Impact of Working Memory Constraints on Speech Monitoring in Healthy Children

by

Tanya Lentz
BA, University of Winnipeg, 2001
MA, University of Victoria, 2003

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTORATE OF PHILOSOPHY

in the Department of Psychology

© Tanya Lentz, 2013
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.
Impact of Working Memory Constraints on Speech Monitoring in Healthy Children

by

Tanya Lentz
BA, University of Winnipeg, 2001
MA, University of Victoria, 2003
Abstract

Supervisory Committee
Dr. Kimberly Kerns, Department of Psychology
Supervisor
Dr. Mauricio Garcia-Barrera, Department of Psychology
Departmental Member – Revision Supervisor
Dr. John Walsh, Department of Educational Psychology and Leadership Studies
External Member
Dr. Leslie Saxon, Department of Linguistics
Additional Member

Abstract

The purpose of the current study was to examine the impact of working memory on speech monitoring processes in the primary language of school-age children using the framework of Levelt’s Perceptual Loop Theory of speech production (1983). A community sample of eight children aged 6-8 and fourteen children aged 10-12 completed 4 verbal description tasks under different conditions; control, working memory load, white noise and combined working memory load and white noise. Participants also completed measures of listening span, digit span and spatial span. The results indicate that with increasing working memory load, children make significantly more speech errors, silent pauses and repetitions. No relationship was found between working memory and total repairs per errors or between working memory and total number of editing terms used. Group differences across the conditions were not significant; however, age-related trends were notable. Younger children had greater difficulty monitoring their speech with the introduction of working memory load; whereas, older children had greater difficulty with the introduction of white noise. A revised speech production model incorporating aspects of working memory is recommended and implications for clinical populations are discussed.
# Table of Contents

Supervisory Committee ........................................................................................................ ii
Abstract ................................................................................................................................... iii
Table of Contents .................................................................................................................. iv
List of Tables .......................................................................................................................... v
List of Figures ........................................................................................................................ vi
Acknowledgments ................................................................................................................... vii

Impact of Working Memory Constraints on Speech Monitoring in Healthy Children ...... 1

Chapter 2 ................................................................................................................................. 41

Chapter 3 ................................................................................................................................. 57

Chapter 4 ................................................................................................................................. 75

References ................................................................................................................................. 96

Appendix A ............................................................................................................................... 124

Appendix B ............................................................................................................................... 127

Appendix C ............................................................................................................................... 128

Appendix D ............................................................................................................................... 129
List of Tables

Table 1. Intraclass correlation values per network with 95% confidence intervals........ 55

Table 2. Language and Schooling Demographic Information for Children Aged 6-8 and 10-12 as Frequency Counts........................................................................................................ 58

Table 3. Language samples raw score characteristics per groups ................................ 59

Table 4. Means and standard errors of raw scores of repetitions and editing terms of language samples. ........................................................................................................ 69

Table 5. Time characteristics of language samples. .................................................... 72

Table 6. Raw scores on working memory tasks by group........................................... 74
List of Figures

Figure 1. Spatial Span Board ........................................................................................................ 46
Figure 2. Sample network of network task ..................................................................................... 49
Figure 3. Total errors per group per condition .............................................................................. 63
Figure 4. Total repairs per errors per group per network condition ............................................ 64
Figure 5. Total transformed pauses ((network condition pauses + 0.5)^{-1/8}) per network condition per group. .............................................................................................................. 66
Figure 6. Total transformed repetitions (square-root) per network condition per group .. 68
Figure 7. Model of the role of working memory in speech monitoring. ................................. 90
Acknowledgments

First, I would like to thank the Coast Salish people for allowing me to live and work on their traditional lands. I further wish to thank all of the children and families who volunteered their valuable time and effort to support this project. Without our volunteer participants, psychology research could not continue.

I also acknowledge my dissertation committee for all their efforts in guiding this project. In particular, I would like to thank Dr. Garcia-Barrera for taking on the responsibility of my final revisions. As well, I would like to recognize my volunteer research assistants, Brook Parlby, Kasia Gwiazda, Jamie Piercy and Branda Guan for their hard work in assisting with advertising and data collection. To the Canadian Institute for Health Research, I appreciate the financial support provided via the Health Professional Student Research Award. Lastly, I wish to express appreciation to Dr. Janet Bavelas for the use of her laboratory and her advice on the project.

To my family, I would like to thank-you for your support and your patience with the long hours and absences that it required to get the project completed and for your encouragement along the way. I also wish to thank Robin Beninger for his prowess in computer programming and graphic design that allowed my vision of the network task and speech monitoring model to come to fruition. For all her assistance with editing, I am very grateful to Antoinette Beninger.

To my colleague and friend, Dr. Jacqueline Bush, I thank-you for all the hours of discussion of concepts and coping. Further, your assistance with the reliability coding was essential to the completion of this project. As a true paediatric clinician, I have the utmost respect for your clinical and research skills. You were a pillar throughout the program and for this, I am eternally grateful.

To Dr. Sworkdal and Dr. Mason, your assistance with this journey has been imperative. Without you, I could not have completed this project and I am forever indebted to both of you.

I thank all of the mentors and professors who shaped my research skills so that I was prepared for this journey. I also thank all of my students who pushed me to become better with each of their questions and who will always be a reminder that the process of learning is never completed.
Impact of Working Memory Constraints on Speech Monitoring in Healthy Children

The purpose of speaking is to provide information to listeners. To ensure that the message speakers desire to impart is understood by the listener, speakers must continuously inspect the verbal output and make any necessary corrections when errors occur, a process referred to as speech monitoring (Hartsuiker & Kolk, 2001; Levelt, 1989; Oomen & Postma, 2002). Monitoring and adjusting of contextual appropriateness, and of semantic, syntactic, phonologic and prosodic aspects of speech are necessary to ensure that the listener receives the proper message (Dell & Kim, 2005; Hartsuiker & Kolk, 2001; Levelt, 1989).

All of these potential foci for speech monitoring are likely not attended to simultaneously while a speaker is speaking, a task that would require vast attentional resources (Levelt, 1989). In fact, even when the goal of the task is accuracy, speakers may make numerous errors and repair only some of those errors (Levelt, 1983). In a seminal study, Willem Levelt analyzed 2,809 verbal samples for error rates and speech repairs (1983). In the study, healthy adults observed and then described visual pathways through a network of coloured circles with the goal of providing a description that another person could draw the pathway without having seen the stimulus (Levelt, 1983). In 17% of these descriptions, errors in colour naming occurred and of those errors, speakers only corrected 46%; other studies have substantiated this level of correction (Levelt, 1983; Nooteboom, 1980; 2005). Thus, even when the emphasis of a task is accuracy, adult speakers continue to make and not repair errors.
Levelt (1983; 1989) suggested that errors made and not repaired are failures of detection, rather than failures of correction. He theorized (1989) that the focus of attentional resources is on the aspects of speech that would be potentially most detrimental to the understanding of the listener. For example, in a social context, one may be less concerned with mild naming errors (e.g., saying truck instead of van) than in a professional context (e.g., saying agraphia versus aphasia), especially when the errors do not impede the overall meaning of the message (Motley, 1980). However, if conditions reduce the ability to monitor the most pertinent aspects of speech, then the potential for numerous speech errors increases. These errors affect listeners’ comprehension and/or the listeners’ perception of the speaker. Speakers who make frequent speech errors are judged more negatively by listeners with respect to aspects of personality, social status, beliefs, and competence (Engstrom, 1994; Kreuz & Roberts, 1993; Small & Burroughs, 1995; Yang, 2002). As negative evaluations have an impact on an individual’s academic, occupational, and social well-being, it is important to determine the factors that impact speech monitoring and potentially, to develop remediation approaches to address these factors.

Cognitive factors affecting speech monitoring are attention and working memory. Attentional processes influence information processing at the initial sensory stages and during the “postperceptual stages of processing” (Awh, Vogel & Oh, 2006, p. 201). Working memory is the ability to temporarily maintain information while using it to complete a mental function (e.g., calculating a restaurant tip without the assistance of paper or a calculator) (Baddeley, 1992). Attention assists in focusing our cognitive resources towards goal-relevant information and biases the information that gains access
to working memory (Awh et al., 2006). When attentional resources are limited, working memory capacity reduces; however, this relationship is likely to be bidirectional as when cognitive load prevents the attention resources necessary to activate long-term memory representations and monitor output (i.e. working memory load) is increased, attentional processes are less efficient (Oh & Kim, 2004; Woodman & Luck, 2004).

Many children (and adults) suffer from attention and working memory deficits following trauma (e.g., brain injury due to cerebral malaria) (Boivin et al., 2007), developmental constraints (e.g., preterm birth) (Vicari, Caravale, Giovanni, Casadei, & Allemand, 2004), and/or environmental constraints (e.g., pre-natal exposure to alcohol) (Rasmussen, Soleimani & Pei, 2011). Even in individuals without neurological constraints, tasks requiring multiple mental manipulations while retaining information can strain attention and working memory resources. If attention and working memory is important for speech monitoring, then a clear understanding of the relationship could provide key information necessary for the development of remediation techniques.

**A Model of Speech Monitoring**

One of the most well-studied and well-supported theories of speech monitoring is the perceptual loop theory (PLT) proposed by Willem Levelt and colleagues (Levelt, 1983; 1989; Levelt, Roelofs & Meyer, 1999). The PLT is an editor theory of monitoring, which proposes that an editor system external to the primary speech production system monitors the production output (Levelt, 1989; Levelt et al., 1999). The PLT proposes a “double perceptual loop” which allows speakers to covertly review their speech (inner loop monitoring) before articulation, as well as monitoring their overt speech (post-articulatory or auditory loop monitoring) for errors (Levelt, 1989, p. 469). The editor(s)
parse out the information and compare the comprehended message to the intended message (Levelt, 1989). Levelt’s theory provides a solid foundation for the study of speech monitoring, and was the primary model conceptualized in the current study; however, modifications were applied based on the current speech monitoring research evidence.

**Conceptualization.** Preceding the initiation of speech, the speaker must first plan the utterance with respect to meaning and purpose, a task completed by the *conceptual loop* (Levelt, 1983; 1989). This is a metacognitive process of evaluating one’s goals in a particular situation with a specific listener (Postma, 2000). This process requires taking on the perspective of one’s interlocutor to determine the most pertinent information needed (Levelt, 1999); for example, knowing to say “Susan” instead of “my wife” when the person to whom you are speaking is familiar with your relationship to Susan. The conceptual loop or conceptualizer retains the plan in working memory for later comparison to the formed message (Levelt, 1983). Errors produced at this stage would reflect errors in context appropriateness monitoring (Postma, 2000) such as using slang when speaking in a formal meeting.

**First stage of formulation.** The formulator receives the plan from the conceptualizer and translates the preverbal message into the necessary instructions for articulation (Levelt, 1983; 1989). This process consists of two main stages. The first stage is the activation of “multiple conceptually similar lexical-semantic representations” (Belke, 2008, p. 357) and selection of the appropriate lemmas which correspond to the elements within the planned message. Lemmas consist of a superordinate form of a word that represents semantically analogous words with different forms for specific contexts
(e.g., “run” for running, ran, etc.) (Levelt, 1983). In order to distinguish the correct lemma from the overall category and related information, a person must have the ability to define the distinctive features of the particular lemma (e.g. what makes a dog, a dog and not a cat) (DeLeon et al., 2007). The process of lemma selection is constrained by attentional resources in dual task paradigms (Ferreira & Pashler, 2002).

During the first stage of formulation, semantic errors may occur. Semantic errors that commonly occur are coordinate semantic errors (i.e., erroneous word belongs to the same category as target) (e.g., naming a picture of a dog as a cat) or associative semantic errors (i.e., erroneous word is related but not of the same category as target) (e.g., naming a dog as a bone) (Cloutman et al., 2009). A dissociation between these two types of errors had been determined using picture naming and word-picture matching tasks in two adult patient groups, semantic dementia and post-stroke comprehension aphasia. Individuals with semantic dementia (commonly show atrophy of anterior temporal lobe) showed increased sensitivity to familiarity and made coordinate errors whereas individuals with post-stroke comprehension aphasia were insensitive to familiarity and made associative errors (Jefferies & Lambon Ralph, 2006). Jefferies and Lambon Ralph (2006) suggest that the performance of those with post-stroke aphasia indicates a loss of semantic control and not a loss of connection to semantic knowledge as proposed in semantic dementia.

**Semantic control.** Semantic control allows a person to determine what information is pertinent to the particular situation (e.g., when naming a dog, irrelevant information about dog food, leashes, the park, etc. can be ignored) (Jefferies & Lambon Ralph, 2006). Semantic control was differentiated from general working memory
functions in a patient with a history of resolved post-stroke transcortical sensory aphasia (Hoffman, Jefferies, Haffey, Littlejohns & Lambon Ralph, 2013). The patient, JB, performed adequately on non-semantic tasks of working memory (e.g., digit span) but demonstrated impaired performance on semantic control tasks (e.g., category fluency) (Hoffman et al., 2013). While the authors postulated that this means semantic control is separate from general working memory, it is possible that JB had impairment in the working memory’s connection to long-term memory for semantic information. This would result in the inability to determine the most appropriate semantic category for the task.

The left inferior frontal gyrus (LIFG) is one of the areas postulated to be key in the selection of the correct lemmas from all of the possibilities held in working memory (Moss et al., 2005; Thompson-Schill, D’Esposito, Aguirre & Farah, 1997). When generating verbs associated with pictures, the LIFG activity is greater than at baseline (Moss et al., 2005; Thompson-Schill et al., 1997). The activity of LIFG was much higher when a picture was associated with a greater number of possible verbs (i.e., a greater selection demand) (Thompson-Schill et al., 1997). Thus, the LIFG may only be required in situations requiring greater semantic control. The cognitive control role of the LIFG with respect to language was reviewed by Novick, Trueswell and Thompson-Schill (2005). Specifically, Novick et al. (2005) suggested that the LIFG is essential for the allocation of attentional resources to task-relevant information, which is not only specific to language. This is one theoretical basis of working memory. However, it must be noted that using verbal activities to measure brain regions for verbal working memory
may create challenges in determining if one is simply measuring a basic language region or a region more specific to working memory.

**Attention, working memory and verbal fluency.** Daneman (1991) used working memory performance (speaking span test) to predict verbal fluency performance (speech generation, oral reading and oral slip tasks). Speaking span predicted the number of words produced and speech rate in the speech generation task and reading rate in the oral reading task (Daneman, 1991). Individuals with lower speaking spans produced more spoonerisms (error in speech in which corresponding consonants, vowels, or morphemes are switched) in the oral slip task compared to individuals with higher speaking spans (Daneman, 1991). Rosen and Engle (1997) found that undergraduate students with high working memory (based on an operation span task) were able to provide more animal exemplars compared to students with low working memory. Students with low working memory showed a pattern of higher number of repetitions, despite instructions to avoid repetitions (Rosen & Engle, 1997). Further, Rosen and Engle (1997) found that a concurrent task reduced the retrieval of students with high working memory, whereas students with low working memory did not show this pattern. The authors speculated that only students with high working memory have the capacity to monitor for errors (e.g., repetitions) and retrieve words under normal conditions (Rosen & Engle, 1997). With a concurrent task, which requires cognitive resources, the attempt to continue to monitor and retrieve reduces the retrieval efficiency. This is again consistent with the theory that the role of working memory in language is cognitive control of attention focus towards task-relevant information, in this case, retrieval of word exemplars (Novick et al., 2005). As the students with low working memory were likely not engaging in
monitoring to the same degree, the attentional load had little to no impact on retrieval (Rosen & Engle, 1997). However, one caveat to these conclusions is that these studies used language based working memory tasks to find the relationship between working memory and language performance. As a result, this may create spurious relationships between verbal working memory and language. A discussion of this issue occurs later in this document.

The impact of a cognitive load, which prevents the attention resources necessary to activate long-term memory representations and monitor output (i.e. attention and working memory load), on fluency is also notable. Healthy young adults completed tasks of semantic and letter fluency with a simultaneous memory task (Azuma, 2004). A manipulation of the relationship of memory load words to fluency task words was created by providing either semantically related words or words starting with the same letter (Azuma, 2004). On the task of semantic fluency, use of the semantically related words for the memory load doubled the rate of perseverations compared to not-semantically related words (Azuma, 2004). Azuma (2004) postulated that it is not the total amount of load on working memory but in fact, the type of information in the load. On the task of letter fluency, three times as many perseverations occurred in the same-letter-memory load condition as the different-letter-memory load condition (Azuma, 2004). Letter-fluency is less automatic than semantic fluency and requires greater attention resources in order to activate the exemplars in long term memory as well as the ability to monitor and suppress previously retrieved information (Azuma, 2004). This suggests that the greater need to monitor the production of words distinct from those on the memory list created
by related-word-memory lists is an attention demanding task, which then reduces the
ability to monitor for repetitions.

**Developmental trends in semantic control.** A developmental trend of increasing
word production during semantic fluency tests has been found when comparing children
aged 7-8, aged 9-10, aged 11-12 and aged 13-14 (Riva, Nichelli & Devoti, 2000;
Sauzéon, Lestage, Raboutet, N’Kaoua & Claverie, 2004). When completing a free
fluency task (e.g., naming as many words as possible), children aged 11-12 produced
more categories than younger children; however, at age 11-12, children produced both
category names and exemplars, which, Sauzéon et al. (2004) suggested, reflects an
ineffective exploration of semantic categories. In children aged 13-14, a shift occurs
towards greater production of exemplars (Sauzéon et al., 2004). While functional
developmental differences in fluency are evident, by the age of seven, children show the
same pattern of activation shown on fMRI in the left inferior frontal gyrus and left middle
frontal gyrus when completing a category fluency task as that seen in adults (Gaillard et
al., 2003). This suggests that while the same regions of the brain are involved in the task,
the process by which those brain regions process the task is less efficient in children.

