VERBAL WORKING MEMORY AND DISCOURSE COMPREHENSION IN OLDER ADULTS

by

Janine K. Fitzpatrick

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
June 2016

Dedication Page

To Carol-Anne, Alan, and Joey: I could not have done this without your love and support, from the very beginning. You have given me everything I could ever need, and I am so thankful. To Lewis: Your patience and understanding cannot be overstated. Thank you for keeping me grounded and focused. To Emily: We did it! I am so proud of what we've accomplished, and I'm happy to have come so far with you. To my friends, classmates, and family: Thank you for your encouragement over the past three years. I'm lucky to have such kind, thoughtful people in my life.

I am indebted to my supervisors for their support. To Dr. Janet Ingles: Thank you for the time and energy you have invested in me. Your guidance and encouragement have been invaluable. To Dr. Gail Eskes: Thank you for welcoming me into your laboratory, and providing me with the direction and resources necessary to complete this project. Thank you to Dr. Patricia Cleave and Dr. Ray Klein for serving on my committee.

Finally, to Kyle, the most determined person I know.

Table of Contents

List of Tables

List of Figures

Abstract

This study addresses the following question: How do different types of linguistic working memory relate to discourse comprehension among older adults? Community-dwelling participants (n = 34, mean age = 67.5 years) completed two assessments: first, a computerized visual working memory task adapted from Wright et al. (2007) that measures phonological, semantic, and syntactic subprocesses; and second, the *Discourse Comprehension Test* (DCT, Brookshire & Nicholas, 1993). We found correlations between phonological and syntactic working memory and total score of the DCT ($rs_{\tau} = .331 - .343$, $p < .01$); working memory's role in auditory rehearsal and sequencing may contribute to its relationship with discourse comprehension. There was no significant relationship between semantic working memory and DCT score; single word semantic processing as measured by this task may not be involved in understanding longer passages. Future research can extend this project to individuals with aphasia to understand these relationships in clinical populations.

List of Abbreviations Used

DCT	*Discourse Comprehension Test* (Brookshire & Nicholas, 1993)
RT	Reaction Time

Acknowledgments

This project has taken shape thanks to many people's assistance and support. Thank you to Dr. Heather Harris Wright for providing the stimuli for this project. I am very grateful to members of the Cognitive Health and Recovery Research Laboratory for their time and expertise: Jacob Kroeker and Joshua Salmon for developing and troubleshooting the stimulus presentation program; Anne Johnson for support with statistical analysis; Sarah Dolan, Stephanie Jones, and Elaine Skene for participant recruitment and screening; and Emily Bedford for support with test administration. Many thanks to Kathleen Cairns for her assistance with formatting. Thank you to the Nova Scotia Aphasia Association and the Chebucto Links group for your help with promoting this study among your members. Finally, I am grateful to all participants in this study; thank you for contributing your time and energy to my project. This research was conducted while JF was supported by a Scotia Scholars$^{\mathrm{OM}}$ Medical Research Award from the Nova Scotia Research Foundation (NSHRF).

Chapter 1: Introduction

1.1 Working Memory and Comprehension

Working memory involves the temporary storage and manipulation of information (Baddeley, 1995). We rely on working memory to calculate tips in a restaurant, or to remember directions while driving. A key feature of the working memory system is its limited capacity for information (e.g., Caplan & Waters, 1999; Daneman & Carpenter, 1980; Miller, 1956). This capacity is constrained by the availability of cognitive resources, and/or the efficiency with which these resources are allocated, both of which decline with age (see Abrams & Farrell, 2011, for review). Working memory is involved in different forms of complex thinking, including problem solving, reasoning, and language comprehension.

To understand language, we need to be able to process incoming information while keeping earlier information in mind. Temporary storage of information, a function of working memory, is therefore crucial to comprehension (Carpenter & Just, 1989). The relationship between working memory and language processing in healthy adults has been explored in depth (e.g., Miyake, Just, & Carpenter, 1994; Caplan & Waters, 1999). This past research focuses heavily on sentence-level processing. However, there is evidence that the cognitive systems for sentence processing may be separate from, although related to, those underlying discourse processing (Ulatowska, Allard, & Chapman, 1990). Discourse refers to a message containing multiple sentences that contribute to an overall meaning (Nicholas & Brookshire, 1995). Examples of discourse include a set of instructions for how to complete a task, a narrative piece, or a conversation. In this study, we investigate how working memory relates to discourse processing in older adults.

Language comprehension occurs via multiple modalities, including visual (reading) and auditory (listening). Penney (1989) reviewed evidence of separate processing streams for short-term memory of verbal information presented visually and aurally. Auditory presentation of verbal information led to better recall in short-term memory tasks compared to visual presentation. However, Penney's (1989) review focused solely on retention and recall of individual letters, numbers, and words, not discourse. Rubin and colleagues (2000) investigated discourse comprehension and perceived mental effort associated with reading and listening in college students. Participants read or listened to passages of oral-based (a corporate speech) and literate-based (a magazine article) discourse, and answered comprehension questions about the passages. Compared to listening, reading was associated with greater mental effort and better comprehension of both oral-based and literate-based material. The authors suggested orthographic decoding requires more cognitive processing than phonetic decoding (Rubin et al., 2000). Taken together, these studies suggest that listening may have an advantage over reading at the single word level, but the increased effort needed to read material improves comprehension of discourse.

