (Kail, 1994). If this view holds, it may be that a sentence imitation task involving sentences
requiring more cognitive operations than those used in the present study would reveal performance differences associated with slowed processing in adults with SLI.

The kinds of sentences that involve more cognitive operations, or planning units, remain unclear (Levelt, 1989), but one approach would be to consider the clause as a planning unit for speech production (Gilchrist et al., 2008). If this is the case, then sentences with embedded clausal units may require more operations for production, taxing the capacity of participants with slower processing speeds. Embedding is an important measure of complexity in sentence processing (Quirk et al., 1985). Observing the association between processing speed and sentence imitation accuracy for complex sentences, or observing effects of speeded processing requirements on complex sentence imitation, may reveal a potential connection between processing speed and sentence imitation performance.

**Online sentence processing of adults with specific language impairment**

This study found no group differences for sentence processing times in the self-paced listening task. This is perhaps the only study to apply the self-paced listening technique to adults with SLI. The finding of no group difference should be replicated with this and with other techniques with the potential to be sensitive to variation in online sentence processing to strengthen the conclusion that the argument-adjunct contrast does not present particular difficulties for adults with SLI. Additional research should consider replication of this task with younger individuals with SLI, given the possibility that adjunct processing abilities are delayed for children with SLI.

An additional direction is to test sentence processing effects for structural variations that may have stronger effects on processing for adults with SLI, such as structures involving long-distance dependencies (van der Lely, 2005). Explorations of structures that are predicted to cause
problems would add to the very limited information on sentence processing in adults with SLI, and would also further validate the sensitivity of the self-paced listening technique for this population. Findings from self-paced listening may be validated by comparison with other techniques that may also be suited to studies of sentence processing effects in adults with SLI, such as event-related brain potentials, the visual world method (Trueswell & Gleitman, 2007) and cross-modal priming techniques (Marinis & van der Lely, 2007; Seiger-Gardner & Schwartz, 2008).

**General conclusions**

An early study of adults with SLI observed that the language differences in this population were revealed by tasks at the sentence level that placed demands on processing and retention of information (Tomblin et al., 1992). The results of this study reinforce that observation. Sentence imitation, a task that clearly places demands on information retention, resulted in quite different performance for the affected and unaffected groups. Listening times during the course of processing structurally simple sentences did not show group differences.

The lack of group differences in the self-paced listening processing times may indicate one of the language processing areas where adults with SLI are relatively capable. These results were surprising given the evidence that children with SLI perform differently than typical language peers in producing adjuncts. If studies of children with SLI using well controlled tasks confirm that adjunct processing is problematic, this may represent an area of delay, rather than a deviation from typical language development in SLI.

The similar processing times in the self-paced listening task masked differences in the way that adults with SLI processed adjunct constituents. Whether this is strong evidence of compensation is unclear, given the assumption that adults with SLI were compensating by relying
on normal declarative memory (Ullman & Pierpont, 2005; Thomas, 2005). Even the limited evidence for compensatory processing is important because compensation has received so little attention in the literature on developmental disabilities (Thomas, 2005). Finding an ability that is similar in adults with SLI and adults without a history of language disorders suggests a way to explore adaptive compensation. An understanding of adaptive compensation among adults with developmental language disabilities could be very helpful, particularly in understanding why some adults with a history of language disorders make such adaptations and some do not.

This study challenged the predictions of the PDH, but the results also highlight issues in falsifying any theory explaining SLI. Researchers need to debate the value of rigorous matching or statistical control to rule out potentially confounding effects from outside the language domain. This position must be considered in light of the reality that individuals with developmental disabilities, particularly those at later stages of development, have had different experiences that make elegantly matched affected and unaffected groups unrealistic.

Finally, this study demonstrates the complex nature of sentence imitation performance. Sentence imitation appears to be a simple matter of listening and remembering. It has been used as a language research and assessment method for at least 40 years (Slobin & Welch, 1971). The fact that it differentiates language ability groups is well documented, but the opportunity is to understand the task so that it can also reveal more about the nature of the language and cognitive systems of participants. This study demonstrated the different kinds of information revealed by sentences of different lengths and structures. Further, it demonstrated that while working memory is an important factor in the task, sentence imitation is mischaracterized if it is reduced to a measure of short-term or rote memory. Performance clearly depends on multiple factors. Factors in the structures of the sentences interact with characteristics of the participants. Careful tailoring of the task can lead to measures that contribute to an understanding of the language and cognitive systems of participants.
### Appendix A