**Developmental trends in the relationship between attention, working memory
and semantic control.** In children aged 6 to 13, Brocki and Bohlin (2004) found that
measures of working memory and verbal fluency converged onto a single factor when
analyzed via factor analysis. Brocki and Bohlin (2004) postulated that verbal fluency
tasks, by nature, require the maintenance of information in mind (i.e., working memory).
Further, developmental trends for the working memory/fluency factor show significant
improvements occur at age 8 and age 12. The increase at age 8 may reflect the strategy
switch from coding information visually to a phonological approach, which improves recall (Brocki & Bohlin, 2004). Whereas, the increase at age 12 may reflect the greater proficiency with which children of this age group can access information using phonological information (e.g., first letter of word) (Brocki & Bohlin, 2004). The greater automaticity of information access reduces the load on attention, which would then reduce constraints on working memory processes.

Stage two of formulation. The second stage of the formulator’s creation of a speech plan is the activation of syntactic building procedures, which correspond to the syntactic blueprint included in the lemmas (Levelt, 1983; 1989). The syntactic or grammatical encoding ensures the use of the proper form of the word necessary to express the message determined by the conceptual loop (Levelt, 1983; 1989; Postma, 2000).

The frontal operculum and Broca’s area (Brodmann’s areas 44, 45 and 47) have well-established roles in syntactic structure creation and syntactic violation monitoring (Friederici, 2002; Friederici, Fiebach, Schlesewsky, Bornkessel & von Cramon, 2006). Syntactically erroneous sentences elicit an anterior negativity between 150 and 400 ms with a later positivity between 300 and 900ms (for review, see Friederici, Hahne & Saddy, 2002). The early anterior negativity may reflect the identification of an error and the later positivity may reflect the reformulation required to make a repair to the syntax (Friederici et al., 2002). Makuuchi, Bahlmann, Anwander and Friederici (2009) found that the left pars opercularis showed greater activation to syntactically complex structure of auditory sentences and the left inferior frontal gyrus showed greater activation to sentences requiring the transfer of specific syntactic information over a longer distance
within a sentence (e.g., greater number of words between the main subject and the verb). This suggests that the left inferior frontal gyrus may provide, in addition to the previously mentioned role in semantic control, the working memory support for the function of parsing syntax.

**Developmental trend of syntax acquisition.** An active period of acquiring base syntax occurs between age 18 months and four years (Brown, 1973). Complex syntax forms tend to arise in spoken language as early as age 2 or 3 years of age (Bloom, Tackeff, & Lahey, 1984; Diessel, 2004). As the child matures, understanding and use of syntactically complex sentence structures increases. For example, in understanding agent (performs the action)-patient (immediately affected by event) relationships, children of different ages depend on different information. In sentences where the animacy (aliveness) contrasts and word order coalesce, 2 year olds are able to determine that the first animate noun is the agent (e.g., “The dog wams the hat” where dog is the agent and hat is the patient) (Chan, Lieven & Tomasello, 2009). However, when the semantic and syntactic cues conflict, 2-year-old children are unable to use either cue systematically to determine the agent (Chan et al., 2009). In children age 3-4, especially English speaking children, word order was the preferred over animacy (e.g., “The book geens the goat” where children choose the book rather than the goat as the agent) (Chan et al., 2009). This is consistent with the English language format, which has a dominant word order cue (Chan et al., 2009). While pre-school children make tremendous gains in terms of syntactic use, gains continue to occur throughout middle to late childhood and adolescence (Berman, 2004).
As children mature, their ability to create and parse syntactic structures consistent with their dominant language increases. A slow progression occurs from age 8 into adulthood in terms of the number of T-units (a main clause attached to any subordinate clause or nonclausal structure) and in terms of the subordination index (total number of clauses / number of T-units) are present in expository and narrative discourse (Berman & Lerhoeven, 2002; Nippold, Hesketh, Duthie, & Mansfield, 2005; Scott, 1988; Scott & Windsor, 2000). This progression is key in terms of children learning to express information in the most efficient and efficacious manner as syntactic complexity allows a person to combine several small sentences into a single, descriptive sentence (Nippold et al., 2005).

**Errors at level of formulation.** In the first stage, activation of “multiple conceptually similar lexical-semantic representations” (Belke, 2008, p. 357) can lead to *semantic context effects* (Belke, Brysbaert, Meyer & Ghyselinck, 2005). Semantic context effects occur when a person views superordinately related objects in succession and interference occurs, which increases the latency of object naming (Belke, 2013). Picture naming latency is longer when a person successively names semantically related objects than when naming pictures of unrelated objects; however, there is not a similar effect when word naming (Belke, 2013; Damian, Vigliocco & Levelt, 2001). Damian et al. (2001) suggest lemma access is necessary for picture naming but not for word naming. This may be due to the necessity of accessing the superordinate category for the object in picture naming to identify it correctly, which may demand more attention resources. When word naming, it is possible that only phonological and syntactic information is necessary to correctly produce the word, a skill that is more automatic and thus, less
attention demanding. Some have hypothesize that this explains why when attention and working memory load is increased, both picture naming and word naming latencies increase due to a lack of attention resources for phonological and syntactic retrieval (Belke, 2008).

The syntactic or grammatical encoding ensures the use of the proper form of the word necessary to express the message determined by the conceptual loop (Levelt, 1983; 1989; Postma, 2000). Lexicality and syntax monitors may signal the need for corrections of word choice or word form at this stage (Levelt, 1989; 1999; Postma, 2000). Errors at this stage likely reflect mistakes in the choice of sub-lexical elements due to the superordinate nature of lemmas (Levelt, 1983; Postma, 2000). This process creates lexical bias, “the fact that phonological speech errors tend to make more real words than non-words” (Nooteboom, 2005, p. 44). One explanation for the lexical bias is that the lexicality and syntax monitors are more likely to detect non-word errors (e.g., moog instead of moose) than real-word errors (e.g., goose instead of moose) (Nooteboom, 2005; Nozari & Dell, 2009). This occurs as real words are only inappropriate due to the context of the present moment whereas non-words are rarely or never appropriate. However, the lexical editor is not very accurate, even when with an emphasis on accuracy (Nozari & Dell, 2009).

In childhood, the use of the lexical editor may not be as efficient and flexible as in adults. In terms of metalinguistic ability, children’s ability to detect and correct phonological errors as well as semantic errors in a consistent manner develops around age 7 (Clark, 1978; Smith & Tager-Flusberg, 1982). Children find it more difficult to identify and revise syntactic errors in spoken and written sentences compared to semantic
or phonological errors (Kamhi & Koenig, 1985; Cairns, Schlisselberg, Waltzman & McDaniel, 2006). For syntactic errors, children age 4 to 7 correct approximately 25% (Kamhi & Koenig, 1985). Around age 8, the metalinguistic ability to quickly judge and formulate repairs to syntactic errors increases substantially (Edwards & Kirkpatrick, 1999).

The articulatory buffer. The formulator uses these two stages, in combination with the morphological and phonological structures, to create a structure of a pronounceable phonetic plan (Levelt, 1983; 1989). The formulator sends the phonetic plan to both the speech comprehension system for monitoring and to the articulator via the articulatory buffer (a form of working memory) for retention until the time that the speaker is able to articulate the information (Levelt, 1983, 1989). Levelt (1989) proposed that the articulator is able to articulate the plan approximately 200-250 ms following creation of the phonetic plan. A buffer-articulation timing monitor may track the timing of new to-be-articulated material (Blackmer & Mitton, 1993; Postma & Kolk, 1993). If a person speaks too quickly or a person’s information processing resources are strained and no new to-be-articulated information is yet available, this monitor may enlist a restart program, leading to repetition of already articulated material (Blackmer & Mitton, 1993; Postma & Kolk, 1993).

Inner Loop Monitoring

Levelt (1983) proposed that the speech comprehension system completes pre-articulatory monitoring, in addition to post-articulatory monitoring. Detection of errors in one’s own speech is very similar to detection of errors in others’ speech when the intended meaning is clear; however, we tend to be more accurate in detecting the errors in
others’ speech (Oomen & Postma, 2002; Oomen, Postma & Kolk, 2001). The latter finding suggests that detecting errors in our own speech may be more taxing on our resources than parsing others’ speech (Oomen et al., 2001). However, it may be that the resources necessary for detection of errors in our speech differs from that of other’s speech. Marshall, Rappaport, and Garcia-Bunuel (1985) reported a case study of a patient with auditory agnosia who was unable to understand other’s speech; however, she demonstrated an ability to detect many of her own speech errors. This may be due to preserved internal monitoring abilities or this is an indication of separate processes of self versus other’s speech monitoring (Marshall et al., 1985).

Interruption of erroneous words occurs quite early in speech production, approximately 150 ms following generation of the phonetic plan (Levelt, 1989). Approximately 50 ms after an erroneous word is heard, an error-related negativity forms on ERP with peaks at 100 ms for frontal electrodes and 150 ms at posterior electrodes (Schiller, Horemans, Ganushchak & Koester, 2009). The speed at which interruptions of the erroneous words occur suggests that it could not be due to overt error detection (Blackmer & Mitton, 1991; Hartsuiker & Kolk, 2001; Levelt, 1983, 1989). If accurate speech monitoring requires overt error detection, error detection time would need to include time to articulate the word, perceive the error, register the error, and make the interruption (Levelt, 1983). In order to explain the quick rate of error detection, Levelt (1983; 1989) proposed an inner monitoring loop, accessible to attention, which monitors the phonetic plan. Later revisions of the theory proposed that the inner loop monitor has access to both the phonetic plan and a more abstract, phonemic and metrical (e.g., pitch, duration, loudness or combinations of these features) representation (Drescher, 2004;
Levlt et al., 1999; Wheeldon & Levlt, 1995). The internal monitoring loop detects and interrupts errors before they are articulated, a form of covert monitoring (Hartsuiker & Kolk, 2001). As previously noted, for each step of language production (as proposed by Levlt, 1983), specific monitors exist to detect specific types of errors before articulation.

White noise presentation during speech production prevents the overt monitoring system from detecting articulated errors (Hartsuiker, Bastinaase, Postma & Wijnen, 2003; Postma & Kolk, 1992). The person cannot accurately hear his/her external speech, or external speech is degraded so as not to provide sufficient auditory feedback, and thus, errors detected reflect those identified via the inner monitoring loop (Civier, Tasko & Guenther, 2010; Hartsuiker et al., 2003; Postma & Kolk, 1992). Studies employing white noise and dual task conditions have found that healthy speakers detect errors in speech at a decreased rate and make less repairs, which may reflect that the speakers no longer have access to the information via auditory feedback (Lackner & Tuller, 1979; Oomen, Postma & Kolk, 2001; Postma & Kolk, 1992; Postma & Noordanus, 1996). In a large proportion of individuals who stutter, the use of masking noise significantly reduces the frequency of repetition, especially sound/syllable repetitions (Civier et al., 2010; for review, see Lincoln, Packman & Onslow, 2006). If repetitions indicate an attempt to repair as suggested by Levlt’s (1983) model, this would support the findings in healthy individuals. While an extensive review of all theories of stuttering is beyond this paper, a number of views alternative to Levlt’s (1983) model do exist. One alternative view of repetitions in stuttering was theorized by Max, Guenther, Gracco, Ghosh and Wallace (2004). These authors suggest that early in development, “children have a high threshold for sensory error-based motor resets” to prevent constant resets while learning new
sounds (Max et al., 2004, p. 115). As children mature, they begin to use a feed-forward mechanism, which encodes the expected sensory consequences of the sound created and as the speaker becomes skilled, few sensory errors occur (Max et al., 2004). In children who stutter, the feed-forward mechanism may not develop or may insufficiently develop. Children who stutter continue to use a feedback mechanism based on parallel generation of motor commands with the error signal produced via a sensory information comparison of actual and target position of the motor system as such, sensory errors continue and trigger motor resets (Max et al., 2004). In children who stutter, auditory masking reduces the efficacy of the feedback circuit due to the continual auditory feedback that is inconsistent with the speech motor movements (Max et al., 2004). Alm (2005) proposes another view and suggests that white noise or other altered auditory feedback changes the feedback to the basal ganglia resulting in a de-automatization of the motor programs for speech, which are impaired in individuals who stutter. While there are a number of theories, it is important to note that masking auditory feedback does not increase fluency in all individuals who stutter, so future research will need determine if this is a valid interpretation of the evidence (Lincoln et al., 2006).

The current study employed white noise to determine developmental trends in errors and repairs when the auditory monitoring loop is masked. A major critique of the use of white noise to isolate the inner monitoring loop is that this approach fails to account for tactile and proprioceptive feedback that occurs when creating a sound. As a person learns and becomes fluent in a language, a “somatosensory target region” or somatosensory criterion develops for each sound (Guenther, 2006, p. 353). The somatosensory system uses this criterion to determine if our current articulation meets
expectation (Guenter, 2006). While the measurement of tactile and proprioceptive
feedback is beyond the scope of the current paper, when making conclusions about the
inner monitoring loop via the use of white noise masking, it requires consideration in
conclusions about the role of white noise.

If, before articulation, an error is determined based on the linguistic rules of the
language spoken, the speaker may require time to formulate the repair, causing a pause in
speech or may repeat sections of a word as an attempt to restart (Hartsuiker & Kolk,
2001; Postma & Kolk, 1992). Between the ages of 5 and 17, the duration of pauses in
speech reduces by as much as 50%, potentially reflecting an improvement in speech
planning efficiency (Nip & Green, 2013; Pavão Martins, Vieira, Loureiro & Santos,
2007; Singh, Shantisudha & Chatterjee Singh, 2007). This is further supported by the
developmental trend that 5-year-old children, when completing narrative tasks, tend to
repeat sentences and reformulate frequently whereas 17-year-old youth tend to focus on
salient details of the story without the need for frequent repetition or reformulation
(Pavão Martins et al., 2007). Repetitions in spontaneous speech show a similar
developmental trend with higher rates in younger children (Bjerkan, 1980; DeJoy &
Gregory, 1985).

Levelt (1983) proposed that interruptions would occur immediately after the
detection of an error (error to cutoff time) and the latency reflected the time to detect the
error. However, other authors have found a delay in the interruptions of speech after
error detection with latency dependence on repair characteristics (e.g., removal of
erroneous word or addition of missing word) (Berg, 1986; Blackmer & Mitton, 1991;
postulated that the time between interruption of speech and initiation of a repair (cutoff to repair time) reflected the time needed to create the repair. However, Blackmer and Mitton (1991) reported cutoff to repair times that were zero. Again, it is likely that cutoff to repair times also reflect the nature of the repair and contextual constraints on speech production (e.g., time pressure) (Oomen & Postma, 2001).

Repairs of a covert error are simply the “correction of errors without external prompting frequently within a short period of time from the moment of error occurrence” (Postma, 2000, p. 98). For example, if the error was a word with similar initial phonemes as the correct word, the person may say the initial phonemes and then stop mid-word to make the correction (e.g., “the car was bl, black” - where the first colour to be said was blue). A further signal of a covert error is via the use of editing expressions. Levelt (1983; 1989) refers to the use of editing expressions (e.g., er, that is, um, sorry, I mean) as a way by which the speaker lets the listener know that something is problematic in the message and that a correction will occur. At the present time, researchers are unable to measure these moments of covert repairs directly and as such, typically rely on these pauses, repetitions, or use of editing expressions as ways to measure these unspoken occurrences (Levelt, 1983; 1989; Postma 2000).

**Articulation.** A motor plan is created that corresponds to the oral structure movements necessary to create the sounds in the phonological plan. In order to produce speech, more than 100 muscle movements need to coordinated (Simonyan & Horwitz, 2011). During childhood and into late adolescence, there is greater variability in the synergy of functional oral muscle groups than in adulthood (Kleinow & Smith, 2006; Smith, Goffman, Zelaznik, Ying, & McGillem, 1995; Smith & Zelaznik, 2004; Walsh &
Smith, 2002). This lack of consistency may make it more difficult for children prior to mid- to late adolescence to detect problems with articulation. Further, during speech, somatosensory feedback is provided from the tongue, larynx, jaw, and other oral structures that guide alterations in speech motor movements, in addition to auditory feedback (Lametti, Nasir, & Ostry, 2012; Tremblay, Shiller & Ostry, 2003). If one type of feedback, auditory or somatosensory, is altered by use of white noise or other experimental tools, a greater reliance on the other type of feedback may occur (Lametti, Nasir, & Ostry, 2012). Further, with more complex utterances and greater processing demands, greater variability in the tactile and haptic feedback during speech occurs until mid-adolescence (Sadagopan & Smith, 2008; Smith, 2006). When white noise is presented and the auditory feedback is disrupted during speech production, children prior to mid-adolescence may make more errors due to the variable tactile and proprioceptive feedback.

**Auditory Loop Monitoring**

The PLT proposes that an auditory loop monitors articulated (overt) errors (Levelt, 1983; 1989). The speech comprehension system parses and compares auditory information to the intended message in the conceptualizer (Levelt, 1983; 1989). Auditory recognition of words occurs around 200 ms after word onset (Marslen-Wilson & Tyler, 1980; 1981). Therefore, this is the fastest that the auditory loop could potentially identify that an error has occurred and initiate the process of error correction. For researchers, it is easier to identify articulated errors than covert errors. However, as with covert repairs, a speaker may also pause, repeat, or make use of editing expressions in overt repairs as an attempt to clarify meaning (Levelt, 1983; 1989).
The Role of Attention and Working Memory in Speech Monitoring

In the PLT, speech monitoring is capacity limited (Oomen & Postma, 2002), which may reflect limited resources of attention and working memory. While Levelt (1983, 1989) includes working memory in the PLT, the nature of that role is vague within the model. One purpose of the current study is to integrate the current literature on attention and working memory with the literature on speech monitoring in order to provide an updated and more detailed version of Levelt’s model.