How is working memory structured? Two main theories of working memory have been proposed: the first proposes that working memory is a general processor of information constrained by a finite amount of cognitive resources. That is, working memory is thought to be part of a domain-general cognitive system for different functions, including attention and executive control (Engle, 2002). The second theory argues that working memory can be broken into distinct subprocessors responsible for manipulating different kinds of information. For example, Baddeley and colleagues' (Baddeley & Hitch, 1974; Baddeley, 2003, 2012) model of working memory contained modality-specific storage and buffers for visuospatial and verbal information, with a central executive for attentional control.

The domain-general/domain-specific distinction can also be found in theories of working memory specific to verbal information. Just and Carpenter's (1992) capacity theory of comprehension described verbal working memory as a single capacity, for which deficits are caused by a generalized resource reduction. Different elements of language, including words, grammatical structure, and thematic roles, are thought to be stored in long-term memory; each has an activation threshold that allows it to be retrieved. There is no distinction between elements in the memory system. Instead, all are constrained by generalized activation resources, which vary among individuals. In contrast, Caplan and Waters (1999)□ proposed further segmentation for verbal information in the working memory system. They distinguished interpretive processing of linguistic information – extracting meaning from linguistic signals by processing their prosodic, syntactic and thematic features – from post-interpretive processing, which involves activities like long-term storage, reasoning, and planning. Their hypothesis included a separate-language-interpretation-resource (SLIR) responsible for interpretive processing that is not influenced by extrinsic loads on general verbal working memory. Regardless of how it is conceptualized, there is strong evidence that working memory is central to cognitive functioning and language processing (Mayer & Murray, 2012).

1.2 Working Memory in Aphasia

One way to study the structure of cognitive processes is to examine how they break down as a result of brain damage. In aphasia, a language disorder usually acquired as a result of left-hemisphere stroke (Papathanasiou & Coppens, 2013), working memory is affected by a general reduction of cognitive resources and/or reduced efficiency of allocating those resources. As a result, people with aphasia often have impaired working memory ability in conjunction with their deficits in language production and/or comprehension (Mayer & Murray, 2012; Potagas, Kasselimis, & Evdokimidis, 2011; Wright & Shisler, 2005).

A number of studies have investigated the relationship between working memory and language deficits in aphasia. Caspari and colleagues (1998) measured working memory in people with aphasia using listening and reading spans. In these tasks, participants were asked to remember words that appeared at the end of a series of sentences. Next, participants tried to identify these words by selecting a picture showing the target word from a set including foils of unrelated pictures. The authors found a strong positive correlation between listening span and reading comprehension of syntactically complex sentences (i.e., morphosyntax subtest of the *Reading Comprehension Battery for Aphasia*; LaPointe & Horner, 1979). Counterintuitively, the authors found reading comprehension was more strongly correlated with listening span than reading span. They suggested listening span can be used as an index of working memory capacity, and is predictive of sentence-level reading comprehension in people with aphasia.

Similarly, Sung and colleagues (2009) used a listening span task to divide participants with aphasia into high and low working memory groups. The low working memory group had worse performance on the complex subtests of the *Computerized Revised Token Test* (CRTT; McNeil et al., 2008), a test requiring participants to follow increasingly complex commands (e.g., "put the little black circle to the right of the white square"). This relationship was consistent for both listening and self-paced reading versions of the CRTT.

In Caspari et al. (1998) and Sung et al.'s (2009) studies, participants with smaller working memory capacities had more difficulty understanding complex sentences than participants with larger capacities. Group differences in understanding simpler sentences, thought to put fewer demands on the working memory system, were less pronounced. These studies demonstrate the effects of working memory on comprehension are most evident when task demands exceed the capacity of the working memory system. While Caspari et al. measured verbal working memory using visual and aural spans, Sung et al. only used aural spans. Visual span tasks for verbal information have not been well studied in aphasia.

People with aphasia can have difficulty processing different types of linguistic information (e.g., phonological, semantic, syntactic), which may lead to distinct aphasia syndromes (Papathanasiou & Coppens, 2013). If working memory capacity affects sentence comprehension in people with aphasia, it follows that this capacity could affect processing in other areas of language as well. Friedmann and Gvion (2003) found that people with a verbal working memory deficit and conduction aphasia, characterized by disproportionate impairment in speech repetition, had more difficulty understanding sentences that required phonological processing than those requiring semantic-syntactic processing. Based on these findings, the authors suggested the effect of a verbal working memory deficit on sentence comprehension may depend on the type of processing needed to understand the sentence. This study lends support to modular theories of verbal working memory (e.g., Caplan & Waters, 1999), which include specialized processors for linguistic information.

1.3 Measures of Working Memory

There are a number of ways to measure working memory. One measure commonly used in aphasia research is the complex span task (Daneman & Carpenter, 1980), which involves a serial recall task combined with a concurrent processing load (Conway et al., 2005). For example, participants may be required to remember a series of letters while also making grammaticality judgments about sentences. *N*-Back tasks also measure working memory. In these tasks, a series of items is presented one at a time; participants look for matches between the current item to one presented ‘n’ items ago in the sequence (e.g., 1-back, 2-back, etc.). *N*-Back tasks require that information be continuously updated to accommodate new items in the sequence (Jonides et al., 1997; Wright & Fergadiotis, 2012). Participants must also inhibit responses for matching items that do not correspond to the level being tested (Cansino et al., 2013). That is, if an item matches one presented immediately before it but the task is to match the one presented two items ago (i.e., 2-back level), the participant must refrain from responding. As such, *n*-back tasks use the storage, processing, and inhibition functions central to the working memory system (Mayer & Murray, 2012).

Performance on *n*-back tasks declines with age in healthy populations (Cansino et al., 2013). Daffner and colleagues (2011) used a 2-back letter-matching task to distinguish high and low performers in younger and older adults. They suggested poor performance in both age groups is due to reduced efficiency of the working memory system, leading to fewer resources allocated to the decision-making process needed to complete the task.