#### Self-Paced Listening Materials

Self-paced listening materials consisted of the following sentences. Conditions in each set of four are, in order of listing here, 1 verb argument condition, 2 direct-object noun argument condition, 3 verb adjunct condition, and 4 noun adjunct condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>No.</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>The committee hid the secretary’s inclusion from the boss very easily.</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>The committee hid the secretary’s inclusion in the process very easily.</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>The committee hid the secretary’s inclusion in their expressions very easily.</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>The committee hid the secretary's accident with the coffee very easily.</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>The leader withheld his support from the candidate without warning.</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>The leader withheld his support for the candidate without warning.</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>The leader withheld his support for the moment after all.</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>The leader withheld his comment about the candidate without warning.</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>The two friends revealed their tattoos to the woman without concern.</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>The two friends revealed their delight in the story without concern.</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>The two friends revealed their tattoos at the party without concern.</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>The two friends revealed their tattoos on their arms without concern.</td>
</tr>
<tr>
<td>1</td>
<td>13</td>
<td>The busy shopper expressed his interest to the manager before leaving.</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>The busy shopper expressed his interest in a wallet before leaving.</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>The busy shopper expressed his interest through his smile before leaving.</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>The busy shopper expressed his complaint about the service before leaving.</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
<td>The new guard concealed his surprise from the visitor with difficulty.</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>The new guard concealed his surprise at the decision with difficulty.</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>The new guard concealed his surprise at the time with difficulty.</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>The new guard concealed the book without a cover, with difficulty.</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>The losing team gained some respect from the coach, once again.</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>The losing team gained some respect for the sport, once again.</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>The losing team gained some respect in the end, once again.</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>The losing team gained some respect under the circumstances, once again.</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>The witness admitted the error to the police under questioning.</td>
</tr>
<tr>
<td>2</td>
<td>26</td>
<td>The witness admitted the error in his statement under questioning.</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>The witness admitted the error in the morning under questioning.</td>
</tr>
<tr>
<td>4</td>
<td>28</td>
<td>The witness admitted the error by the expert under questioning.</td>
</tr>
<tr>
<td>1</td>
<td>29</td>
<td>The physician suggested a remedy to the traveler this morning.</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>The physician suggested a remedy for the flu this morning.</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>The physician suggested a remedy at the station this morning.</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>The physician suggested a remedy with good evidence this morning.</td>
</tr>
</tbody>
</table>
1 33 The speaker described people’s alienation to the class, in detail.
2 34 The speaker described people’s alienation from the system, in detail.
3 35 The speaker described people’s alienation from his perspective, in detail.
4 36 The speaker described people’s alienation in the south, in detail.
1 37 The environmental agency offered some exemptions to the business this year.
2 38 The environmental agency offered some exemptions from the law this year.
3 39 The environmental agency offered some exemptions from the start this year.
4 40 The environmental agency offered some exemptions over a thousand dollars this year.
1 41 The teacher communicated much enthusiasm to the children during class.
2 42 The teacher communicated much enthusiasm for the calculators during class.
3 43 The teacher communicated much enthusiasm over the week during class.
4 44 The teacher communicated much enthusiasm from the faculty during class.
1 45 The police detective proposed an investigation to the captain, right away.
2 46 The police detective proposed an investigation of the company, right away.
3 47 The police detective proposed an investigation for a day, right away.
4 48 The police detective proposed an investigation through the city, right away.
1 49 The lawyers delivered employee demands to the president, despite criticism.
2 50 The lawyers delivered employee demands for a raise, despite criticism.
3 51 The lawyers delivered employee demands for a while, despite criticism.
4 52 The lawyers delivered employee demands from last week, despite criticism.
1 53 The administrator announced many cuts to the reporter, confirming rumors.
2 54 The administrator announced many cuts in the staff, confirming rumors.
3 55 The administrator announced many cuts in the meeting, confirming rumors.
4 56 The administrator announced many cuts of some significance, confirming rumors.
1 57 The famous author mailed the revisions to the editor last week.
2 58 The famous author mailed the revisions of the poem last week.
3 59 The famous author mailed the revisions for his wife last week.
4 60 The famous author mailed his revisions from the draft last week.
1 61 The producer kept the conversion in his notes, under wraps.
2 62 The producer kept the conversion of the Americans, under wraps.
3 63 The producer kept the conversion until morning, under wraps.
4 64 The producer kept the conversion at the ranch, under wraps.
Appendix B

Sentence Imitation Materials

Sentence imitation materials consisted of the following sentences. Conditions were 10 short/argument loaded, 20 long/argument loaded, 30 short/adjunct loaded, 40 long/adjunct loaded.