Attentional control has a primary importance in the maintenance of information in the context of potential distractors, both internal and external, and is a key determiner of one’s ability to monitor one’s speech (Buchsbaum & D’Esposito, 2008; Engle & Kane, 2004; Kane, Conway, Hambrick & Engle, 2007; Unsworth & Engle, 2007; Waters & Caplan, 1996). Attention “selectively update[s] relevant information brought into the focus of attention while ignoring irrelevant information that is outside the focus of attention,” (Magimairaj & Montgomery, 2013, p. 2). This attentional process prolongs activation of information so that it is readily accessible for use within working memory (Engle & Kane, 2004). As noted previously, this process creates a bias in terms of what information is available for working memory (Awh et al., 2006).

Demands of the environment such as secondary tasks, priority of tasks, and potential rewards for those tasks likely influence the allocation of attention (Buchsbaum & D’Esposito, 2008). The time-based resource-sharing (TBRS) model of working memory proposed by Barrouillet and colleagues posits that cognitive load impairs working memory as the brain’s processing activities capture attention in such a manner that reduces or prevents the refreshing of memory traces (Barrouillet, Bernardin &
Camos, 2004). Attentional refreshing reactivates both the semantic and phonological information of the intended speech plan (Camos, Mora & Oberauer, 2011). Cognitive load increases with the number of times memory traces need to be refreshed, the speed required to process the task, and the ratio between the two (Barrouillet et al., 2004; Camos, Lagner & Barrouillet, 2009). The working memory spans of children younger than age 7 are not impacted by cognitive load, which suggests that attentional refreshing is not a consistently used strategy for working memory maintenance in this age range (Barrouillet, Gavens, Vergauwe, Gaillard & Camos, 2009). From the age of 7, the ability to refresh information in working memory via attention continues to improve into adolescence, which corresponds to the developmental trend of working memory span (Barrouillet et al., 2009). However, even in older children and adults, while tasks with a higher attention-demanding nature impede attentional refreshing, working memory is not entirely impeded. These findings suggest attentional refreshing is not the only strategy to maintain information in working memory (Camos et al., 2011).

Articulatory rehearsal, the sub-vocalization of language that serves to maintain verbal information, is a second necessary component of verbal working memory specifically (Camos et al., 2009; Magimairaj & Montgomery, 2012). When a task demands attention to a degree that attentional refreshing is not available, a strategic switch to a less attention-demanding strategy, articulatory rehearsal, occurs (Camos et al., 2011). However, this switch comes with a cost. Articulatory rehearsal merely allows the reactivation of the specific phonological information, which can lead to phonological errors when the information contains phonologically similar content (Camos et al., 2011). Further, as with attentional refreshing, in young children (< 7 years), articulatory
rehearsal does not occur reliably, which may be another reason, children in this age range have lower working memory spans than older children (Gathercole, 1998).

Further, a certain level of cognitive flexibility is required to switch attention between task processing and storage, a skill believed to increase with the development of the frontal lobes after age seven (Camos & Barrouillet, 2011). In 6-year-old children, verbal recall performance depends solely on the time delay between “encoding and recall without any effect of the cognitive load of the intervening activity,” which may suggest an inability to utilize the attentional refreshing switching strategy seen in older children (Camos & Barrouillet, 2011, p. 903). This provides further support for the progressive developmental trajectory of working memory.

One model that has examined the name of articulatory rehearsal in greater depth and is one of the most widely used working memory models in the child development literature is that of Baddeley and Hitch (1974). This multicomponent model of working memory proposed that working memory consisted of the central executive, and two subsidiary components, the phonological loop and the visuospatial sketchpad. A later revision of the model also included a third subsidiary component, the episodic buffer (Repovš & Baddeley, 2006). In this model, the general allocation of resources is separate from memory maintenance (Baddeley & Hitch, 1974). The central executive component “is responsible for the manipulation of information within working memory” via the control of the three subsidiary systems and the allocation of attentional resources (Baddeley, 1992; Repovš & Baddeley, 2006, p. 6). The central executive aspect is contrary to the formerly mentioned models as it separates temporary storage from attention refreshing, which some theorists argue limits its application (Camos et al.,
However, as this model continues to be one of the most widely used models in the child development literature and specifically in the language development literature, it is worth expanding upon.

The first subsidiary system, the phonological loop, consists of a temporary store of acoustic and phonological forms as well as an articulatory control or rehearsal system (Baddeley, 1992; Repovš & Baddeley, 2006). The maintenance of phonological and acoustic information in the phonological store is brief (1 – 2 seconds) and so the articulatory rehearsal system acts to refresh the information via subvocal repetition, which allows the system to maintain information over a longer period (Baddeley, 1992; Buchsbaum & D’Esposito, 2008). Baddeley (1992) further postulated that the phonological loop enables the temporary storage of visual images (e.g., pictures of objects or name of objects) “in the phonological store by subvocalization” (p. 558).

These characteristics of the phonological loop have a clear link to Levelt’s (1983, 1989) theory of a necessary storage and activation of verbal information for the comparison of the intended and the actual verbal message when self-monitoring for errors.

Other authors have proposed that maintenance of information in verbal working memory is due to activation of long-term storage within the language production system (Acheson, Hamidi, Binder & Postle, 2011; Buchsbaum & D’Esposito, 2008). Studies demonstrating an influence of the language production system on short-term recall provide support for this theory. Words are easier to recall than non-words (Hulme, Maughan & Brown, 1991). Concrete words are easier to recall than abstract words (Walker & Hulme, 1999). Words of higher frequency are easier to recall than words of low frequency (Roodenrys, Hulme, Lethbridge, Hinton & Nimmo, 2002). Non-words
with higher frequency phonemes are easier to recall than words with lower frequency phonemes (Gathercole, Frankish, Pickering & Peaker, 1999). Presentations of words in grammatically correct sentence structures leads to more words recalled (Gilchrist, Cowan & Naveh-Benjamin, 2009). More word pairs following English syntax form (e.g., adjective-noun) are recalled by fluent English speakers than pairs following an alternative syntax form (e.g., noun-adjective) (Perham, Marsh & Jones, 2009). In consonant pair stimuli, a lower frequency first consonant (with respect to second consonant) elicited more errors compared to those pairs with a high frequency first consonant (Levitt & Healy, 1985). This body of research indicates that short-term recall is lower for verbal information that follows a format that has a zero or low frequency in an individual’s language production system. While the findings of these studies do not necessarily negate the role of the phonological loop, the process by which the long-term memory influences short-term recall of linguistic information may be better accounted for by the episodic buffer as discussed later in this paper.

Load on phonological working memory may lead to the production of speech errors. In 1980, Albert Ellis proposed that potentially a single “phonemic response buffer” was involved in both verbal short-term memory and speech production as similar types of errors occur in both (p. 625). Ellis (1980) presented lists of syllables to participants and had them repeat the sequence in correct order. The lists of syllables varied with respect to having all the same vowels or different vowels across the lists, repeated or not repeated vowels within the lists, and same or different consonant-vowel combinations across lists (Ellis, 1980). During immediate recall, participants transposed the following: (1) consonants more frequently than vowels; (2) vowels more frequently
than whole syllables; (3) phonetically similar consonants more frequently than phonetically distinct consonants ("feature similarity effect"); (4) consonants sharing the same vowel more frequently than those with different vowels ("contextual similarity effect"); and (5) phonemes in similar syllable positions more frequently than those in different positions ("syllable position effect") (Ellis, 1980, p. 633). Subsequent studies using verbal working memory tasks corroborate Ellis’ (1980) findings (Page, Madge, Cumming & Norris, 2007; Vousden, Brown & Harley, 2000). These findings suggest that tasks measuring phonological storage elicit the same types of errors found in language research.

Also consistent with the linguistic literature, Saito and Baddeley (2004) found that the presentation of irrelevant words, which interrupts short-term memory performance, interferes with speech production when the words are phonologically similar (Levelt et al., 1999). These authors suggest that for speech production, timing of the presentation of irrelevant words to a period of “phonological planning” would create maximum interference, whereas with the “phonological store,” timing is less relevant (Saito & Baddeley, 2004, p. 1335). The “phonological planning factor,” as measured by memory span, reading rate and a tongue twister task, negatively correlated with speech error rate (Saito & Baddeley, 2004, p. 1329). Difficulty establishing strong phonological representations while planning speech may lead to higher rates of speech errors (Saito & Baddeley, 2004). This phonological planning factor would require more than the phonological loop and likely involves the third subsidiary system, the episodic buffer, to be discussed later in this paper. The phonological planning factor corresponds closely to the formulator in Levelt’s model.
Evidence for Development of the Phonological Store. As expected for a system integrated into the functions of the language system, the phonological store has an early developmental onset. In children, aged 3 years, strong phonological storage abilities predicted the production of a broader range of vocabulary, more grammatically complex sentences and overall longer speech samples when compared to children with weaker phonological storage abilities (Adams & Gathercole, 1995). A longitudinal study of children (ages 4, 5, 6 and 8) examining the relationship between phonological memory and vocabulary development found that between the ages of 4 and 5, phonological memory seems to be the factor predicting vocabulary development (Gathercole, Willis, Emslie & Baddeley, 1992). After the age of five, the nature of the relationship suggests that vocabulary performance predicts phonological memory development (Gathercole et al., 1992). Another developmental trend in the relationship between the phonological loop and vocabulary is the change in the relationship with respect to ordering and item information. Majerus, Poncelet, Greffe and Van der Linden (2006) found that serial order recall tasks predicted vocabulary development for children aged 4 and 6 years; however, they also found that for 5-year-old children, vocabulary development was better predicted by non-word delayed repetition. Majerus et al. (2006) additionally found that a significant increase in short-term verbal memory task performance occurred between ages 5 and 6, whereas no differences were present between ages 4 and 5. This suggests that an important shift in processing of phonological information occurs at age 5, which then leads to gains in short-term verbal memory ability and vocabulary knowledge (Engel de Abreu, Gathercole & Martin, 2011; Majerus et al., 2005). Around age 7, children begin to use the strategy of subvocal rehearsal, which allows for greater maintenance of
the information in the phonological store (Gathercole & Alloway, 2008). Speed of subvocal rehearsal, memory reactivation, and storage capacity increase as children’s cognitive abilities develop (Gaillard, Barrouillet, Jarrold & Camos, 2011; Gathercole & Alloway, 2008; Gilchrist et al., 2009). Healthy younger children, on average, have lower scores on measures of digit span than healthy older children (Gathercole et al., 2004; WISC-IV; Wechsler, 2003). For English speaking individuals, digit span performance progresses to adult levels around age of 15 (Gathercole & Alloway, 2008). These findings suggest that verbal working memory abilities increase over development until reaching adult levels in middle to late adolescence.

The second subsidiary system of the Baddeley and Hitch (1974) model, the visual-spatial sketchpad, consists of a temporary store and maintenance of visual and spatial forms (Repovš & Baddeley, 2006). While a full review of this area is outside the scope of this paper, for individuals who use American Sign Language, the visual-spatial sketchpad is proposed as the system that takes the place of the phonological loop for verbal language (Wilson & Emmorey, 1997). Signing errors have similarity effects (e.g., hand position, movement and location) (Wilson & Emmorey, 1997), manual articulatory suppression effects (Losiewicz, 2000) and irrelevant signed input effects (Wilson & Emmorey, 2003). These results suggest a similar relationship of working memory in American Sign Language as in oral English.

The visuospatial sketchpad was targeted for the filler task for the current study’s network language task via a visual-spatial span task in the desire to reduce language interference while maintaining working memory (Shah & Miyake, 1996). Visual-spatial tasks differentiate between children with healthy typical development and attention-
deficit disorder and show a developmental trend similar to that of the phonological loop (Westerberg, Hirvikoski, Forssberg & Klingberg, 2010; Gathercole et al., 2004). Children in the younger age group were expected to have lower scores than those in older age group in the current study.

The third subsidiary system is the episodic buffer, a limited-capacity, temporary store of integrated information from various systems including other temporary memory stores and long-term memory (Baddeley, 2000; Repovš & Baddeley, 2006). The episodic buffer integrates a wide variety of differently coded information (e.g., visual, auditory) into a multi-dimensional code representing complex, coherent summaries of information such as a scene or episode (Baddeley, 2000; Repovš & Baddeley, 2006). Using repetition of meaningful sentences as a measure of the episodic buffer, with children aged 4 to 6, this ability is related to but distinct from unrelated verbal item span (Alloway, Gathercole, Willis & Adams, 2004; Baddeley & Wilson, 2002). This finding illustrates a key characteristic of the episodic buffer, the binding of the semantic information from long-term memory with the working memory representation (Baddeley & Wilson, 2002; Rudner & Rönnberg, 2008).

Binding is the linking or combining of features into a whole concept (Oberauer & Lange, 2009). Binding of letters, phonemes and syllables is necessary to maintain appropriate placement of spoken language components and to avoid speech errors such as “cog, doat” for “dog, coat” on serial word recall (Oberauer & Lange, 2009). Further, the binding of more complex elements (e.g., a whole event) to contextual information (e.g., temporal reference), syntactic and semantic information is essential for more complex language forms (Oberauer & Lange, 2009). One example of a complex
language form, in which the episodic buffer’s ability to bind information together is necessary, is the ability to understand dependency relations. For example, when hearing the following sentence: “**Carol** laughed at **Crystal** who spilt juice on **herself**,” while holding the sentence in mind, a person must be able to bind the lexical-semantic and syntactic information together to understand that the word **herself** represents **Crystal** and not **Carol** (Santi & Grodzinsky, 2007).

While some researchers have argued for modeling the lexical-semantic connections at the level of the phonological store (Acheson, Postle & MacDonald, 2010), the episodic buffer’s theoretical connection to long-term memory makes it an ideal candidate for this function. Double dissociation of phonological span deficits and semantic span deficits have been found in case studies of traumatic brain injuries which suggests that the phonological loop is separate from connections to lexical-semantic information (Martin, Shelton & Yaffee, 1994). Using electroencephalogram (EEG) and a short-term verbal memory task requiring a semantic judgment in the recall phase, Cameron, Haarmann, Grafman and Ruchkin (2005) found that temporary retention of verbal information is accompanied by a prolonged activation of the semantic representations in long-term memory, which was determined based on the prolonged activation of posterior association cortices. Further, Cameron et al. (2005) found semantic relationships between the incidental probes and words to be retained increased the speed at which participants were able to recognize the words to be retained. This indicates a priming of information in the long-term storage of the language system, which then facilitates recognition. These findings relate back to those previously discussed with respect to facilitation of information in the long-term memory system of the language
production system. Within Baddeley’s updated model of working memory, the episodic buffer has reciprocal relationships with the phonological store and with long-term memory (Repovš & Baddeley, 2006). Through these reciprocal relationships, the episodic buffer is able to act as an intermediary in which the binding of both sources of information can be stored and then utilized (Baddeley, 2000; Baddeley, Allen & Hitch, 2010; Baddeley & Hitch, 2000).

In terms of the episodic buffer development, there is a paucity of information as this is a relatively new area of research. Children age 4 to 6 show a system differentiated from, but associated with, the phonological loop and central executive, which relates to Baddeley’s description of the episodic buffer (Alloway et al., 2004). This suggests that early on, this system is in place.

In order to measure the episodic buffer, the current study utilized the Competing Language Processing Task (CLPT; Gaulin & Campbell, 1994), a task developed from the sentence span paradigm from Daneman and Carpenter (1980). This task requires the integration of representations from working memory, short-term memory, and the long-term language memory system, which is the current standard for testing the episodic buffer (Alloway et al., 2004; Baddeley & Wilson, 2002). Healthy children’s performance on the veracity component in the CLPT is consistently high across ages; however, there is a strong developmental trend in extent of target word recall (Gaulin & Campbell, 1994). Performance on the CLPT in children predicts sentence imitation accuracy (Poll et al., 2013), and complex sentence comprehension (Montgomery & Evans, 2009). The ability to parse and segment language into smaller units as well as children’s semantic knowledge (vocabulary) significantly predict children’s scores on the CLPT (Mainela-
Arnold, Evans & Coady, 2010; Mainela-Arnold, Misra, Miller, Poll & Sook Park, 2012). The CLPT word recall score differentiates children with Specific Language Impairment (SLI) from typically developing controls (Mainela-Arnold & Evans, 2005; Weismer, Evans & Hesketh, 1999). Children with SLI, despite high levels of scores on veracity component, recall fewer low frequency words on the CLPT compared to healthy peers (Mainela-Arnold & Evan, 2005).

**Evidence of working memory’s relationship to language.** In healthy adults, speech production difficulties increase with experimental constraints on attention and working memory (Kemper & Sumner, 2001; Kemper, Herman & Lian, 2003; Kerns, 2007; Jou & Harris, 1992; Oomen & Postma, 2001; 2002). When working memory is constrained by task demands, reduction in sentence length, grammatical complexity, propositional density and speech rate are methods of managing the increased demands (Kemper & Sumner, 2001; Kemper et al., 2003). These findings indicate working memory is important for our ability to use complex forms of language.