N-Back tasks have been used to study working memory in healthy (Cansino et al., 2013; Daffner et al., 2011) and clinical populations, including in those with aphasia (Christensen & Wright, 2010; Mayer & Murray, 2012), Parkinson's Disease (Miller, Price, Okun, Montijo, & Bowers, 2009), and traumatic brain injury (Perlstein et al., 2004). Mayer and Murray (2012) describe a number of advantages of using *n*-back tasks to measure working memory. First, *n*-back tasks require recognition vs. recall; participants must identify a target item as it is presented instead of retrieving it without external cues, as in a complex span task. Second, *n*-back tasks can be completed using response buttons to indicate matches, reducing the need for verbal output, which is particularly important when testing people with aphasia. Third, the *n*-back format can host a variety of stimulus types, including letters, numbers, words, and shapes; this allows researchers to selectively measure working memory for specific information without confounding demands on other systems. Fourth, *n*-back items are presented at equal time intervals, allowing reaction time to be collected and analyzed to detect subtle cognitive impairments (Crerar, 2004)□. Finally, *n*-back stimuli can be presented aurally and/or visually, allowing the task to measure processing in different modalities.

Wright and colleagues (2007) studied the relationship between auditory *n*-back performance and sentence comprehension in aphasia. Their *n*-back tasks measured three types of linguistic working memory (phonological, semantic, and syntactic) at 1-back and 2-back levels. Participants matched items based on phonological similarity, semantic categories, and sentence structure. Accuracy scores were calculated for each type of processing. Sentence comprehension was assessed using the *Subject-Relative, Object-Relative, Active, Passive Test of Syntactic Complexity* (SOAP; Love & Oster, 2002). This task required participants to point to one of two pictures that corresponded to a sentence presented aurally. Performance on the syntactic 2-back task and the non-canonical sentence forms of SOAP (i.e., sentences that did not follow subject-verb-object format) were positively related. Participants who performed worst on the syntactic working memory task had the most difficulty understanding syntactically complex sentences. The authors interpreted their results as being supportive of Friedmann and Gvion's (2003) findings and concluded that distinct working memory subsystems underlie different facets of language processing. Further, they concluded that verbal working memory deficits affect comprehension when both tasks require the same type of processing, which in this case was syntactic.

Because Wright et al. studied a clinical population, their sample size was limited. They tested 8 people with a variety of different types of aphasia and had no control group. An additional limitation was that the authors did not report the correlations between the semantic and phonological *n*-back performances and SOAP scores, which limited the degree to which their results can be interpreted. That is, they indicated that syntactic working memory scores at the 2-back level related to SOAP performance, but did not demonstrate that phonological and semantic scores did not relate to SOAP performance. Given that correlations were found between two syntactic processing tasks, the relationship may not be related to working memory specifically, but rather to a general syntactic processing skill used in both tasks. Finally, the SOAP task measured auditory comprehension of different syntactic structures in a controlled format. If different processing streams exist for aural and visual information (Penney, 1989), it is worth exploring these relationships in the visual modality. Sentence and discourse processing in more functional formats may place even greater demands on working memory (Nicholas & Brookshire, 1995).

1.4 Discourse Comprehension

Discourse processing places higher demands on cognitive systems compared to lexical or simple sentence-level processing. In the latter types of processing, linguistic units are presented in isolation and do not need to be held or manipulated in the working memory system in order to extract meaning across units. Understanding discourse involves integrating current material with what has been processed earlier, in order to create a cohesive representation of the message (Abrams & Farrell, 2011). Therefore, comprehension crucially requires the holding and accessing of information in working memory. Nicholas and Brookshire (1995) applied a resource allocation model of working memory to discourse comprehension, suggesting that performance involves a number of cognitive processes and would be limited by increasing demand for resources. They proposed that we use heuristic processes dependent on memory, including context and background knowledge, to extract meaning from discourse. In this way, the need for word-by-word analysis is reduced. Comprehension relies on general knowledge and relationships between different elements of the discourse (e.g. use of scripts and inferences).

Older adults may have difficulty understanding and retaining longer sentences and texts compared to younger adults (see Burke & Shafto, 2008, for review). Age differences become especially apparent in situations in which information is presented quickly, by increasing demands on the encoding efficiency of the working memory system (Hasher & Zacks, 1988). Age-related difficulties in discourse comprehension have been attributed to decline in cognitive processes, including working memory (Thornton & Light, 2006). There are several theories concerning age-related cognitive decline (see Abrams & Farrell, 2011, for review). Theorists have proposed decline is due to generalized cognitive slowing (e.g., Salthouse, 1996), or weakened inhibitory processes for irrelevant information (Hasher & Zacks, 1988).

Stine-Morrow and colleagues (2008) demonstrated working memory decline in older adults affects processing of surface-level (individual words and syntactic structure) and text-based (links of semantic meaning across several concepts) material in sentence-level reading. However, age effects were less pronounced in narrative and expository reading, presumably due to the facilitating effects of context and general knowledge. The authors concluded older adults depend on their existing knowledge during discourse comprehension more so than younger adults, instead of processing meanings of individual words and sentence structures. As a result, they may remember the gist of a text but not specific details. In this study, verbal working memory was conceptualized as a unitary construct, and was measured using a combination of span tasks.