<table>
<thead>
<tr>
<th>Condition</th>
<th>No.</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1</td>
<td>The worker carved the wood with his knife but denied the evidence.</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>The widow painted the shed and frightened the mice with her hammer.</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>The president urged the army to act and blocked their concerns.</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>The observer located the park and taught the kids the method.</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>The soprano played the passage and read the joke to the public.</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>The professor assigned the woman the verse to check her vision.</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>The dentist baked her assistant a cake and added a cherry.</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>Our grandmother injured her thumb yet knit our mother new linen.</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>My brother studied the item then filled the bowl with liquid.</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>This generation tempted sin and grasped success with both hands.</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>My cousin ate an orange and inserted the seeds in sand.</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>The principal saddened the child but coaxed him into class. The gentleman amused the earl with his description and impressed his servant but his cat frightened the mouse.</td>
</tr>
<tr>
<td>20</td>
<td>13</td>
<td>The navy crushed the resistance and chased the enemy then guarded the republic with heavy guns. Management chose a plan and excited loyalty but entertained the executives with concerts.</td>
</tr>
<tr>
<td>20</td>
<td>14</td>
<td>The amateur chopped the stump with an axe while professionals estimated his ability.</td>
</tr>
<tr>
<td>20</td>
<td>15</td>
<td>The teacher typed his welcome and drew a village with chalk then shut his eyes and designed his lecture.</td>
</tr>
<tr>
<td>20</td>
<td>16</td>
<td>The graduates delighted in the trip but annoyed my aunt while the minister adjusted the tent.</td>
</tr>
<tr>
<td>20</td>
<td>17</td>
<td>The school required a loan and asked their banker for a note but the executive required assurance.</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
<td>The girl mentioned the crime and expected an exception so the company attacked the evidence. My grandfather added the stables and paid for the ferry but accepted assistance from the men.</td>
</tr>
<tr>
<td>20</td>
<td>19</td>
<td>Our candidate invited the nation to a fight but lost the primary and conceded defeat.</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>
The committee allowed the amateur a fight but covered his flesh and emphasized good judgment.

Management praised the vision and permitted an exception but the company shut the gallery.

The genius realized the date at dawn and wrote in the cold.

My patient protested in the bath and dripped down the long hall.

The queen visited before the judgment but prayed at great length.

People hurried at the accident but crumbled at the main square.

The author cheated the agency at first and bragged in public.

The dog struggled in the pet motel and jumped at the signal.

The staff tired before the farewell and complained in the cellar.

The driver rushed for the fat priest but crashed on the dreaded curve.

Our guide worried at the emergency but decided later.

The specialist froze at the call then replied in confusion.

The competition melted at our strength but rose after the snub.

The corporal flew at noon over the sound and drove in a rush.

The student sailed for miles on the ocean but drifted to the west while his brother sang quietly.

The physician proposed the cure in his statement but his patient swore at the complication.

My uncle disclosed the scheme during the call but the company cheated as usual at the port.

Congress attempted the deal in the winter but objected after the primaries and revolted.

The government refused the drugs from the border base but the army swayed after the clash at the harbor.

The lady confessed in a whisper after the trial while the household grieved in the afternoon at church.

The natives waited at the platform while the military marched by in respect after the concert.

The secretary disclosed the bet in the afternoon and still raced the next morning at the brick track.

The soldier rotated to a new danger in the summer and walked for the season without a coat.

Council declared its purpose despite the emotion and concluded construction after the election.

The beauty dreamed of romance during vacation but her family moved after the church holiday.

The couple confided a breakdown in the spring and surrendered their house in the valley after the split.


VITA

Gerard H. Poll, III
ghp110@gmail.com

Education

2012  Doctor of Philosophy, Communication Sciences and Disorders
The Pennsylvania State University, University Park, PA

2007  Master of Science, Communication Sciences and Disorders
The Pennsylvania State University, University Park, PA

1987  Master of Business Administration
The University of Chicago, Chicago, IL

1982  Bachelor of Arts, English Language and Literature, with Honors
The University of Chicago, Chicago, IL

Publications


Selected Research Support

2010 – 2012  National Research Service Award (F31).
National Institute on Deafness and Other Communication Disorders
1F31DC010960 *Sentence Processing Factors in Specific Language Impairment*