Jou and Harris (1992) examined the effects of limited resources on error detection in others’ speech in healthy adults. In the first condition, participants heard passages read aloud and then verbally recalled what they had heard. In the dual task condition, the participants had to perform listen to 15 single digit numbers presented through earphones and add the numbers cumulatively while verbally recalling a short story. With divided resources between tasks, participants’ quantity of verbal production reduces (i.e. lower recall), as does quality of production (Jou & Harris, 1992). In the dual-task condition, participants made more frequent and longer (minimum 5 seconds) within-clause pauses, which are not typical of normal speech (Jou & Harris, 1992). In normal spontaneous
speech, integrity of clauses is typically maintained and the maximum length of between-sentence pauses is 2.5 seconds (Grosjean, 1980). The authors suggest that this may reflect that when completing a clause, greater attentional resources are required (Jou & Harris, 1992). A further finding was that the number of sentence fragments tripled in the dual task condition. The authors speculated that participants had difficulty retrieving the correct lexical items in order to complete the sentence as well as having difficulty monitoring what they had already said (Jou & Harris, 1992). Both of these difficulties would result in the participant abandoning the sentence and restarting (Jou & Harris, 1992). The inability to connect to long-term language stores while monitoring articulation likely reflects constraints on the episodic buffer.

Finally, in the dual task condition, participants retraced with greater frequency (e.g., repeated words, phonemes, etc.), which the authors suggest may be an attempt to keep track of information previously said or to provide time to plan the next clause (Jou & Harris, 1992). During other divided attention, dual task paradigms (tactile-form recognition task and story telling), speakers made more pauses and repetitions (Oomen & Postma, 2001). As the need to allocate cognitive resources for access to long-term storage for other tasks simultaneously (i.e. working memory load) increased, speakers had more difficulty maintaining online comparison of the prearticulatory and postarticulatory plan to that of the intended speech plan. This resulted in the need to restart the plan from the beginning or stop speech in order to make the comparison. Again, the load on the episodic buffer of Baddeley (2000) is the likely barrier to connecting real-time articulation to long-term memory stores.
Oomen and Postma (2002) found that limiting processing resources, by a dual task paradigm (i.e., random finger tap generation and speech production), leads speakers to generate more errors, detect errors more quickly, and repair fewer errors. The results suggest limited processing resources lead to reduced accuracy in error detection (Oomen & Postma, 2002). Further, the authors suggest that the faster speed of error detection under the dual task paradigm may reflect that the speakers reduced the time spent monitoring their speech, which resulted in faster speed but also increased the rate of errors (Oomen & Postma, 2002). The speed-accuracy trade-off may reflect that the speakers chose to focus their attention on the faster process of prearticulatory monitoring, which may be more autonomous from the central resources, and thus, less susceptible to the central resource demands (Oomen & Postma, 2002). Previous research has found that speakers can shift their attention to different components of speech output and focus on specific types of errors depending on the context (Motley, 1980; Motley, Baars & Camden, 1981; Power, 1985), and thus, can reduce the time necessary for monitoring.

Oomen and Postma (2002) examined the effects of limiting processing resources on error detection in other-produced speech in the same group of young, healthy adults. In this test of the perception based monitoring, using a dual task paradigm (random finger tap generation while listening to a speech sample), participants detected a smaller number of errors in the speech sample and latency of error detection increased (Oomen & Postma, 2002). Increased latency of error detection may result from interference with the motor aspects of the random finger-tap generation task (indicated errors by tapping a key) or from limited resources reducing processing speed in one modality (Oomen & Postma, 2002). Overall, the authors concluded that on divided attention tasks, a reduction in the
ability to detect errors results in fewer repairs (Oomen & Postma, 2002). Further, the authors proposed that speaking and monitoring your own speech is likely to be a more challenging task than only monitoring another person’s speech, due to its dual task nature (Oomen & Postma, 2002). Supporting this proposal, participants monitoring their own speech did less well at the distracter task than when monitoring others’ speech (Oomen & Postma, 2002). Overall, the adult research literature suggests that if working memory and attentional resources are constrained, then speech monitoring is impaired. The present study expands this literature and investigates the role that working memory has on speech monitoring within healthy children.

Dissociation of Language and Working Memory

A critique of the literature on working memory’s relationship to language is the use of tasks that require language to test working memory (Mainela-Arnold, 2013, personal communication). The relationship of these tasks to language tasks may simply be due to the shared output mechanism (i.e. speech production). If working memory is simply the information from long-term storage that is the focus of attention, then verbal working memory could be defined as activated long-term language knowledge (Cowan, 1999). However, the likely main distinction between verbal working memory and language would be the level of attentional control required for the task, which is consistent with the previously reported findings regarding semantic control (Cowan, 1999; Leonard et al., 2007).

There is currently a paucity of research on methods by which to directly differentiate the two cognitive functions, which can make interpretation of the literature supporting the relationship challenging. One solution has been to use working memory
tasks that use spatial or non-nameable (abstract) objects. Nystrom et al. (2000) attempted to dissociate the prefrontal brain regions activated by different stimuli on fMRI during \( n \)-back tasks using letters, shapes and locations. At the highest level of working memory load (3-back) or level at which attentional resources need to be divided between the primary task and other tasks to activate different information from long-term memory, the same approximate cortical areas responded to all types of stimuli (Nystrom et al., 2000). This may indicate that with sufficiently challenging tasks, the brain must engage a core set of regions to provide the attention and working memory required, regardless of the type of stimuli. If this initial finding is indeed reflective of the true relationship, then verbal working memory is separate from the core language systems. That said, these findings need to be replicated and expanded in order to make conclusions.

There is also evidence for verbal working memory’s unique contribution to language determined via statistical analysis. A longitudinal study of the relationship between working memory, verbal abilities, reading comprehension component skills (inference making, comprehension monitoring, story structure knowledge) and reading comprehension in children at ages 8, 9, and 11 found that working memory uniquely predicted reading comprehension once verbal abilities and component skills were removed (Cain, Oakhill and Bryant, 2004). Working memory was reported to uniquely contribute to reading and language acquisition in a second language beyond the explanation provided by attention, naming, and phonological processing in children in grades 1 to 3 (Swanson, Orosco, Lussier, Gerber & Guzman-Orth, 2011). In adults, working memory uniquely predicted significant variability in oral metaphor production, a language task thought to be more challenging, when vocabulary knowledge and exposure
to print materials were controlled (Chiappe & Chiappe, 2007). These preliminary findings may provide support for working memory’s unique contribution to language beyond the base language system. While not the focus of the current study, more research is needed to provide solid support for the unique contribution of verbal working memory to language beyond the core language abilities.

**Normal Language Development and Working Memory**

As previously mentioned, working memory has a well-established role in language development. In healthy 3-year-olds, greater proficiency in verbal working memory predicted higher productive vocabulary, greater use of sophisticated grammatical structure in spontaneous speech and increased mean length of utterances (Adams & Gathercole, 1995; Blake, Austin, Cannon, Lisus & Vaughan, 1994). The ability to make use of the phonological loop at age 4 predicts vocabulary knowledge at age 5 but not vice-versa, which suggests that working memory is an essential aspect of language acquisition (Gathercole & Baddeley, 1989). Further, in 4- and 5-year-olds with normally developing language, phonological working memory predicts sentence length and information provided when retelling a narrative story (Adams & Gathercole, 1996).

Adams and Gathercole (1996) proposed that verbal working memory might be necessary for language development as an output buffer for phonological information. The output buffer aids in the programming of an articulatory plan, especially prior to the development and automatic activation of long-term phonological representations. This proposal relates back to Levelt’s (1983, 1989) theory that working memory is necessary to maintain and monitor the articulatory plan during spontaneous speech, which requires
access to the long-term memory stores. This integrative role again is consistent with Baddeley’s (2000) episodic buffer.

Prior to the age of 7, children have been found to rely on the sensory aspects of verbal information that is presented visually (e.g., picture of a red balloon) whereas older children begin to rely on the phonological loop to support recall of the verbal material (e.g., name of the object) (Hitch, Halliday, Schaafstal & Scraagen, 1988). However, by the age of 6, children begin to show the individualization of the components of the tripartite model of working memory (central executive, visual-spatial sketchpad, and the phonological loop) (Gathercole, Pickering, Ambridge & Wearing, 2004). As children develop, they are better able to maintain larger amounts of verbal material in the phonological store as they increase their rate of rehearsal of information (Hulme, Thomson, Muir & Lawerence, 1984). Further, the association between the central executive and the phonological loop increases as children age, which may reflect developmental increases in processing efficiency (Gathercole et al., 2004; Jarvis & Gathercole, 2003). The present study examined the impact of developmental differences in working memory on speech monitoring in two age groups, 6 to 7 and 10 to 12. The basis of the choice of these two age groups was their significantly different abilities with respect to multiple aspects of working memory (Gathercole et al., 2004).

**The Present Study**

In the present study, on a computer screen, children saw a map of a network of coloured pictures (depicting simple everyday objects) connected with lines and they described the route a moving dot took on the map. In the first condition, the children described the route the dot took along the lines and pictures of the network. In the
experimental conditions, children completed the same route description task (with different routes) with (1) working memory demand, (2) white noise, and (3) a combination of both constraints.

White noise is not believed to impact working memory; however, this constraint was included in the current study as noted previously, white noise may prevent a person from using the auditory loop to monitor for errors. As such, the combination of both in the third condition will potentially allow for the isolation of the impact of working memory demands on the predominately production-based prearticulatory monitoring system as posited by Oomen and Postma (2002). White noise or irrelevant noise is a commonly experienced phenomenon when speaking (e.g., in a classroom). One must inhibit noise in order to attend selectively to the target stream of speech as a speaker and a listener. Thus, for the purposes of providing recommendations for individuals coping with working memory demands, it is useful to gain understanding of how these constraints affect speech monitoring in realistic conditions.

Speech error rates were expected to be higher for all children in the working memory condition and the dual-task (working memory and white noise) condition than in the control condition due to reduced capacity to monitor. The younger children were predicted to have the highest rate of errors in all conditions. White noise increases the rate of speech errors due to reduced auditory monitoring (Oomen & Postma, 2002), thus, the rate of speech errors during this task was expected to be higher than the control condition.

Under all three experimental conditions, the number of covert (based on speech interruption with no error) and overt self-repairs was expected to be less than in the
control condition. The rate of covert and overt self-repairs was predicted to be lowest in the combined spatial span and white noise condition due to the higher constraints on working memory. Younger children were expected to make the least amount of self-repairs in this condition due to lower working memory capacity.

As the cognitive demands of language production increase, the semantic work completed decreases and latency of response increases (for review, see Oomen & Postma, 2002), which likely is due to the load on the episodic buffer and reduced ability to access long-term memory storage. Thus, it was hypothesized that the speed of response (error to cutoff, cutoff to repair, and speech rate) would increase as task complexity increased except for the younger children whose working memory subsystems are less independent and more constrained by central resources (as posited by Levelt, 1983).

Each child completed a series of working memory span tasks: digit span, listening span, and spatial span. As previously discussed, verbal working memory span is closely related to language performance and as such, it was predicted that children with higher verbal working memory spans would have lower rates of speech errors and higher rates of repairs per errors. The spatial working memory span task was included as a method by which to increase working memory load (i.e. requirement of allocation of resources away from the language task to another stimulus for the purpose of activating other information) for the language tasks in a manner that would not elicit basement effects in the younger children due to overload on their language resources. Consistent with previous findings, it was hypothesized older children in comparison to younger children would have large spans on all tasks.
Chapter 2
Method

Participants

Volunteer participant referrals were elicited via flyers, rack cards and ads sent to and posted in 3 medical/developmental clinics, 4 community centers, family-focused community organization websites, university advertising boards, print and online media advertisements (e.g., parenting website), 2 private schools, and via Facebook ‘like’ connections.

Previous research examining spontaneous speech in children, Evans (1985) found a medium effect \((r = 0.35)\) for the difference between children in the second grade and kindergarten with respect to the number of repairs. Verhoeven (1989) found a large effect \((\eta^2 = 0.24)\) for the difference between children aged 6-8 from children aged 8 in making repairs to speech errors in their second language. Non-spontaneous speech production research using adults found a large effect size \((r_{pb} = 0.59)\) for the difference between the speech only and dual-task on rate of speech (Oomen & Postma, 2002). These authors also found a medium effect size \((r_{pb} = 0.48)\) for the difference between speech only and dual task condition on percentage of repaired errors (Oomen & Postma, 2002). The present study used a conservative approach in estimating the possible effect size, due to a paucity of literature examining non-spontaneous speech in children.

However, due to challenges in recruiting children who met the exclusion criteria for the study, the sample of the current study included 22 participants (eight in the age 6-8 group, fourteen in the age 10-12 group). The exclusion criteria for the study was an absence of (a) a history of a chronic medical and/or psychological illness; (b) a history of
trauma to the brain; (c) a history of hearing impairment; (d) a history of learning disability; (e) a history of school non-attendance; (f) non-fluency in English; and (g) colour blindness. Initial screening for the stated exclusion criteria took place over the phone and resulted in the exclusion of five children. Three other children, one in the young group and two in the older group, were also excluded from the analysis due to severe attentional difficulties during testing in two cases and one instance of recording equipment problems resulting in an incomplete language sample. Intelligence and basic language skill testing was not completed in this sample as children in this age range have limited attention and the requirement of multiple sessions would have increased the likelihood of attrition. Testing sessions took between 60 minutes and 85 minutes, depending on the age of the child, with younger children needing more breaks to ensure full attention on the tasks. Intelligence and basic language skill testing was not completed in this sample as children in this age range have limited attention and the requirement of multiple sessions would have increased the likelihood of attrition. The strict screening requirements provided sufficient in ensuring a sample of typically developing children.

Participants received five dollars and small prizes (e.g., stickers, bouncy balls, etc.) for their participation. All participants were treated in accordance with principles 6.6-6.20 of the “Ethical Principles of Psychologists and Code of Conduct” (APA, 1992). The parent/legal guardian of the child provided consent and the child provided assent (see Appendix A for consent form and Appendix B for assent form). Guardians also completed a brief background questionnaire to collect information on date of birth, gender, country of origin, languages spoken at home, primary language, and grade in
school. Table 1 lists the descriptive statistics regarding the sample’s language and schooling history per group.

**Measures**

At the beginning of the session, each child completed three working memory tasks: Digit Span from the Wechsler Intelligence Test for Children – Fourth Edition (WISC-IV; Wechsler, 2003), the Competing Language Processing Task (CLPT; Gaulin & Campbell, 1994), and the experimental Spatial Span task.

**Digit span.** In the digit span task, the examiner presented the child with spoken sequences of digits that the child was asked to recall in the same order (Digits Forward) and then, with different sequences, in reverse order (Digits Backward) (Digit Span - WISC-IV; Wechsler, 2003). The sequences of digits are random and the length of the sequence increases by one digit for every two items the child completes successfully (Wechsler, 2003). A practice session ensured the child’s understanding of the task. The child continued until unable to complete two sequences of the same length (Wechsler, 2003). The score for Digits Forward and Digits Backward was the number of items successfully completed (Wechsler, 2003). The total Digit Span score was the sum of Digits Forward and Digits Backward (Wechsler, 2003).

**Listening span.** In the listening span task, the examiner presented the child with a series of spoken short sentences (Competing Language Processing Task (CLPT): Gaulin & Campbell, 1994). The child judged the veracity of each sentence in turn by responding “yes” or “no,” and then later recalled the final word of each of the sentences within a group (Gaulin & Campbell, 1994). Level 1 of the CLPT, requires the child to comprehend one statement and then recall one word (Gaulin & Campbell, 1994). For
each subsequent level, the number of statements increased by one to the maximum number of six statements (Gaulin & Campbell, 1994). Half of the statements are true and half are false (Gaulin & Campbell, 1994). “Each statement contains three words (subject-verb-object, subject-verb-modifier, or subject-auxiliary-main verb)” (Gaulin & Campbell, 1994, p. 57). The total number of words correctly recalled was used as the measure of working memory in the present study.

Spatial span. The spatial span task, developed by the primary author, used a laminated board separated into quadrants of different colours and patterns. The examiner placed the board in front of the participant with the red square in the top right quadrant. The examiner demonstrated pointing to the quadrants in a specific sequence at a rate of one square per second and then asked the participant to repeat that sequence (See Figure 1 for picture of spatial span board). For the practice trials for the spatial span task, the examiner demonstrated the task and provided corrective feedback to the participant. The examiner could repeat the practice trials as many times as necessary to ensure the understanding of the participant; however, the standard practice trials were sufficient for all participants. The practice trials and the first scored level have a sequence of two quadrants. There are two items in each level. Each subsequent level had an increase of one quadrant in the sequences. The participant continued through the levels until s/he reached a level where s/he was unable to successfully perform both items at the level. The spatial span score was the total number of items completed successfully. The last level where the participant was able to complete both items at a level successfully was the level for the experimental concurrent task (e.g., if a participant is able to successfully
complete a level with a sequence of 4 quadrants, then s/he will do the network task with
the concurrent task using 4-quadrant sequences).