1.5 Current Study

The purpose of the current study was to determine whether working memory for different types of linguistic information is related to discourse comprehension. We adapted Wright et. al's (2007) *n*-back tasks to assess linguistic working memory for phonological, semantic, and syntactic information. Discourse comprehension was assessed using the silent reading version of the *Discourse Comprehension Test* (Brookshire & Nicholas, 1993). We evaluated the relationship between working memory and discourse in older adults, rather than a clinical population, to allow for a larger sample size. Findings in older adults can then be applied to research in aphasia and other clinical populations. We expected that linguistic working memory would be positively related to discourse comprehension. Semantic processing involves extracting meaning and considering relationships between words, which could rely on general knowledge. Syntactic processing involves sequencing units to form a cohesive message. Therefore, we expected working memory for semantic and syntactic information would relate to comprehension of discourse in older adults. We expected phonological working memory would not relate to discourse comprehension.

Chapter 2: Method

2.1 Participant Characteristics

Thirty-four community-dwelling adults (28 females and 6 males) participated in this study. Participants were between 56 and 82 years old (M = 67.03, SD = 7.74), and had completed between 12 and 20 years of education (M = 15.74, SD = 2.29). The *Montreal Cognitive Assessment* (MoCA; Nasreddine, 2010) and the Word Reading subtest of the *Wide Range Achievement Test 4* (WRAT-4; Wilkinson & Robertson, 2006) were administered to ensure participants had sufficient cognitive (MoCA total score $\geq$ 24) and reading (WRAT-4 Word Reading score $\geq$ 45, corresponding to 6th grade equivalent) ability to complete the experimental tasks. MoCA scores ranged from 24 to 30 out of 30 (M = 27.68, SD = 1.71), and WRAT-4 Word Reading scores ranged from 47 to 68 out of 70 (M = 62.74, SD = 4.17). All participants reported themselves to be in good health, spoke English as their first language, had normal or corrected vision, and sufficient upper limb mobility to use a computer mouse. Participants had no self-reported history of neurological disorder, and were not currently taking medication that could influence cognitive performance (e.g., certain antidepressants). All participants provided informed consent in accordance with procedures approved by the Nova Scotia Health Authority Research Ethics Board. Participants were recruited via a laboratory database of individuals who had agreed to take part in research, online advertisements, and word of mouth. Participants were not reimbursed for taking part in this study.

2.2 Measures

Participants were tested during one session. First, the *n*-back working memory tasks were administered, followed by the *Discourse Comprehension Test*. Breaks were scheduled between tasks as needed.

2.2.1 *N*-Back. *N*-Back tasks were used to measure working memory for three types of linguistic information: phonological (*PhonoBack*), semantic (*SemBack*), and syntactic (*SynBack*). The stimuli for each task were developed by Wright et al. (2007) and are displayed in Appendix A. The tasks were run on a 2011 Apple Mac Mini desktop computer using custom software written in Python 2.5.2. In each task, participants viewed a series of successively presented words or sentences. They were instructed to decide whether the current item on the screen corresponded to the one presented a certain position behind it in the sequence. The ways in which items corresponded to one another differed by task type.

The *PhonoBack* stimuli included 25 consonant-vowel-consonant (CVC) words with five possible final sounds: *-at, -it, -in, -ill,* and *–ig.* Participants were instructed to decide whether the current word in the sequence rhymed with one presented previously (e.g., *fin* and *tin*). The *SemBack* stimuli consisted of 25 words belonging to one of five categories: fruit, tools, animals, furniture, and clothing. Participants determined whether the current word belonged to the same category as one before it (e.g., *apple* and *grape*, in types of fruit category). In the *SynBack* task, a subject-verb-object sentence was presented in either active (e.g., *The actor thanked the golfer*) or passive (e.g., *The golfer was thanked by the actor*) format. Ten nouns and ten verbs controlled for length, frequency of occurrence, and grammatical role were used in the *SynBack* sentences. Participants decided whether the current sentence described the same event as one presented previously, in either active or passive format (e.g., *The actor thanked the golfer* matched *The actor thanked the golfer* as well as *The golfer was thanked by the actor*).

PhonoBack and *SemBack* items were presented one at a time for 500 ms; items were separated by a 1500 ms interstimulus interval. *SynBack* items had a longer presentation time than *PhonoBack* and *SemBack* items because they consisted of sentences instead of single words, and took longer to read. *SynBack* items were presented for 1500 ms and were separated by a 1500 ms interstimulus interval. Participants used two buttons on a computer mouse to indicate a match or no-match for each item, based on whether it corresponded to an earlier item in the sequence according to the designated *n*-level. Participants could respond to a given item from the time it was presented until the end of the 1500 ms interval that followed. The stimulus presentation program recorded accuracy and reaction time (RT) for each item.

Each task was presented at 1-back and 2-back levels: in the 1-back level participants compared the current item to the one that came immediately before it in the sequence; in the 2-back level participants compared the current item to that which came two positions before it, with one item in between those that were being compared. A total score was computed for 1- and 2-back levels on each of the three working memory tasks, taking correct matches and non-matches for each trial into account.

The item presentation sequence used by Wright et al. (2007) was replicated in this study. The *PhonoBack*, *SemBack*, and *SynBack* tasks each contained two practice blocks and two experimental blocks for both 1-back and 2-back levels, totaling eight blocks for each information type. One-back blocks were presented first in each task, followed by 2-

back blocks. Items within each block were presented in the same order across participants.

The first two blocks in both 1-back and 2-back levels were practice trials, each containing 10 items with 2 targets (items which were a positive match to one presented earlier). The same practice block was presented twice in a row for each level. The 1-back level contained a third block of 32 experimental items with 10 targets, and a fourth block of 33 experimental items and 10 targets. The 2-back level contained a third block of 37 experimental items with 10 targets, and a fourth block of 39 experimental items with 10 targets.