**White noise.** The Simply Noise © Signature White Noise ITunes© Application
was used to generate the white noise for the white noise condition (Reactor LLC, 2011).
This white noise was a synthetic noise containing every frequency within the range of
human hearing in equal amounts (personal communication, Reactor LLC, 2011). The
white noise was transmitted via an Apple iPod touch © on headphones that were
disinfectected between participants’ use. The volume level was set to 40 decibels within the
Simply Noise © program for all participants. This volume corresponds to a level of 50
decibels or the level associated with a casual conversation (Olsen, 1998).
Figure 1. Spatial span board for experimental spatial span task created by primary author. Reproduction without permission from primary author is prohibited.
Using pictures of objects with which children in the local culture would be familiar, a modified version of the network task used in Levelt (1983) as well as Oomen and Postma (2002) was developed. The network task consisted of 22 networks, each consisting of nine coloured pictures depicting simple everyday objects (see Figure 2 for sample network). The number and type of connections (e.g. curvy versus straight) between the pictures was kept consistent between networks. The networks were all tested on five pilot participants and modifications were made as needed to the tasks and procedures based on the feedback from the pilot participants regarding network difficulty and condition level difficulty as well as examiner observations. Pilot participant data was used only for modifying the networks and procedures and for training the second rater.

In the current study, each participant was shown each picture separately before the practice trials and asked to name the object in the picture to ensure familiarity with all objects. If needed, the examiner could provide the expected term for the picture; however, this was not necessary for any of the subjects in the present study.

One or more straight or curved lines connected the pictures in the networks. A blue dot moved through the network on the lines, indicating the route the child followed in her/his verbal description. As the dot arrived at a picture, a blue outline appeared around the picture to draw the child’s attention to that picture. A route through the network always consisted of 10 steps. The rate of the dot was set at a specified rate, similar to that of a normal speech rate, 53 seconds for a network (Oomen & Postma, 2002).

The networks were presented on a MacBook Air © 1.8 GHz Intel Core i5 laptop. The graphic design and programming completed by Annual Golf Designs ©. Speech was
recorded using the software, Audacity © for Mac OS X (Version 1.2.5). A back-up was created via a video recording. Recorded speech was transcribed and analyzed via the speech analysis program, ELAN (Wittenburg, Brugman, Russel, Klassman & Sloetjes, 2006). ELAN was used as the speech analysis program based on the recommendations of researchers in this area (Bavelas, 2012, personal communication).
Figure 2. Sample network of network task

A blue ball would begin on green square and end on red square. This network is the sample network demonstrated by the examiner for the participant. Network description is as follows “Start. It goes on the black straight line to the apple. Then it goes on the red curved line to the guitar. It goes on the yellow straight line to the puppy. Then on the purple curved line to the brown straight line to the telephone. It goes on the green straight line to the light bulb. Then on the blue straight line to the yellow curved line to the bike. It goes on the brown straight line to the computer. Then it goes on the purple straight line to the red straight line to the helmet. Then on the brown curved line to the flower. Then on the purple straight line to the computer. Then on the brown straight line to the bike. And then on the black straight line to the end.” Networks were adapted from Levelt (1983) and Oomen and Postma (2002).
Procedure

The following procedure is a modified version of the procedure outlined in Levelt (1983) and experiment 1 of Oomen and Postma (2002). All participants completed all four network conditions (control-network task only; network task and working memory demand; network task and white noise demand; and network task, working memory, and white noise demands).

In the network only task, participants were told “We are now going to do a task on the computer. You will see a group of pictures connected with different coloured lines. There will also be green and red squares. The green square is the start and the red is the end. You will describe the path the blue ball takes on this network of pictures. I will show you on this sample network.” The examiner provided a complete description of the path the ball took on the sample network. When the examiner made an error in the verbal description, s/he provided a speech repair and picked up the description at the ball’s current position. The examiner demonstrated at least one error and repair to each child with a maximum of two errors and repairs. After the sample network, the examiner said to the child, “Okay, now that you have seen how to do it, I want you to try one. Be sure to say start at the beginning and then describe the path that the blue ball takes through the pictures. Tell me the colour of the line and whether it is straight or curved and then tell me the name of the picture. Just like I did on the sample. If you make a mistake, you can correct it but keep going as the ball moves quickly. Try not to fall behind. If you do, pick up where the ball is and go from there. Ready to try one?” Each child completed two practice networks and the examiner provided feedback to the child
about any missing descriptors such as line colour, line trajectory (straight or curved), or pictures and also provided encouragement. For all experimental trials, the experimenter provided encouragement when necessary during the task (e.g., “Just pick up where you left off.”) and rewards at the end of trials (i.e. prizes).

After the practice trials, for the network only task, the examiner provided the following instructions to the child, “Now that you know how to do the task and have practiced it, we are ready to go onto the real ones. They are just like the ones we have done so try your best and remember if you fall behind, just pick up where the ball is at that time.” The child completed networks 1 to 5 for the network only task. The same networks were for each condition for all children in order to ensure that comparisons between children per task were based on the same networks in an attempt to reduce error variance within a task that could potentially result from small differences between the networks.

In the working memory demand condition, the examiner demonstrated a spatial span sequence at the level determined during the initial spatial span task where the child was able to successfully complete 2 items at that level. The participant was asked to remember the spatial span sequence as s/he would be asked to show it later. The participant then proceeded to describe a network. After providing the description, the participant showed the sequence that the examiner had demonstrated earlier. For the network task and working memory condition, the participant completed networks 6 to 10. The number of correct sequences within the condition was used as the working memory condition spatial span score.
In the white noise condition, participants listened to the white noise while completing the network description task and completed networks 11 to 15. In the working memory demand and white noise demand condition, participants listened to white noise while completing the spatial span task and the network description task. The number of correct sequences within the condition was as the network task, white noise and spatial span score.

**Analysis**

For each network of the twenty networks (5 per condition), language errors were coded as phonological, omission, lexical, syntax, or exchange errors. Repairs were coded based on the type of error that had been made. Covert repairs included hesitations, and repetitions, as per Levelt (1983). Pauses, while coded as covert repairs in Levelt’s (1983) coding scheme, were coded separately given the premise that younger children may be less skilled in re-initiating the network task after making an error (Singh, 2007). For a full description of the error and repair coding schemes used, see Appendix A and B.

The author coded all samples. A reliability check of the coding of the errors was completed by a clinical child psychologist with a background in assessment on a random sample of 11 subjects’ language samples. A block randomization procedure was used to generate the list of the 11 subjects’ language samples to be checked (Suresh, 2011; [http://www.graphpad.com/quickcalcs/index.cfm](http://www.graphpad.com/quickcalcs/index.cfm)). Initial training for the error analysis of the rater was completed using pilot subject data that was not used in the main analysis. The rater coded 10 networks of pilot subject data and codings were reviewed by the author. Any discrepancy from the coding system was discussed before the rater completed a further 10 pilot networks. No changes were made to the original coding
system during this process. The rater was provided with the random sample of 11 language samples (20 networks per sample, no identifying information provided). A two-way mixed model intraclass correlation coefficient was calculated to determine the absolute agreement between the raters for the error analysis. The Single Measures ICC was 0.969 with a 95% CI of .735 to .993 (see Table 1 for per network ICC confidence intervals). As per the ICC strength descriptions of Portney and Watkins (1993), ICCs greater than .75 as good and ICCs greater than .90 as reasonable for clinical measures. Therefore, it was concluded that based on the overall ICC, the ratings used in this analysis were acceptable and no adjustments to the coding protocols were made. However, it should be noted that for network 4 and network 18, the lower bound of the confidence interval was lower than the .75 (.745, .704 respectively). In order to obtain ICC confidence intervals with lower bounds above the .75 cut-off, the raters could have been further trained and all networks included in the rating.

While the ICC was acceptable, the agreement between the raters was not 100%. The difference between raters in the number of errors rated ranged from 0 to 29 errors. A graphical examination of the ratings did not reveal any obvious systematic (bias) in the ratings based on error type or participant. Sources of rater variability can arise from the following sources; (a) degree to which raters comply with the scoring rubic; (b) manner in which the rater interprets criteria employed in training scoring sessions; (c) degree of severity and leniency exhibited when scoring; (d) level of understanding and use of rating scale categories; and (e) degree to which ratings from one rater are consistent across examinees, scoring criteria, and performance tasks (Wang, 2010, p. 110). Essentially, variability between raters can reflect variability in the data set due to individual raters and
not due to the experimental manipulation. Experimenters using small samples, such as the one in the current study, should use greater caution with respect to making conclusions in the context of the possibility of non-experimental variability.
Table 1. Intraclass correlation values per network with 95% confidence intervals

<table>
<thead>
<tr>
<th>Network</th>
<th>ICC single</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>0.977</td>
<td>.920-.994</td>
</tr>
<tr>
<td>N2</td>
<td>0.958</td>
<td>.861-.988</td>
</tr>
<tr>
<td>N3</td>
<td>0.961</td>
<td>.869-.989</td>
</tr>
<tr>
<td>N4</td>
<td>0.92</td>
<td>.745-.978</td>
</tr>
<tr>
<td>N5</td>
<td>0.954</td>
<td>.848-.987</td>
</tr>
<tr>
<td>N6</td>
<td>0.978</td>
<td>.923-.994</td>
</tr>
<tr>
<td>N7</td>
<td>0.972</td>
<td>.905-.992</td>
</tr>
<tr>
<td>N8</td>
<td>0.973</td>
<td>.907-.992</td>
</tr>
<tr>
<td>N9</td>
<td>0.978</td>
<td>.924-.994</td>
</tr>
<tr>
<td>N10</td>
<td>0.974</td>
<td>.911-.993</td>
</tr>
<tr>
<td>N11</td>
<td>0.945</td>
<td>.819-.985</td>
</tr>
<tr>
<td>N12</td>
<td>0.972</td>
<td>.904-.992</td>
</tr>
<tr>
<td>N13</td>
<td>0.946</td>
<td>.823-.985</td>
</tr>
<tr>
<td>N14</td>
<td>0.958</td>
<td>.860-.988</td>
</tr>
<tr>
<td>N15</td>
<td>0.954</td>
<td>.847-.987</td>
</tr>
<tr>
<td>N16</td>
<td>0.966</td>
<td>.884-.990</td>
</tr>
<tr>
<td>N17</td>
<td>0.949</td>
<td>.832-.986</td>
</tr>
<tr>
<td>N18</td>
<td>0.906</td>
<td>.704-.973</td>
</tr>
<tr>
<td>N19</td>
<td>0.961</td>
<td>.868-.989</td>
</tr>
<tr>
<td>N20</td>
<td>0.951</td>
<td>.836-.986</td>
</tr>
<tr>
<td>Network Total</td>
<td>0.969</td>
<td>.895-.991</td>
</tr>
</tbody>
</table>
Statistical analysis was completed using IBM © SPSS © Statistics – Version 21, Macintosh version. Power analysis was completed with G*Power 3.1 (Faul, Erdfelder, Buchner & Lang, 2009).
Chapter 3

Results

Eight children (5 males, 3 females) in the 6-8 year old age range and 14 children (9 males, 5 females) in the 10-12 year old range participated in the study. Three other children, one in the young group and two in the older group, were not used in the analysis due to one instance of recording equipment problems resulting in an incomplete language sample and 2 children with severe attentional difficulties that were felt to compromise their ability to complete the tasks. Table 2 lists the descriptive statistics regarding the sample’s language exposure and current school information by group. Summary scores of errors, repairs and pauses across the five networks within each condition were calculated for each condition and for each group (see Table 3 for all raw error, repair, repair per error, and pause values).
Table 2. Language and Schooling Demographic Information for Children Aged 6-8 and 10-12 as Frequency Counts.

<table>
<thead>
<tr>
<th>Demographic Factor</th>
<th>Children aged 6-8 (n = 8)</th>
<th>Children aged 10-12 (n = 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Hebrew</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Primary language spoken at home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Additional languages spoken at home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>French</td>
<td>2*</td>
<td>0</td>
</tr>
<tr>
<td>Hebrew</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Spanish</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>None</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Grade at school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kindergarten</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Grade 1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Grade 2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Grade 4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Grade 5</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Grade 6</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Grade 7</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Language of classroom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English primary</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>French immersion</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>French primary</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type of Classroom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single grade</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Mixed grades</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

* One child spoke 3 languages.
<table>
<thead>
<tr>
<th>Language characteristics</th>
<th>Network Condition</th>
<th>Children aged 6-8</th>
<th>Children aged 10-12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M (SE) (RANGE)</em></td>
<td><em>M (SE) (RANGE)</em></td>
</tr>
<tr>
<td>Total speech errors</td>
<td>Networks 1-5</td>
<td>174.63 (115.88) (4.00 – 383.00)</td>
<td>97.57 (87.97) (19.00 – 319.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>237.13 (159.94) (5.00 – 494.00)</td>
<td>92.35 (87.97) (19.00 – 319.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 11 - 15</td>
<td>241.13 (180.39) (3.00 – 548.00)</td>
<td>141.21 (120.33) (17.00 – 350.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>231.38 (202.53) (10.00 – 528.00)</td>
<td>132.46 (110.24) (15.00 – 306.00)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>221.06 (159.64) (22.00 – 1825.00)</td>
<td>115.95 (93.50) (85.00 – 1175.00)</td>
</tr>
<tr>
<td>Total speech repairs</td>
<td>Networks 1-5</td>
<td>9.63 (7.60) (1.00 – 22.00)</td>
<td>10.36 (5.23) (4.00 – 18.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>6.75 (5.01) (1.00 – 17.00)</td>
<td>9.79 (4.74) (4.00 – 17.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 11-15</td>
<td>4.88 (2.23) (2.00 – 8.00)</td>
<td>10.86 (7.92) (1.00 – 31.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>4.13 (2.36) (1.00 – 8.00)</td>
<td>10.64 (6.48) (3.00 – 26.00)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>6.34 (3.49) (1.00 – 22.00)</td>
<td>10.41 (5.25) (4.00 – 18.00)</td>
</tr>
<tr>
<td>Total speech repairs / errors</td>
<td>Networks 1-5</td>
<td>0.08 (0.08) (0.01 – 0.25)</td>
<td>0.16 (0.12) (0.05 – 0.44)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>0.05 (0.06) (0.01 – 0.20)</td>
<td>0.16 (0.08) (0.03 – 0.31)</td>
</tr>
<tr>
<td></td>
<td>Networks 11-15</td>
<td>0.11 (0.23) (0.00 – 0.67)</td>
<td>0.16 (0.15) (0.02 – 0.48)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>0.07 (0.13) (0.00 – 0.40)</td>
<td>0.14 (0.11) (0.02 – 0.48)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.07 (0.12) (0.00 – 0.25)</td>
<td>0.15 (0.10) (0.02 – 0.48)</td>
</tr>
<tr>
<td>Total pauses</td>
<td>Networks 1-5</td>
<td>12.25 (15.46) (0.00 – 43.00)</td>
<td>4.86 (6.93) (0.00 – 26.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>18.75 (15.93) (0.00 – 49.00)</td>
<td>4.71 (6.63) (0.00 – 22.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 11-15</td>
<td>20.63 (19.16) (2.00 – 59.00)</td>
<td>10.36 (9.93) (0.00 – 37.00)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>25.50 (25.31) (3.00 – 75.00)</td>
<td>10.07 (7.33) (0.00 – 20.00)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19.28 (14.58) (0.00 – 43.00)</td>
<td>7.50 (7.18) (0.00 – 20.00)</td>
</tr>
</tbody>
</table>
**Speech errors.** As seen in table 3, it should be noted that a large degree of variability existed in number of errors between and within the groups. This degree of variability may reflect natural variability in the type of data studied (e.g. a wide range of normal working memory development at a particular age leading to a wide range of speech errors) or systematic variability as noted previously. No systematic variability was noted in the data; however, with a small sample size, caution must be used when analyzing and interpreting data with this level of variability as values at the extreme ends of the range can more greatly affect statistical outcomes.

A 2 (Age Group) X 4 (Network Conditions) repeated measures ANOVA was completed on the total errors. Box’s test of equality of covariance matrices (Box’s M) and Mauchly’s test of sphericity were not significant, $p > 0.05$. Levene’s tests of equality of error variances for networks 1-15 were not significant, $p > 0.05$; however, for network 16-20, Levene’s test for equality of error variances was significant, $p < 0.05$, indicating that error variance in speech errors differed between the groups for networks 16-20. A square-root transformation to the total error scores for each of the conditions was completed to correct for this heterogeneity of variances (Howell, 2010; Cardinal & Aitken, 2006). On the transformed scores, Box’s M, Mauchly’s test of sphericity and Levene’s tests of equality of error variances were all not significant, $p > .05$. The use of ANOVA with transformations was deemed to be appropriate as while the range of the data in each of the groups was variable, it was felt that this reflected true variability in the construct (Howell, 2010). Transformations are a commonly used and acceptable method by which to reduce the measurement range in order to meet the assumptions of the
ANOVA (Howell, 2010; Osborne, 2002). Another approach to challenges with equality of error variances in the context of a small sample is a non-parametric test; however, conversion of data into rankings or other non-score data results in a significant loss of information (e.g. mean, standard error) and non-parametric tests have less power than parametric (Mumby, 2002).

The main effect of network condition on the square-root transformed total errors, $F(3, 60) = 3.892, p = .01, \eta_p^2 = .16, f = .44, I-\beta_o = .80$ was significant. The main effect of age group, $F(1, 20) = 2.94, p = .10, \eta_p^2 = .13, f = .39, I-\beta_o = .37$ and the effect of the Age Group X Network Condition interaction, $F(3, 60) = 1.817, p = .15, \eta_p^2 = .08, f = .42, I-\beta_o = .45$ were not significant. Simple contrasts using the network task, white noise and spatial span condition (condition 4) as the comparison condition revealed a significant difference between the control condition and condition 4, $F(1, 20) = 5.683, p = .03, I-\beta_o = .62$. No other contrasts were significant; however, power was low for both the condition 2 versus condition 4 contrast, $I-\beta_o = .13$ and the condition 3 versus condition 4 contrast, $I-\beta_o = .10$. In the non-transformed scores, different trends of errors were evident between the two age groups. Children in the younger age group had a substantial increase in errors with the introduction of working memory demands, which remained consistent across the remaining two conditions. Children in the older group made substantially more errors in the last two conditions after white noise was introduced; whereas in the first two conditions, errors made were similar (see Figure 3 for patterns of non-transformed total error scores).