The presentation order of the *PhonoBack*, *SemBack*, and *SynBack* tasks was counterbalanced across participants.

2.2.2 *Discourse Comprehension Test.* A paper-and-pencil version of the *Discourse Comprehension Test* (DCT; Brookshire & Nicholas, 1993) was administered to measure comprehension of narrative discourse in silent reading. The DCT was developed for neurologically impaired populations and was standardized on aphasia, right-brain damage, traumatic brain injury and no brain-damage groups. The stories were written between a fifth and sixth grade reading level. The test contained two sets of five short stories (mean number of words in each story = 205.6) with eight yes/no comprehension questions for each story. The comprehension questions were grouped into four subtests along two dimensions: salience (questions that concerned main ideas vs. details) and directness (questions that concerned directly stated vs. implied information). Main Idea questions concerned information that was central to the story and elaborated upon, while Detail questions focused on peripheral information that was only mentioned once. Stated questions concerned information that appeared directly in the story, in the same form in which it appeared in the question. Implied questions involved information that was not directly stated in the story, but must be inferred from other information (Nicholas & Brookshire, 1995). Each of the four subtests generated an accuracy score out of 20; the four subscores were summed to generate a participant's overall DCT score, out of 80.

Participants were not given a time limit to read the stories, but were instructed to read each story only once. After they finished reading, the story was removed and the comprehension questions were immediately provided.

The two story sets were counterbalanced across participants.

Chapter 3: Results

3.1 Statistics and Analysis

N-Back accuracy and reaction time (RT) data were each analyzed using repeated measures 3 x 2 analysis of variance (ANOVA), with information type (phonological, semantic, syntactic) and *n*-back level (1-back, 2-back) as within-subject factors. Accuracy is reported as percent correct across all trials within each condition, including both correct matches and non-matches. Accuracy was computed out of a possible 63 trials at the 1-back level and 72 trials at the 2-back level. RT analyses included data for correct trials only. Trials with RTs less than 200 ms or greater than 3 standard deviations above the participant's mean RT for a given condition were excluded from RT analysis (1.2 % of total trials). DCT data were analyzed using 2 x 2 repeated measures ANOVA with salience (questions about Main Ideas vs. Details) and directness (questions about Stated vs. Implied information) as within subject factors. The dependent variable was DCT total score, out of 80.

Correlation analyses were first computed between performance on the working memory and discourse comprehension measures and demographic variables (age, years of education, MoCA and WRAT-4 Reading scores). Next, correlation analyses were conducted to examine relationships between performance on each *n*-back task and the DCT. Kendall's Tau correlations were computed because homogeneity of variance could not be assumed. Finally, hierarchical regression analyses were conducted to determine whether the *n*-back tasks predicted DCT total score after parceling out the influence of demographic variables.

3.2 *N*-Back

Participant mean accuracy across *n*-back conditions is displayed in Figure 1 and Table 1. Accuracy was lower for 2-back tasks than 1-back; 1-back performance approached ceiling across all information types. *SynBack* accuracy was highest at both 1-back and 2-back levels, while *PhonoBack* and *SemBack* accuracy were similar to one another at both levels. The difference in accuracy between *SynBack* and the other tasks was greater at the 2-back level than 1-back. These observations were confirmed by statistical analysis which revealed significant main effects for *n*-back level, $F(1, 33) = 316.13$, $p < .01$, and information type, $F(2, 32) = 26.70$, $p < .01$. The interaction between information type and *n*-back level was significant, $F(2, 32) = 18.45$, $p < .01$. Analysis of the interaction revealed greater accuracy for *SynBack* compared only to *PhonoBack* at the 1-back level ($M_{\text{Difference}} = 3.22$, 95% CIs [0.56 – 5.88], $p < .02$). For 2-back trials, *SynBack* accuracy was greater than both *PhonoBack* and *SemBack* ($M_{\text{Difference}} = 10.25, 9.76$, 95% CIs [7.31 – 13.20], [6.73 – 12.80], respectively, $p < .001$). *PhonoBack* and *SemBack* accuracy were not significantly different from one another in 1-back or 2-back levels ($p > .16$). Overall accuracy was computed for each condition by combining 1-back and 2-back scores. Overall accuracy scores on all three *n*-back tasks were significantly correlated with one another ($rs_{\tau} = .415 – .464$, $p < .001$).

Table 1

Accuracy and RT for N-Back Tasks (N = 34)

	PhonoBack		*SemBack*		*SynBack*	
	1-back	2-back	1-back	2-back	1-back	2-back
M Accuracy	90.76	70.14	92.86	70.63	93.98	80.39
(*SD*)	(10.19)	(11.05)	(7.19)	(12.46)	(5.79)	(7.33)
M RT	973.90	1277.94	1111.57	1492.87	1483.48	1670.30
(*SD*)	(227.91)	(318.69)	(223.77)	(266.89)	(333.80)	(256.05)

Note. N-Back accuracy variables represent % correct trials for each task. RT is displayed in ms.

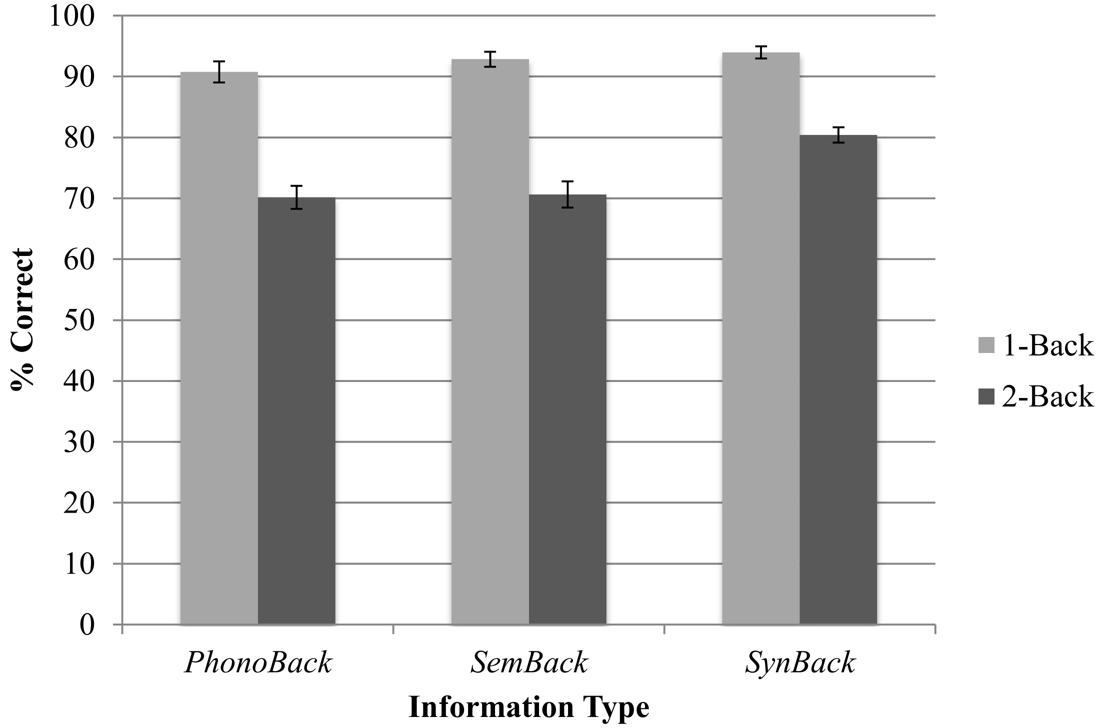


Figure 1. Mean accuracy for *n*-back tasks for each information type and level. Error bars represent standard errors of the mean.

Age was negatively correlated with accuracy across all *n*-back conditions (rs_τ = -.253 – -.45, $p < .04$) except the *SynBack* 1-back task ($p > .07$). A scatterplot demonstrating the relationship between age and *n*-back accuracy (collapsed across information type and n-back level) is displayed in Appendix B. MoCA score was correlated with accuracy on the *PhonoBack* 1-back, and *SynBack* 2-back conditions (rs_τ = .325 – .371, $p < .02$). Neither education nor WRAT-4 Reading score were significantly correlated with *n*-back accuracy for any condition (see Table 2).

These analyses were repeated using *d*-prime scores instead of accuracy (as in Cansino et al., 2013 and Hayes, 2011). *D*-prime scores provide an estimate of participants' discrimination level, taking both correct 'hits' and false alarms into account. The overall patterns of *d*-prime data were the same as the accuracy data reported above. A breakdown of participant accuracy, including correct 'hits', correct rejections, false alarms, and false rejections is displayed in Appendix C.

Table 2

Correlations for Demographic Variables and N-Back Accuracy (N = 34)

	PhonoBack		*SemBack*		*SynBack*	
Variable	1-back	2-back	1-back	2-back	1-back	2-back
Age	$-.45^{***}$	$-.38^{**}$	$-.37^{**}$	$-.28^{*}$	-.18	$-.25^{*}$
Yrs. Education	.14	.21	.12	-.20	.05	-.01
MoCA Score	$.37^{**}$	.15	.08	.18	.22	$.33^{**}$
WRAT-4 Score[a]	.04	.16	-.08	.04	.04	.09

Note. Two-tailed Kendall's-tau correlation values are presented. *N*-Back accuracy variables represent % correct trials for each task.

[a] WRAT-4 word reading subtest score.

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Participant mean RT across *n*-back conditions is shown in Figure 2 and Table 1. RT was higher for 2-back than 1-back across all information types. RT was highest for *SynBack* tasks, and lowest for *PhonoBack* tasks. The difference in RT between *SynBack* and the other tasks is greater at the 1-back level than 2-back. These observations were confirmed by statistical analysis which revealed Greenhouse-Geisser corrected main effects of *n*-back level, $F(1, 33) = 100.64$, $p < .01$, and information type, $F(1.68, 26.88) = 62.59$, $p < .01$. There was a significant interaction between information type and *n*-back level, $F(1.68, 26.88) = 7.87$, $p < .01$. Analysis of the interaction revealed the difference between RT for *SynBack* compared to *PhonoBack* and *SemBack* was greater at the 1-back level ($M_{\text{Difference}} = 509.58$, 371.91 ms, 95% CIs [399.19 – 619.97], [263.87 – 479.96], respectively, $p = .001$) compared to 2-back ($M_{\text{Difference}} = 392.36$, 177.43 ms, 95% CIs [287.52 – 497.19], [74.35 – 280.51], respectively, $p = .001$).

Age was correlated with RT for *SemBack* 1-back and *SynBack* 2-back tasks only ($rs_{\tau} = .248 – .309$, $p < .04$). Education was negatively correlated with RT for *PhonoBack* 2-back and *SynBack* 1-back tasks ($rs_{\tau} = -.268 – -.295$, $p < .04$). Neither MoCA score nor WRAT-4 Reading score were significantly correlated with *n*-back RT for any condition (see Table 3).