**Speech repairs per errors.** A 2 (Age Group) X 4 (Network Condition) repeated measures ANOVA was completed for total repairs per errors. A per error correction was
used to account for the fact that higher levels of working memory are associated with fewer errors and therefore, there are fewer errors to repair. All repair scores were divided by errors prior to analysis. Box’s M was not significant, $p > .05$. Mauchly’s test of sphericity was significant, $p = .001$. As the Greenhouse-Geisser estimate of sphericity ($\varepsilon$) $= .644$, the Greenhouse-Geisser correction was used (Howell, 2010). Levene’s test for equality of error variances was not significant for any of the network conditions, $p < .05$.

The main effect of network condition, $F(3, 60) = .518, p = .59, \eta_p^2 = .03, f = .16, I-\beta_o = .13$, the main effect of age group, $F(1, 20) = 2.56, p = .13, \eta_p^2 = .11, f = .35, I-\beta_o = .33$ and the effect of the Age Group X Network Condition, $F(3, 60) = 0.38, p = .68, \eta_p^2 = .02, f = .14, I-\beta_o = .11$ on repairs per errors were not significant. In the uncorrected (for errors) repair scores, children in the younger age group demonstrated a linear decrease in repairs as the condition difficulty increased; whereas children in the older age group demonstrated a consistent level of repairs across the conditions. In the corrected (for number of errors) repair scores, the pattern in the younger children was different and demonstrated higher repairs per errors in the network task only (control) and network task and white noise conditions (condition 3) (see Figure 4 for patterns of mean repair per errors scores).
Figure 3. Total errors per group per condition

Note: Error bars represent the 95% confidence interval.
Figure 4. Total repairs per errors per group per network condition

Note: Errors represent the 95% confidence interval.
Speech pauses. A 2 (Age Group) X 4 (Network Conditions) repeated measures ANOVA was performed on the total pauses. Box’s M and Mauchly’s test of sphericity were significant, $p < .05$. Levene’s test of equality of error variances was significant for network conditions 1-5, 6-10, and 16-20. Based on slope and power values determined using a spread versus level plot, Levene’s test, and Tukey’s power ladder, a power transformation of $(\text{network condition pauses} + 0.5)^{1/8}$ was determined to be the most appropriate method by which to correct the differences in variance and covariance (Tukey, 1977; Howell, 2010). For the transformed data, Box’s M and Levene’s test of equality of error variances were not significant. Mauchly’s test of sphericity was still significant, $p = .03$. As the Greenhouse-Geisser estimate of sphericity ($\epsilon = .729$), the Greenhouse-Geisser correction was used (Howell, 2010). The main effect of network condition was significant, $F(3, 60) = 10.780, p = .00, \eta_p^2 = .35, f = .73, 1-\beta_o = .99$. The main effect of group, $F(1, 20) = 2.814, p = .11, \eta_p^2 = .12, f = .37, 1-\beta_o = .36$ and the effect of the Network Condition X Group interaction, $F(3, 60) = 1.053, p = .36, \eta_p^2 = .05, f = .23, 1-\beta_o = .23$ were not significant. Simple contrasts using condition 4 as the contrast condition revealed significant differences between the control condition and condition 4, $F(1, 20) = 16.630, p = .00$ as well as between the spatial span and network task condition (condition 2) and condition 4, $F(1, 20) = 7.441, p = .01$. No other contrast was significant. Examination of the non-significant trends in the transformed scores, children in the younger age group demonstrated decreasing pauses as complexity of the task increased; whereas children in the older age group demonstrated a decrease of pauses on tasks where white noise was introduced (see Figure 5 for the pattern of transformed score of pauses).
Figure 5. Total transformed pauses \((\text{network condition pauses} + 0.5)^{1/8}\) per network condition per group.

Note: Error bars represent the 95% confidence interval.
**Speech repetitions.** A 2 (Age Group) X 4 (Network Condition) repeated measures ANOVA was completed for repetitions. A square root transformation was completed as Box’s M, \( p < .001 \), and Mauchly’s test of sphericity, \( p < .05 \), were significant. After the transformation, Box’s M was not significant, \( p > .001 \). Mauchly’s test of sphericity was significant, \( p = .021 \). As the Greenhouse-Geisser estimate of sphericity (\( \varepsilon \)) = .781, the Huynh-Feldt correction was used (Howell, 2010). Levene’s test of the equality of error variances was not significant for any of the networks, \( p > .05 \).

The main effect of network was significant, \( F(2.81, 60) = 4.073, p = .01, \eta_p^2 = .17, f = .45, 1-\beta_o = .80 \). However, neither the interaction, \( F(2.81, 60) = 1.871, p = .15, \eta_p^2 = .09, f = .31, 1-\beta_o = .44 \) or the group main effect, \( F(1, 20) = .120, p = .73, \eta_p^2 = .01, f = .10, 1-\beta_o = .06 \). Simple contrasts using condition 4 as the contrast condition revealed a significant difference between the control condition and condition 4, \( F(1, 20) = 8.231, p = .01 \). No other contrasts were significant. See Figure 6 for patterns of transformed repetitions (square-root) across network conditions by group and Table 4 for repetition and editing terms means and standard errors.

**Editing terms.** A 2 (Age Group) X 4 (Network Condition) repeated measures ANOVA on the editing terms was completed. Box’s M was not significant, \( p > .001 \). Levene’s test was not significant \( p > .05 \). Mauchly’s test of sphericity was significant, \( p = .02 \) and as the Greenhouse-Geisser estimate of sphericity (\( \varepsilon \)) = .675, the Greenhouse-Geisser correction was used (Howell, 2010). The main effect of network condition, \( F(2.03, 60) = 1.77, p = .18, \eta_p^2 = .08, f = .29, 1-\beta_o = .35 \), the interaction of group and network condition, \( F(2.03, 60) = 0.26, p = .78, \eta_p^2 = .01, f = .10, 1-\beta_o = .09 \), and the
Figure 6. Total transformed repetitions (square-root) per network condition per group.

Note: Error bars represent the 95% confidence interval.
Table 4. Means and standard errors of raw scores of repetitions and editing terms of language samples.

<table>
<thead>
<tr>
<th>Language characteristics</th>
<th>Network Condition</th>
<th>Children aged 6-8 (n = 8)</th>
<th>Children aged 10-12 (n = 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetition without error</td>
<td>Networks 1-5</td>
<td>5.38 (5.53)</td>
<td>3.57 (3.96)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>3.26 (4.80)</td>
<td>3.29 (4.08)</td>
</tr>
<tr>
<td></td>
<td>Networks 11-15</td>
<td>1.50 (0.53)</td>
<td>3.93 (6.90)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>1.50 (2.27)</td>
<td>3.50 (6.39)</td>
</tr>
<tr>
<td></td>
<td>Summary Score</td>
<td>11.63 (10.01)</td>
<td>14.29 (20.66)</td>
</tr>
<tr>
<td>Editing Terms</td>
<td>Networks 1-5</td>
<td>8.38 (8.50)</td>
<td>9.86 (11.84)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>9.38 (7.69)</td>
<td>9.57 (11.71)</td>
</tr>
<tr>
<td></td>
<td>Networks 11-15</td>
<td>9.88 (9.42)</td>
<td>9.71 (11.36)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>12.00 (13.44)</td>
<td>11.43 (10.37)</td>
</tr>
<tr>
<td></td>
<td>Summary Score</td>
<td>39.63 (36.91)</td>
<td>40.57 (43.59)</td>
</tr>
</tbody>
</table>
main effect of group, $F(1, 20) = .003, p = .96, \eta_p^2 = .00, f = .00, 1-\beta_o = .05$, were not significant.

**Speech time variables.** Values for total error to cut-off time and cut-off to repair time were calculated for all participants who made repairs to their speech in all conditions ($n = 18$). Levene’s test of homogeneity of variance was significant for both error to cut-off time and cut-off to repair time, $p < .05$ and not significant for speech rate total, $p > .05$. Spread vs. level plots for both error to cut-off time and cut-off to repair time were created to determine the appropriate transformation. Based on the power values, slope values and Tukey’s power ladder (Tukey, 1977), the most appropriate transformation was a reciprocal transformation (Howell, 2010). Levene’s test of homogeneity of variance was not significant for the transformed variables, $p > .05$. To control for Type I error with multiple t-tests, $p(\alpha)$ was set to 0.01 for the following analyses.

Independent samples t-tests on the transformed variables revealed an non-significant difference between the groups for error to cut-off time, $t(16) = -0.44, p = .66, r^2 = .01, d’ = -.22, 1-\beta_o = .07$, a significant difference between the groups for cut-off to repair time, $t(16) = -2.952, p = .01, r^2 = .35, d’ = -1.48, 1-\beta_o = 1.00$, and a significant difference for speech rate total, $t(20) = -3.514, p = .00, r^2 = .436, d’ = -1.57, 1-\beta_o = 1.00$.

One-way ANOVAs covarying out speech-rate were completed on the transformed cut-off to repair time to determine if rate of speech is important in determining the group differences in these variables. Group differences in cut-off to repair time were no longer significant, $F(1, 15) = 1.637, p = .22$; however, power was low, $1-\beta_o = .22$. In the non-transformed scores, children in the younger group had significantly longer times between speech cut-off and the initiation of the speech repairs than children in the older group.
Children in the younger group had significantly slower speech rate than the children in the older group (see Table 5 for descriptive statistics of raw error to cut-off time, cut-off to repair time and speech rate total).

**Working memory tasks.** As the working memory tasks have different scales, scores from the three working memory tasks were transformed into z-scores based on the total sample’s mean and standard deviation to create a consistent scale for comparison. A 2 (Age Group) X 3 (Working Memory Tasks) repeated measures ANOVA was completed with the z-scores (based on total sample mean) to determine the main effect of working memory task. Box’s M and Mauchly’s test of sphericity were not significant, $p > .05$. No significant within-group differences on the working memory tasks was found, $F(2, 40) = .011, p = .99, \eta^2_p = .00, 1-\beta_o = .05$. Children in each group showed consistent patterns of performance regardless of the type of working memory task and therefore, it was appropriate to create a composite score for the next analysis. For group and total sample raw scores on the working memory tasks, see Table 6.

A composite score for each participant was created using a sum of the z-scores for each of the three working memory tasks (based on the total sample mean). The difference between the working memory score composite z-scores was determined via an independent groups t-test. Levene’s test of equality of variances was not significant, $p > .05$. The between group differences were significant, $t(20) = -3.369, p = .00, r^2 = .36, d' = 1.51, 1-\beta_o = .95$. Children in the younger group performed significantly lower ($M_{z1} = -0.77, SE_{z1} = 0.65$) on tasks of working memory compared to the children in the older group ($M_{z2} = 0.44, SE_{z2} = 0.90$).
### Table 5. Time characteristics of language samples.

<table>
<thead>
<tr>
<th>Language characteristics</th>
<th>Network Condition</th>
<th>Children aged 6-8</th>
<th>Children aged 10-12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$ (SE)</td>
<td>$M$ (SE)</td>
</tr>
<tr>
<td>Speech rate (words/second)</td>
<td>Networks 1-5</td>
<td>9.79 (2.66)</td>
<td>12.83 (1.69)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>7.93 (3.58)</td>
<td>12.93 (1.99)</td>
</tr>
<tr>
<td></td>
<td>Networks 11 - 15</td>
<td>8.16 (3.82)</td>
<td>13.78 (6.51)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>7.89 (4.06)</td>
<td>11.47 (2.13)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8.44 (3.40)</td>
<td>12.75 (2.37)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n = 6$</td>
<td>$n = 12$</td>
</tr>
<tr>
<td>Error to cut-off time (ms)</td>
<td>Networks 1-5</td>
<td>639.31 (343.46)</td>
<td>770.05 (260.98)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>982.96 (566.25)</td>
<td>652.88 (194.16)</td>
</tr>
<tr>
<td></td>
<td>Networks 11 - 15</td>
<td>1289.22 (1649.13)</td>
<td>814.45 (409.82)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>788.58 (411.33)</td>
<td>633.10 (174.15)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>925.02 (490.29)</td>
<td>717.62 (149.17)</td>
</tr>
<tr>
<td>Cut-off to repair time (ms)</td>
<td>Networks 1-5</td>
<td>577.17 (424.46)</td>
<td>532.76 (201.43)</td>
</tr>
<tr>
<td></td>
<td>Networks 6-10</td>
<td>935.68 (296.75)</td>
<td>451.32 (182.51)</td>
</tr>
<tr>
<td></td>
<td>Networks 11 - 15</td>
<td>1395.24 (1614.10)</td>
<td>593.44 (321.57)</td>
</tr>
<tr>
<td></td>
<td>Networks 16-20</td>
<td>582.58 (261.13)</td>
<td>520.31 (203.58)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>872.67 (402.68)</td>
<td>524.46 (123.04)</td>
</tr>
</tbody>
</table>
**Working memory tasks and speech characteristics.** For the entire sample, the total sample working memory raw score total was regressed onto the raw total scores for errors, repairs per errors and pauses for the entire sample. Forty-three point three percent of the variability in the total errors from Networks 1-20 was explained by the working memory composite in a linear regression, \( F(1, 20) = 15.260, p = .00, R^2 = .43, f = .87, 1-\beta_o = .97 \). Forty-one percent of the variability in total repairs per errors score from Networks 1-20 was explained by the working memory composite in a linear regression, \( F(1, 20) = 13.01, p = .00, R^2 = .39, f = .80, 1-\beta_o = .95 \) Thirty-one point nine percent of the variability in total pauses score from Networks 1-20 was explained by the working memory composite in a linear regression, \( F(1, 20) = 9.368, p = .01, R^2 = .32, f = .69, 1-\beta_o = .86 \).
Table 6. Raw scores on working memory tasks by group.

<table>
<thead>
<tr>
<th>Working Memory Task</th>
<th>Children aged 6-8 (n = 8)</th>
<th>Children aged 10-12 (n = 14)</th>
<th>Total Sample (n = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WISC-IV Digit Span Forward (/16)</strong></td>
<td>7.88 (0.99)</td>
<td>9.29 (2.05)</td>
<td>8.77 (1.85)</td>
</tr>
<tr>
<td><strong>WISC-IV Digit Span Backward (/16)</strong></td>
<td>5.63 (1.60)</td>
<td>7.64 (2.41)</td>
<td>6.91 (2.33)</td>
</tr>
<tr>
<td><strong>WISC-IV Digit Span Total (/32)</strong></td>
<td>13.50 (1.93)</td>
<td>16.93 (3.38)</td>
<td>15.68 (3.43)</td>
</tr>
<tr>
<td><strong>CLPT Total Word Recall (/42)</strong></td>
<td>21.63 (5.13)</td>
<td>27.21 (5.29)</td>
<td>25.18 (5.80)</td>
</tr>
<tr>
<td><strong>CLPT Total Correct Sentence Response (/42)</strong></td>
<td>41.38 (0.74)</td>
<td>41.43 (0.76)</td>
<td>41.41 (0.73)</td>
</tr>
<tr>
<td><strong>Spatial Span (Maximum 16)</strong> **</td>
<td>6.50 (1.20)</td>
<td>8.50 (1.45)</td>
<td>7.77 (1.66)</td>
</tr>
<tr>
<td>– Network and Spatial Span Task</td>
<td>2.00 (1.41)</td>
<td>2.43 (1.50)</td>
<td>2.27 (1.45)</td>
</tr>
<tr>
<td><strong>Spatial Span Score (/5)</strong></td>
<td>2.00 (1.51)</td>
<td>2.14 (1.66)</td>
<td>2.09 (1.57)</td>
</tr>
<tr>
<td>– Network, White Noise and Spatial Span Task</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Group difference significant at p < .05.

** Group difference significant at p < .01.

Note: No other comparisons completed due to increased likelihood of Type I Error.
Chapter 4
Discussion

This chapter will discuss the theoretical and practical implications of the results discussed in chapter 3. The first will provide a number of important limitations of the current study that are key to understand prior to the discussion of the results. The second section will examine the main and supplemental findings in reference to the current state of the literature on speech production and working memory. Further, this section will provide a discussion of the findings with respect to the author’s proposed more detailed model of working memory’s role in speech production. The third section will evaluate the clinical implications of the current study’s findings.

Limitations. The current study’s findings must be considered in light of a number of limitations. The current study used a sample of healthy children recruited from an individual city. While common practice in research, such convenience samples may lead to a bias within the sample (Madsen, 2008). Recruitment strategies such as the use of internet parenting sites may lead to greater recruitment of higher-income families who are able to afford a computer as well as families with literate parents. In the current study, many parents indicated during the consent process that they had attended university for a bachelor’s or a master’s degree. While approximately 60% of adults aged 25 to 64 of the Canadian population have obtained secondary school degrees, including technical training degrees (Statistics Canada, 2006), the current study’s sample may have contained a higher parental education level than the average family in the population. A potential method by which to address the potential cognitive skew that may have been present in the data would have been to add assessment of general cognitive functioning or general
language abilities. In the current study, the highly restrictive inclusion criteria were adequate to ensure the children had the basic minimum abilities to complete the study. However, the sample may also reflect a higher functioning group and intellectual testing could assist in that determination. Further, while testing appointments were arranged at times convenient for the families, the requirement of dropping off and picking up the child at the university may have deterred some parents from volunteering due to the time constraints. Again, while these constraints on the sample are common in child research, any constraints on a sample require caution when applying the information to the population.