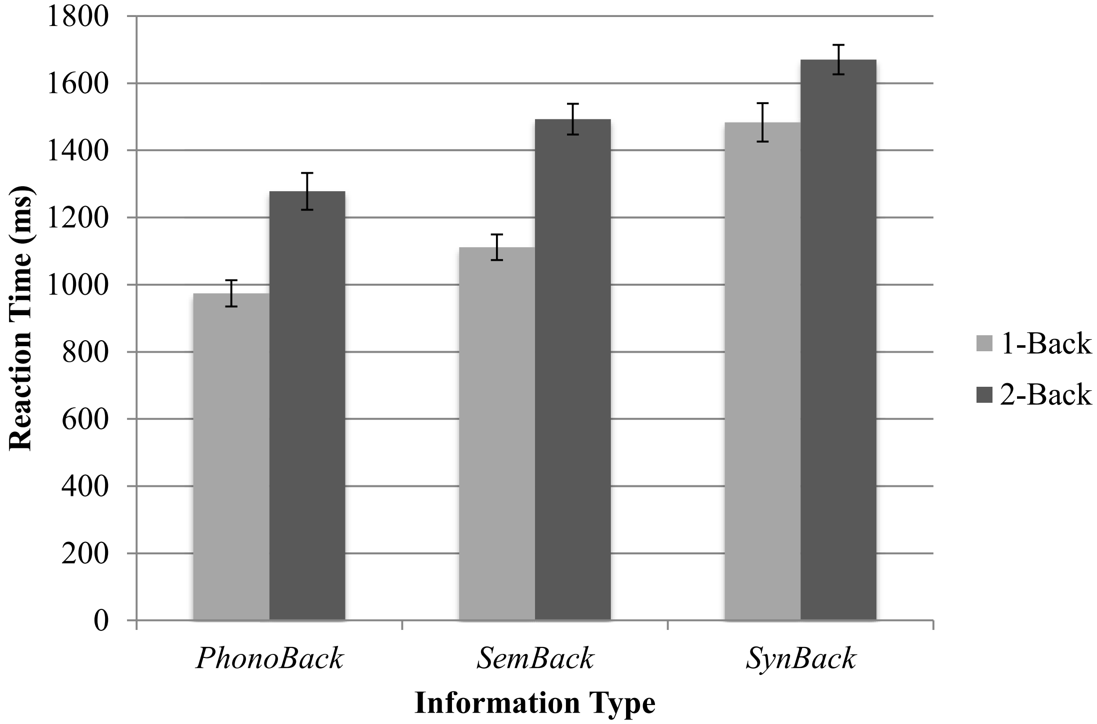


Figure 2. Mean reaction time for *n*-back tasks for each information type and level. Error bars represent standard errors of the mean.

Table 3

Correlations for Demographic Variables and N-Back RT (N = 34)

	PhonoBack		*SemBack*		*SynBack*	
Variable	1-back	2-back	1-back	2-back	1-back	2-back
Age	.10	-.10	.25*	.08	.09	.31**
Yrs. Education	.23	-.30*	-.11	-.11	-.27*	-.09
MoCA Score	-.14	.01	-.09	.04	-.16	-.15
WRAT-4 Score[a]	-.11	-.17	-.11	-.11	-.10	-.09

Note. Two-tailed Kendall's-tau correlation values are presented. *N*-Back RT variables represent mean RT for each task.

[a] WRAT-4 word reading subtest score.

* $p < .05$. ** $p < .01$.

3.3 DCT

Figure 3 shows participant mean scores on each of the four subscores of the DCT. Accuracy for Main Idea questions was at ceiling level, and higher than accuracy for Detail questions. The difference between accuracy for Stated vs. Implied questions was greater for Detail than Main Idea subscores. These observations were confirmed by statistical analysis which revealed main effects of salience, $F(1, 33) = 90.85$, $p < .01$, directness, $F(1, 33) = 21.50$, $p < .01$, and an interaction between salience and directness, $F(1, 33) = 4.50$, $p = .042$. Analysis of the interaction revealed a greater difference between Stated and Implied scores on Detail questions ($M_{\text{Difference}} = 1.44$, 95% CI [.64 – 2.25], $p = .001$) compared to Main Idea questions ($M_{\text{Difference}} = .47$, 95% CI [.10 - .84], $p = .014$).

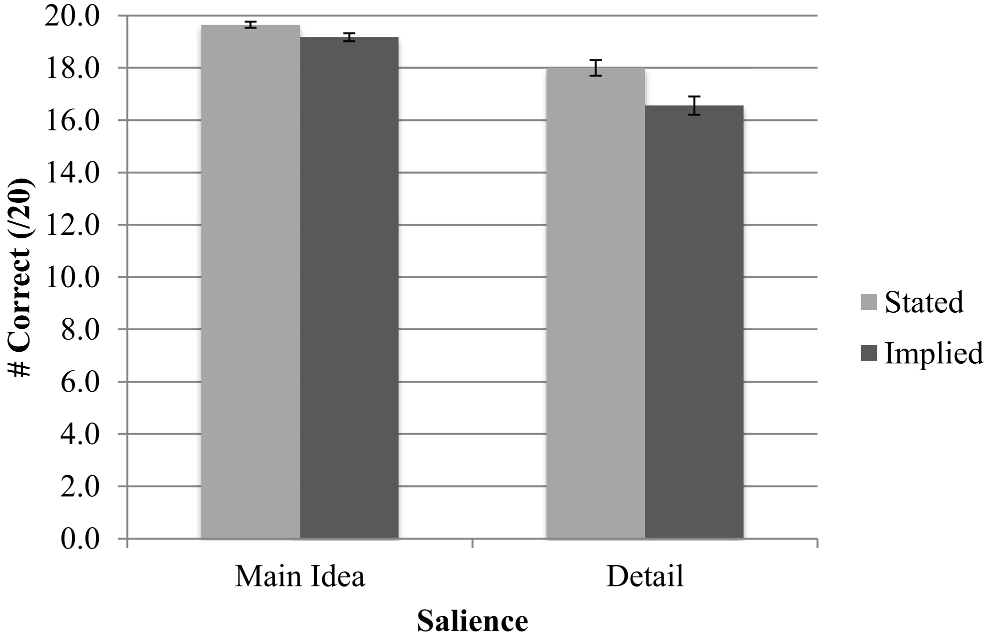


Figure 3. Accuracy on *Discourse Comprehension Test* subtests, based on salience and directness of information in comprehension questions. Error bars represent standard errors of the mean.