A second and major limitation of the current study was power. As Jacob Cohen (1992) suggested, in his seminal paper, a minimum of $1-\beta_o = .80$ is the criteria of an acceptable level of power; however, he also noted that this value was somewhat arbitrary. This reflects an 80% chance of finding a significant result when one exists (Bernstein, 2008). The power levels in this study were variable, ranging from .05 to 1.00. The analyses of the network task main effect were generally sufficient. However, the analyses of the group main effect and the interaction effects did not have sufficient power. The effect sizes of these effects were, on average, moderate as per Cohen’s (1992) guidelines. The primary explanation for the power deficiency of these effects is the small sample size.

A concern with reduced power is the fact that when an effect is deemed statistically insignificant ($p < .05$), there is still a “50% chance that a real effect has not been detected” (Cleophas & Zwinderman, 2012, p. 79). This can be a particular concern as non-significant results are not typically published, potentially creating bias and
incorrect models in the literature. Analyses using small samples may not detect small to moderate effects and as such, may not provide information about the level of the true effect of a variable (Cleophas & Zwinderman, 2012). Further, researchers may not pursue avenues of research that do not produce significant results and as such, true effects may be ignored.

Another concern with a small, low-powered study is that of the winner’s curse, the inflation of effect sizes (Button et al., 2013). The winner’s curse occurs due to the bias created by the fact that small, low-powered studies are only able to detect large effect sizes with reliability and thus, can create a literature citing an over-estimated effect size for a particular effect (Button et al., 2013; Forstmeier & Schielzeth, 2011). Button et al. (2013) suggest that this leads to a lack of reproducibility in future studies as the true effect may be lower. This should be taken into account if researchers use the effect sizes from this study as the true effects may be smaller. As such, a conservative approach to estimating sample sizes is recommended.

The third limitation of this study is the small sample size for reasons beyond reduced power. While the basic sample size assumptions were met within the statistical analyses, small samples inherently have greater sampling error. Sampling error reflects the precision of the estimate of the population from a particular sample (Howell, 2010). With a small sample such as in the current study, generalizations back to the population may be limited. This can be especially problematic in the context of high variability in a construct as with speech production. While techniques such as transformations are legitimate methods by which to address heterogeneity of covariance and variance, these techniques make it more challenging to directly interpret the statistical outcomes in a
practical manner (Howell, 2010). Thus, caution should be used when applying the results of this study to other normative samples.

The fourth limitation of the current study was the inherent language component of all verbal working memory tasks and potentially, the non-verbal working memory task. Verbal working memory tasks require the use of speech production for the recall phase of the task and this creates an inevitable overlap between constructs of speech production and verbal working memory (Acheson & MacDonald, 2009). While theoretically meaningful to use verbal working memory tasks as this is the type of working memory believed to be important in language, in order to clarify models, it may be useful to add non-language based working memory tasks to separate out some of the components. In the current study, the use of a coloured board in the spatial span task may have allowed children to use a verbal strategy such as listing the colours in their head, which may have reduced the working memory load of maintaining the spatial locations in mind.

Consistent with the theories of a domain general working memory component, a possible solution to resolve the problem is the use of a working memory tool that is less likely to elicit activation of brain processes associated with language production. Nystrom et al. (2000) noted that participants reported using a verbal strategy during \( n \)-back tasks using letters and abstract shapes but not location. This occurred even though participants were required to complete an articulatory repression task during the period between the stimulus and recall (Nystrom et al., 2000). A tool that may provide a partial solution to the problem of verbal strategy approaches is the Corsi Block-Tapping task. The Corsi Block-Tapping task requires a participant to replicate the examiner’s pattern of taps on nine identical spatially separated blocks in the forward direction and then in
reverse (Kessels, van den Berg, Ruis & Brands, 2008). This type of task is well established as a measure of visuo-spatial working memory and has been differentiated from verbal working memory tasks in multiple studies (for a review, see Kessels, van den Berg, Ruis & Brands, 2008). However, a concern about this task is the lack of an information manipulation component that is an aspect of the verbal working memory tasks used in the current study. The Corsi Block Tapping task and variants have identical forward and backward spans, which limits the information manipulation requirement (Kessels, van den Berg, Ruis & Brands, 2008; Wilde & Strauss, 2002; Wilde, Strauss & Tulsky, 2004). This could create a further methodological problem as the verbal working memory tasks chosen for the current study were chosen as they specifically require the executive attention needed in speech monitoring. However, using the Corsi Block Tapping task in addition to the other working memory measures in this study may be a useful method by which to differentiate the language component from the other aspects of working memory.

A fifth limitation of this study is the focus on an English-only analysis. While the focus of the current study was to establish a relationship between working memory and speech monitoring in children, evidence suggests that future research should integrate multi-lingual children to get a richer picture of the relationship. Bilingual children have more proficient working memory abilities than monolingual children, especially under higher-demand tasks (Morales, Calvo & Bialystok, 2013). Hernández, Costa and Humphreys (2012) found that bilingual children are less distracted by task-irrelevant information held in working memory while speaking than monolingual children. In this study, four children spoke two languages and one child spoke three. With a larger pool
of multilingual children with whom to compare to monolingual children, a comparison in speech monitoring abilities would potentially provide further clarification with respect to the relationship between working memory development and speech monitoring. A particularly interesting future avenue of research could evaluate speech monitoring across languages in multi-lingual children as differences may exist as part of language proficiency but also potentially of language type.

**Discussion of results.** This study examined the relationship between speech monitoring and constraints on working memory, developmental and experimental, in a child sample. With the above caveats in mind, one finding of the study was the significant relationships between the degree of working memory load of the network task condition and the following speech production phenomena: (1) total overt errors; (2) total pauses; and (3) total repetitions. A related finding was the significant relationship of performance on working memory tests and the following speech production phenomena: (1) total errors and (2) total pauses. Further, an interesting finding is the lack of significance of the relationship between degree of working memory load of the network task condition with (1) total repairs per errors and (2) total number of editing terms used. This finding is particularly interesting in the context of a significant relationship between the working memory composite and total repairs per errors. Group status was not found to be significantly related to the speech phenomena in this sample with the exception of the speech time phenomena; however, different non-significant trends were notable for the two age groups. These non-significant trends will be discussed throughout this chapter.
Total errors, silent pauses and repetitions. Higher cognitive load of the most complex network conditions and lower performance on working memory tasks were both associated with greater errors in speech. In both age groups, as predicted, children made more speech errors in the final condition (white noise and spatial span inference) than in the network only condition (control). There was no statistical difference between the white noise only condition and the final condition with respect to number of errors. This was contrary to the author’s prediction that greater complexity of task would elicit a greater number of errors. However, due to the low power of said contrast, with a larger sample, one may find a difference between the two levels. As such, a cautious approach to interpretation is simply that the most complex task elicited more speech errors than the control task.

While the group effect and interaction effect for speech errors were not significant, there were trends in the data that are theoretically of interest for future studies. For the younger age group, the introduction of the spatial span interference task in the second condition was sufficient to cause an increase (approximately 35.0% increase) in speech errors. However, this is with the caveat that the spatial span task may have been mediated by a verbal strategy. This said, if the younger children utilized a verbal strategy, the increase in errors may reflect less cognitive control than necessary to allocate attentional resources to long-term memory retrieval (i.e. working memory) required for the language task and the verbal strategy, especially in the context of reduced lexical access efficiency in younger children (Gathercole & Alloway, 2008; Novick et al., 2005; Sauzéon et al., 2004). Younger children may have difficulty holding the information in the phonological loop while the language monitors utilize the episodic
buffer for the prolonged access to the lexical-semantic system. The finding of lower levels of working memory associated with higher speech errors is consistent with the adult findings demonstrating that higher working memory spans are related to more proficient access to the lexical-semantic system and lower rates of speech errors (Daneman, 1991; Rosen & Engle, 1997).

The non-significant trend in the older children is not consistent with the findings in adults that suggest auditory noise masking does not affect error rates in speech (Postma & Kolk, 1992a; 1992b; Postma & Noordanus, 1996). A possible explanation for the detriment to speech monitoring is a lack of consistent tactile and proprioceptive feedback during speech production as seen in adults (Guenther, 2006). Until the age of 14-16, children are much more variable in their articulation patterns (i.e., movements of oral structures in relation to a particular sound) (Smith, Goffman, Zelaznik, Ying, & McGillem, 1995). With more complex utterances and greater processing demands, greater variability in this feedback is seen until mid-adolescence (Sadagopan & Smith, 2008; Smith, 2006). When white noise is presented and the auditory feedback is disrupted during speech production, children age 10-12 may not be able to efficiently use the variable tactile and proprioceptive feedback in order to detect errors. Thus, white noise may effectively blocks the internal speech monitor and one of its input monitors, a somatosensory monitor. However, due to the small sample size and non-significance of these trends, it is key not to assume a similar finding would occur with a larger sample. Thus, it is key to re-investigate these findings with a large normative base and the methodological changes noted in the limitations.
Higher cognitive load of the more complex tasks elicited less repetitions than the less complex tasks. If as according to Levelt’s model (1983; 1989; 1999), repetitions reflect repairs to covert errors via the internal monitoring loop, the pattern of fewer repetitions with greater load on the working memory system indicates a reduction in ability to monitor via the internal system, resulting in externalized errors. Further, repetitions may specifically reflect an attempt to restart a temporally degraded sentence plan (Kolk, 1995; Kolk & Van Grunsven, 1985) and the load on working memory may prevent subvocal rehearsal, which results in less successful restarts. While children as young as 4-5 years of age have demonstrated rehearsal, they tend to be less likely to rehearse and be less proficient with new information (Gathercole & Adams, 1994). Strategic subvocal rehearsal emerges around age 7 (Gathercole & Hitch, 1993; Gathercole & Pickering, 2000). However, extensive maturation of the speed and proficiency of covert rehearsal occurs between the ages of 6-12 (Kail & Ferrer, 2007).

In the current study, non-significant age group trends in production of repetitions occurred across the network task conditions. Younger children demonstrated a non-significant pattern of fewer repetitions after introduction of spatial span and another reduction after introduction of white noise. Older children demonstrated a non-significant pattern of consistent repetitions across the conditions, which is concordant with the pattern found by Oomen and Postma (2002) for adults. This non-significant trend may indicate that the level of subvocal rehearsal within the phonological loop is developed enough by age 10 to facilitate sentence plan restarts (i.e., effective use of repetition). Further study into this phenomenon should be designed clarify whether these non-significant trends reflect true differences or merely random error.
The number of non-filled (silent) pauses increased as network task difficulty increased in this sample of children. As the relationship between working memory load and verbal fluency has been previously established, the finding of increased pauses with increasing task complexity was expected (Grosjean & Deschamps, 1975; McDaniel, McKee, & Garrett, 2010). McDaniel et al. (2010) compared groups of 3-5 year olds, 6-8 year olds, and adults in an elicited sentence production task and more pauses for all groups occurred with increasing speech complexity (McDaniel et al., 2010). In the current study, non-significant age trends indicated that younger children had an increase in silent pauses after the introduction of the spatial span task and another during the last task. Older children had a non-significant increase in silent pauses when white noise was introduced, a finding consistent with the adult literature (Brookshire, 1969; Jameson et al., 2010). As the level of background noise increases, adult speakers increase the number of silent pauses (Lu & Cooke, 2008). These pauses may serve to assist the listener with comprehension of speech (Jameson et al., 2010; Lu & Cooke, 2008). However, they may also serve to provide a speaker with greater time to enact a speech plan or complete a repair (Levelt, 1983; 1989). The non-significant age group patterns of silent pauses is potentially theoretically interesting as with conditioning, silent pauses in speech can be reduced (Howell & Sackin, 2001). It may be that through interactions with other people, children become conditioned over time to reduce the use of silent pauses as this may lead to another speaker taking over the conversation (Howell & Sackin, 2001). Again, further research into these non-significant age-group trends should be designed to clarify if these differences reflect true differences or random error.

There are a number of other possibilities to explain the differences in the task complexity effect. One is the idea of goal neglect. In children with lower levels of working memory, maintenance of task goals and execution of demands of the task is less likely, even when the children can recall the task instructions (Marcovitch, Boseovski,
Knap & Kane, 2010). In terms of the current study’s results, while the children with lower levels of working memory knew that they needed to monitor their errors when speaking, it is possible that the ability to enact that goal and maintain it over the period during which they were speaking was beyond their working memory capabilities. Those children with higher levels of working memory potentially were able to keep the goal of both internal and external speech error detection in mind while speaking (Marcovitch et al., 2010).

The impact of working memory load on detecting errors in the second condition is a particularly interesting result considering the similar low performance (< 50% correct in both) of the groups on the distractor spatial span task. The low performance on the dual task is consistent with the findings in adults (Lackner & Tuller, 1979; Oomen & Postma, 2002). Best, Miller and Naglieri (2011) found no relationship between speed and accuracy on a coding task for children aged 5 and 7, whereas for children age 8-12, the speed-accuracy trade-off was much stronger. The age-differences found by Best et al. (2011) may reflect age-related differences in working memory and attention. In the current study, children with lower working memory abilities, the low score on the distractor task may reflect the inability to maintain the information in working memory (Best et al., 2011). Whereas for the children with higher working memory abilities, the need to quickly detect errors, make corrections, and keep up with the task may elicit a metacognitive strategy of reducing accuracy on the secondary task to preserve speed (and fluency) as appropriate for the network task (Best et al., 2011).

Another possible explanation for the network complexity effect is that working memory is an “emergent product of various components of cognitive and behavioural
systems organizing themselves over time within a specific stimulus and task context” (Simmering & Perone, 2013, p. 11). For speech monitoring, working memory capacity may partially reflect the ability for each language monitor (e.g., lexicality monitor) to effectively use the episodic buffer to combine information from long-term memory (e.g., language rules, social context, etc.) with information in the phonological store while comparing that information to intended and/or articulated speech. It may also reflect the constraints of attention capacity for all of the language monitors’ activities, which is consistent with the time-based resource-sharing model of working memory proposed by Barrouillet et al. (2004) or other more attention capacity focused models of working memory. The attention load of the various language monitors (e.g. syntax) during speech production impairs working memory. The processes needed to compare the intended plan to the pre-articulated or articulated message do not allow for refreshing of the material (Barrouillet et al., 2004). As such, differences in efficient processing of the specific language monitors likely affect how well children detect both internal and external speech errors. Children may simply need more time and effort to process information, which reduces their attention capacity and working memory (Nettelbeck & Burns, 2010).

A further constraint on processing is that young children have holistic lexical representations, which as language skills develop, change to segmented/phonemic representations of lexical items (Metsala, 1999). This holistic representation, or the rime (word family), is a superordinate unit acquired earlier in development than the ability to segment these units into phonemes (Brooks & MacWhinney, 2000; Coch, Grossi, Coffey-Corina, Holcomb, & Neville, 2002). Children age 7-8 are slower at monitoring phonemes within consonant clusters than children age 10-12, which is thought to be due
to the differences in segmenting ability (Sasisekaran & Weber-Fox, 2012). Further, children age 7-8 are much slower at rhyme and phoneme monitoring than children age 10-12 (Sasisekaran & Weber-Fox, 2012). Many of the language monitors require effective and efficient segmentation of larger units into smaller units (e.g., phonemes, morphemes) to compare on-line speech with the intended message (Levelt, 1983; 1989). Therefore, segmentation ability likely affects the efficiency of language monitoring. As both storage capacity and general processing speed are strong predictors of working memory and attention (Magimairaj, Montgomery, Marinellie & McCarthy, 2009), the lower efficiency in the ability to segment language is likely a constraint as it leads to slower language processing speed and thus, longer time periods of storage. As a result, children with slower processing are less likely to detect internal errors in speech before externalization, especially with working memory and attention constraints. For the children with greater skill with segmentation of language, the load on the working memory and attention system is reduced when monitoring language and thus, there is greater capacity to detect errors.

A further potential constraint of monitoring of one’s own internal and external speech errors is that the process is completed while the person is speaking (i.e., an on-line process with dual task component). Sentences are essentially long complex sequences (Palmer & Pfordresher, 2003). Long complex sequences require planning in terms of content, temporal aspects (e.g., according to the language’s syntax), situational context and when to produce it, in addition to the phonemic and lexical aspects of the sentence (Levelt, 1983; 1989; Palmer & Pfordresher, 2003). Younger children pause twice as often, make more frequent repetitions and require more time when describing a picture than older children (Pavão Martins et al., 2007). This suggests that even without a secondary task, providing a description of a picture can be challenging for younger
children, which may potentially reflect the reduced capacity of their working memory and attention.

Based on the findings of the current study and with considerations of the limitations of the study, a tentative proposal of a more detailed speech production model is provided that combines the updated Baddeley and Hitch (1974) model and Levelt’s (1989) model with additional monitors as per the research literature. Figure 7 illustrates this proposed network connectivity model. With this proposed model, the episodic buffer acts as a multimodal temporary store via which the language monitors are able to access necessary information from long-term memory, the phonological store and potentially from other monitors in order to complete efficient monitoring and repairing of language. This model expands on the description of the working memory and attention component of speech monitoring originally proposed by Levelt (1989).

**Repairs per errors and editing terms.** The lack of a significant relationship between working memory load on the network task and the two speech phenomena, repairs per errors and editing terms (an indicator of covert repairs) may have been due to the low power due to small group sizes. The importance of sample size is reflected in the finding that repairs per errors were significantly related to the working memory composite, which had increased power for the regression analysis due to the use of the entire sample.

The lack of a relationship between repairs per errors and working memory constraint is in contrast to the adult literature examining constraints of attention and working memory on speech repairs. In adults, Postma, Kolk and Povel (1990) found that speakers tended to reduce rate of speech, make less errors and more repairs as well as
show higher rates of dysfluencies (e.g., pauses) under time constraint and high required-accuracy level manipulation. Postma et al. (1990) theorized that the focus of greater attention increases towards speech programming under time constraint and in situations with an emphasis on accuracy. The increased attention to speech programming leads to greater detection and correction of both internal and external errors via covert and overt repairs (Postma et al., 1990). The non-significant trend in the current study suggests that older children repaired at a consistent level regardless of network task condition; whereas, the younger children reduced speech repairs with the introduction of spatial span and reduced again with the introduction of white noise. The use of editing terms increased during the last condition which may indicate that increasing the working memory constraints does not impact the ability to repair externally but when working memory constraints are combined with white noise, children may use editing terms (e.g., um, ah, I mean) to fill the pauses created by the greater processing demands. The non-significant editing term finding is consistent with findings in adults, where editing terms are used to indicate that an upcoming pause in speech is due to speech disruption so that the interlocutor does not begin to speak while the speaker is correcting the problem (Clark & Fox Tree, 2002). However, research with larger sample sizes is necessary to confirm these trends as true findings and not as a result of random error.
Figure 7. Model of the role of working memory in speech monitoring.
A consideration for the model is that the non-significant relationship between repairs and task complexity in children may in fact reflect a true finding rather than simply a power issue. Seyfeddinipur, Kita, and Indefrey (2008) found that adult speakers prefer fluency over accuracy and will interrupt speech at the point an appropriate repair has been created and not necessarily at the time of actual error. Children, with greater processing demands, may have developed a similar strategy or more likely, were not able to complete the simultaneous task of creating a repair while speaking. Seyfeddinipur et al. (2008) suggested that after making an error, a process of replanning occurs in which the proper repair is created. This process must occur, in part, simultaneously with the creation of the error as some repairs occur within zero milliseconds of making the error (Seyfeddinipur et al., 2008). The working memory and attention demands of the simultaneous repair may simply be too extensive for children. However, this will need to be confirmed in a study with a larger sample size.

Further, while working memory capacity predicts self-repairs in adults’ first language, Mojavezi and Ahmadian (2014) found that the direction of the particular relationships (e.g., positive or negative) depended on the type of repair made. Higher working memory capacity was strongly predicted more appropriateness repairs (e.g., repair of ambiguity in sentence). However, higher working memory capacity was strongly predicted lower speech error repairs. If different types of repairs have different relationships with working memory, this may explain the non-significant findings for speech repairs. For future studies, it may be more appropriate to separate the different types of repairs for the analysis. However, a limitation of the finding is that individuals with higher working memory capacity tend to make fewer errors and thus, they will not
have to make as many repairs. Therefore, a correction that accounts for number of errors made is necessary in order to determine the true direction of the relationship and this was not completed by Mojavezi and Ahmadian (2014). In the current study, due to power concerns, repairs were not analyzed by type. The use of the total score may have reduced the likelihood of finding a significant relationship if the relationships between working memory and different types of errors are in reverse directions.

**Speech time phenomena.** A third major finding is the significant relationship between group status and speech time phenomena, cut-off to repair time and total speech rate. Error to cut-off time was not related to group status in this study. Both groups performed slower (approximately 200-300 ms) in the control condition than has been reported for adults with respect to cut-off to repair time (Pillai, 2006). This may suggest that children have greater difficulty with the simultaneous re-planning that is postulated to occur in the creation of repairs and thus, require more time to complete re-planning once a cessation in speech occurs. Older children maintained a consistent cut-off to repair time, suggesting that the repair time is not constrained by working memory load. The cut-off to repair time for younger children increased when spatial span was introduced and again in the white noise condition. These findings are consistent with the view postulated in the perceptual loop theory that when resources are limited, the speed at which repairs are created is slowed (Levelt, 1983; 1989). However, evidence from the second language literature in advanced and intermediate adult speakers suggests that no differences exist between mono-task and dual-task conditions on cut-off to repair time (Declerck & Kormos, 2012). In children, the working memory constraints on information processing speed during complex tasks may be sufficient to reduce the speed
at which repairs are created. Contradictory to this view, when white noise and spatial span were combined, younger children’s cut-off to repair time reduced to the control condition speed. This finding is more consistent with that of Oomen and Postma (2002) in adults, where the dual-task condition elicited trends of faster error to cut-off time and cut-off to error time. With a high processing load, it may be that fewer resources go to re-planning. A future analysis could examine the nature of the attempts to repairs in the higher load conditions. If children put fewer resources into re-planning, attempts to repair would potentially include new errors such as awkward grammar. In the current study, due to power concerns, this analysis was not completed. As well, an examination of the role speech rate in these phenomena was ambiguous due to lack of power.

Speech was consistently slower in the younger age group. While rates were relatively consistent across conditions, during the network task and white noise condition, the older children spoke more words per minute. This may reflect that when the inner monitoring system is blocked, speech becomes less succinct due to reduced efficiency of grammatical monitoring. A future analysis should examine the grammatical complexity and succinctness of the sentences produced by the older children to test this hypothesis.

Clinical implications of findings. The relationship between working memory and attention constraints and speech errors has implications for how we measure linguistic abilities in children, particularly those with working memory challenges. Standardized language assessments currently used in neuropsychological assessment do not address speech monitoring. In speech-language assessments, these working memory and attention constraints on speech errors and repairs may not be detected as this is not an area traditionally assessed by those in speech-language pathology. As a result, these
types of speech monitoring difficulties may not be readily assessable at this time and potential opportunities for remediation lost. An adaptation of the network task for use by clinicians may provide a possible area for development in assessment. There is a substantial literature for the use of such tasks in adults (for review, see Postma, 2000) and with this study, support in children. Future studies in children with clinical conditions known to impact working memory and attention such as Attention Deficit Hyperactivity Disorder would provide further support for the task’s clinical use.

A further area of clinical study using the new model of speech production is the area of stuttering. Stuttering is a primary disfluency disorder. While many theories of stuttering focus on motor execution, recent psycholinguistic research has begun to examine the role of working memory in the creation of disfluencies in individuals who stutter (for review, see Bajaj, 2007). Children who stutter have been found to have lower working memory spans on non-word repetition tasks (with increasing syllables) than children who do not (Anderson & Wagovich, 2010). Further, sentence production has been found to be more cognitively effortful in individuals who stutter than individuals who do not (Bosshardt, 2006). One explanation for additional load during speech production is neuroimaging studies’ results that suggest over-monitoring of speech production in individuals who stutter (Arnstein, Lakey, Compton & Kleinow, 2011). A greater number of restarts or repetitions are the likely result of over-monitoring of the speech plan (Arnstein et al., 2011). Future research could use the network task conditions with appropriate modifications (e.g. reduced speed of ball) in children who stutter to examine the impact of working memory load and white noise on stuttering frequency.
Conclusions. While these preliminary results need to be replicated and expanded in a larger sample, a tentative conclusion to the results is that with the development of the episodic buffer and the phonological loop, children’s ability to monitor speech increases and speech errors decrease. Notable non-significant developmental trends in speech monitoring for the age groups of 6-8 and 10-12, correspond to the different developmental stages of working memory for these age groups (Gathercole et al., 2004). As such, the current speech monitoring models need to expand to include greater discussion of working memory and attention as well as the neuropsychological developmental literature in order to better account for typical and atypical speech monitoring development.
References


http://dx.doi.org/10.1016/S0022-5371(80)90312-6


Hulme, C., Maughan, S. & Brown, G. (1991). Memory for familiar and unfamiliar words:
Evidence for a long-term memory contribution to short-term memory span.  
*Journal of Memory and Language, 30*(6), 685-701.

http://dx.doi.org.ezproxy.library.uvic.ca/10.1016/0749-596X(91)90032-F


http://dx.doi.org/10.1093/geronb/58.5.P260.


doi:10.1006/jmla.1995.1014


Appendix A
Parent Consent Form
Research Consent Form
For Legal Guardian

Please read this form carefully. It tells you important information about a research study. Your child, ______________________, is invited to participate in a study entitled “Working Memory and Speech Production.”

This study is being conducted by Tanya L. Lentz, MA. Ms. Lentz is a graduate student in the department of Psychology at the University of Victoria and is conducting this research as part of the requirement for a PhD in clinical neuropsychology. This study is being conducted under the supervision of Dr. Kimberly Kerns.

Why is this research study being conducted?
The purpose of this research project is to further expand knowledge of the development of children’s ability to monitor their speech for errors and the impact of mental demands on that monitoring. This research is important as it may assist in the development of treatment approaches for children with difficulties in speech monitoring.

How are individuals selected for this research study?
Children with no known history of behavioural, cognitive and/or learning disabilities in two age groups (6-8) and (10-12) are being asked to participate in this study. All children must be fluent English speakers. Children will be excluded from the study if they have (a) chronic medical and/or psychological illness; (b) a history of trauma to the brain; (c) a history of hearing impairment; (d) a history of school non-attendance; and (e) English is not their primary language.

What will my child have to do if she/he is in the study?
While you wait in the waiting room, your child will complete a series of tasks in an adjacent room. Your child will be asked to do a series of tasks including: repeat numbers forward and backwards, verify the truth of sentences and recall the last word of the sentences, repeat a visual sequence, describe a dot’s progression through a pathway of pictures and sometimes these tasks will be completed while listening to white noise as a distraction.

What is involved for the study?
If you and your child agree to volunteer that your child will participate in this research, your participation will include one, approximately hour-long appointment at the University of Victoria, Cornett Building.

Your child’s performance on all tasks will be recorded manually and their verbal descriptions of the dot’s progression through a pathway of pictures will be recorded on a digital video disk (DVD). A written transcription and a digital copy will be made from the DVD.

What are the risks of this research study?
There are no known or anticipated risks to your child by participating in this research.

What are the benefits of this research study?
There is unlikely to be direct benefit to your child, other than the compensation provided, as a result of participation in this research. It is hoped that collecting and studying this information will lead to important knowledge of the relationship of speech production and monitoring in children. This may lead to improved services for children with difficulties in these areas. In addition, children often enjoy these sorts of studies and participating in a science experiment.

**Is my child’s participation in this study voluntary?**
Your child’s participation in this research must be completely voluntary. If you do decide to allow your child to participate, he/she may withdraw at any time without any consequences or any explanation, even if this is right at the beginning of the study. Your child will complete an assent form prior to starting the study and s/he will be asked if s/he wants to participate in the study. If s/he does withdraw from the study her/his data will destroyed and will not used in the conclusions of the study.

**Is there any compensation for this study?**
As a way to compensate your child for their time and any inconvenience related to his/her participation, he/she will be given $5.00, regardless of whether s/he completes the study.

**What will happen to the information obtained from the study?**
The DVD of your child’s information will be kept in a secure place and labeled only with a code number, not your child’s name. Digitalized copies of the tape will be protected by password. This consent form, which connects your child’s DVD with his or her code number will be kept in a separate secure place.

Data from this study will be disposed of five years following the completion of this study. This time period allows for the principal researcher to utilize all of the data for further papers.

The following people will have access to the data:
- Principal Investigator: Tanya L. Lentz, M.A.
- Principal Supervisor: Dr. Kimberly Kerns
- Research Assistant: TBA – access only under the supervision of the above researchers.

The results of this study will be presented to a committee of faculty members in written form for the purpose of completing the requirements of a PhD. Further, the results may be published in a journal or via the University of Victoria’s online research system or used to teach others. Your child’s name or other identifying information will not be used for these purposes without your specific permission.

**Who can I contact if I have further questions?**
Principal Investigator: Tanya L. Lentz, MA
Email: tientz@uvic.ca

Principal Supervisor: Kimberly Kerns, PhD
Email: kkerns@uvic.ca
Phone: 250-721-7553

In addition, you may verify the ethical approval of this study, or raise any concerns you might have, by contacting the Human Research Ethics Office at the University of Victoria
Email: ethics@uvic.ca
Phone: 250-472-4545

Your signature below indicates that you understand the above conditions (summarized below) of participation in this study and that you have had the opportunity to have your questions answered by the researchers.
I, _______________________________, consent to my child ______________________ participating, as requested in the research study “Working Memory and Speech Production.”

- I have read the information provided;
- Details of the procedures and any risks have been explained to my satisfaction.
- I am aware that I will be provided with a copy of the consent form and should retain this copy for future reference.
- I understand that:
  o My child may not directly benefit from taking part in this research.
  o My child is free to withdraw from the project at any time and is free to decline to answer particular questions.
  o While the information gained in this study will be published as explained, my child will not be identified, and individual information will remain confidential.
  o My child may ask that the recording/observation be stopped at any time, and she/she may withdraw at any time from the session or the research without disadvantage.
  o I agree to the audio/video recordings of my child’s information and participation.

Please indicate below the ways we can use the digital video made during this experiment. You can select some options and not others. Please note that if a video is played as an example (options 2-4), the child’s face in the video will be covered so that he/she is not recognizable. Please initial in the column that indicates your choice.

| 1. Viewing and analysis by Tanya Lentz, Dr. Kimberly Kerns and research assistants under their supervision conducting communication research. | YES, I consent | No, I do not consent |
| 2. Playing as an example for professional audiences (e.g. at a professional conference) |  |  |
| 3. Playing as an example for a class at UVIC. |  |  |
| 4. Playing as part of instructional material for communication research. |  |  |

Legal Guardian’s Name: _______________________________ Date: ______________________

Legal Guardian’s Signature: _______________________________

I certify that I have explained the study to the legal guardian of the participant and consider that she/he understands what is involved and freely consents to his/her child’s participation.

Researcher’s Name: _______________________________ Date: ______________________

Researcher’s Signature: _______________________________
Appendix B
Child Assent Form

My name is _______________________________.

Today I will be working with ___________________ at the University of Victoria. They are researchers in psychology and would like me to do some activities with them. For some of the activities, I will be asked to remember some information as best as I can such as numbers, words, or places. For other activities, I will watch a ball move from picture to picture on a path and I will be asked to tell them where the ball is on the path as it moves. Sometimes I will listen to a noise while saying where the ball is on the path and other times I will have to remember something and say where the ball is on the path. While I am doing the activities, there will be cameras that record what I say and what I do.

Doing these special activities will help researchers to understand more about what happens when kids speak and try to remember information. This may help in the future to figure out how to help kids who have difficulty with speaking and remembering.

I will do these activities because I decided I would like to participate in the study. I can change my mind at any time. If I decided at any time today that I no longer want to participate, I just have to tell ________________ and they will let me stop if I want to. Nothing I do here today will affect my grades in school or my health.

All of my data (scores, numbers, and any other information) collected from me today will remain confidential. That means that not even my parents or teachers will be able to know what my scores are. In fact, instead of using my name, they will use a secret code.

If I have any questions, my parents or I can contact Tanya Lentz at tlentz@uvic.ca or her supervisor, Dr. Kimberly Kerns at 250-721-7553.

Date: _____________________

Signature: _________________________________
## Appendix C
### Error coding scheme

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lexical Error</strong></td>
<td>1. Exchange errors: Meaningful items spoken in the wrong position. Example: red chair line on the chair for red line on the chair.</td>
</tr>
<tr>
<td><strong>Omission Error</strong></td>
<td>1. A unit of speech is not spoken and is missed from the intended target. Example: Blue line for blue curvy line.</td>
</tr>
<tr>
<td><strong>Phonological Error</strong></td>
<td>1. Switch of single or group of phonemes. Example: led rine for red line.</td>
</tr>
<tr>
<td></td>
<td>2. Upcoming or previous speech component spoken at wrong time.</td>
</tr>
<tr>
<td></td>
<td>a. Anticipation: Example: Bluervy curvy line for blue curvy line.</td>
</tr>
<tr>
<td></td>
<td>b. Perseveration: Example: Blue cluevy line for blue curvy line.</td>
</tr>
<tr>
<td><strong>Syntactic Error</strong></td>
<td>1. Grammatical errors that impede comprehension. Example: red to line back up now for red line back to the blue line.</td>
</tr>
</tbody>
</table>

*Error definitions adapted from Levelt (1983; 1989).*
## Appendix D

### Repair and related concepts coding scheme

<table>
<thead>
<tr>
<th>Repair Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lexical Repair</strong></td>
<td>1. Interruption of speech and correction of a lexical error.</td>
</tr>
<tr>
<td><strong>Omission Repair</strong></td>
<td>1. Interruption of speech and correction of a omission error.</td>
</tr>
<tr>
<td><strong>Phonological Repair</strong></td>
<td>1. Interruption of speech and correction of a phonological error.</td>
</tr>
<tr>
<td><strong>Syntax Repair</strong></td>
<td>1. Interruption of speech and correction of a syntax error.</td>
</tr>
<tr>
<td><strong>Repetitions</strong></td>
<td>1. Indicator of internally detected errors. Defined for the current study as repetition of correct word. No error expressed externally.</td>
</tr>
<tr>
<td><strong>Editing Terms</strong></td>
<td>1. Interjections in speech. May be used to indicate need for repair. Examples: That is, I mean, Uh, No, Sorry.</td>
</tr>
<tr>
<td><strong>Pauses</strong></td>
<td>1. Short pause – no speech for 60 seconds</td>
</tr>
<tr>
<td></td>
<td>2. Long pause – no speech for greater than 60 seconds.</td>
</tr>
</tbody>
</table>

*Repair definitions adapted from Levelt (1983; 1989).*