DCT total score was correlated with education ($r_\tau = .266$, $p < .05$) and approached significance with WRAT-4 Reading score ($r_\tau = .235$, $p = .068$). Neither age nor MoCA score were correlated with DCT total score (see Table 4).

Table 4

Correlations for Demographic Variables and DCT Total Score (N = 34)

Variable	DCT Total Score
Age	-.09
Yrs. Education	.27*
MoCA Score	.16
WRAT-4 Score^{a}	.24

Note. Two-tailed Kendall's-tau correlation values are presented.
^{a} WRAT-4 word reading subtest score. * $p < .05$.

3.4 *N*-Back and DCT

DCT total score correlated significantly with *n*-back accuracy in two tasks: *PhonoBack* 2-back ($r_\tau = .343$, $p < .01$), and *SynBack* 1-back ($r_\tau = .331$, $p = .01$). Correlations between DCT and *n*-back approached significance for *PhonoBack* 1-back ($r_\tau = .247$, $p = .054$) and *SynBack* 2-back ($r_\tau = .221$, $p = .082$). The DCT and *SemBack* correlations were non-significant ($p > .20$). The *n*-back tasks that related to DCT total score were then correlated with the DCT subscores. The Main Idea – Stated and Main Idea - Implied scores did not correlate significantly with any of the *n*-back scores ($p > .10$), likely due to the fact that DCT performance was at ceiling levels for these questions. The Detail - Stated score correlated significantly with *SynBack* 1-back ($r_\tau = .349$, $p < .01$). The Detail - Implied score correlated significantly with *PhonoBack* 2-

back ($r_\tau = .264, p < .05$) (see Table 5). DCT total score was not correlated with reaction time for any *n*-back task ($p > .24$).

Hierarchical regression analyses examined whether the *n*-back tasks were predictive of DCT total score when the effects of the demographic variables were taken into account (see Table 6). In these models, age and education were entered first and then followed by *n*-back accuracy. WRAT-4 Reading was not included because it correlated with education ($r_\tau = .42, p < .01$), and MoCA score was not included because it did not correlate with the DCT. Each *n*-back task was entered separately because of their high intercorrelations. Model 1, with age and education as predictors, explained 10.9% of the variance in DCT total score and approached significance, $F(2, 31) = 3.01, p = .064$. Adding *PhonoBack* 2-back and *SynBack* 1-back to the regression model both explained significantly more variance in DCT. However, *PhonoBack* 1-back and *SynBack* 2-back, which approached significance in the correlation analysis, were not significant predictors of DCT in the regression model. The *SemBack* tasks were also not significant additions to the regression model. Regression analysis was repeated with *n*-back RT data and demographic variables. Adding *n*-back RT to the regression models after education and age did not explain additional variance in DCT for any *n*-back task ($p > .08$).

Table 5

Correlations for DCT Scores and N-Back Accuracy (N = 34)

	PhonoBack		*SemBack*		*SynBack*	
DCT Score	1-back	2-back	1-back	2-back	1-back	2-back
Total[a]	.33**		.11		.26*	
	.25	.34**	.16	.14	.33**	.22
Subscores						
MIS[b]	-.02	-.01			.01	-.09
MII[c]	.11	.13			.23	.07
DTS[d]	.17	.20			.35**	.10
DTI[e]	.17	.26*			.21	.22

Note. Two-tailed Kendall's-tau correlation values are presented. *N*-Back accuracy variables represent % correct trials for each task. *SemBack* correlation values were not computed for DCT subscores because they did not correlate with total score.

[a] Correlation values for DCT total score are presented for combined 1-back and 2-back scores (top) and individual level scores (bottom) within each condition. [b] Main Ideas – Stated. [c] Main Ideas – Implied. [d] Details – Stated. [e] Details – Implied.

$^{*}p < .05$. $^{**}p < .01$.

Table 6

Hierarchical Regression Analyses Predicting DCT Total Score From Demographic Variables and N-Back Accuracy (N = 34)

Predictor	*B*	*SE (B)*	β	*t*	R^2	ΔR^2	ΔF
Model 1					.16	.16	3.01
Age	-.07	.08	-.15	-0.91			
Yrs. Education	.61	.27	.38	2.30^{*}			
Model 2a					.20	.04	1.44
Age	-.02	.09	-.03	-0.16			
Yrs. Education	.56	.27	.35	2.09^{*}			
Phono 1-back	.08	.07	.23	1.20			
Model 2b					.27	.11	4.55^{*}
Age	.04	.09	.08	0.43			
Yrs. Education	.42	.26	.26	1.59			
Phono 2-back	.14	.07	.42	2.13^{*}			
Model 2c					.20	.03	1.24
Age	-.01	.10	-.01	-0.06			
Yrs. Education	.55	.27	.34	2.05^{*}			
Sem 1-back	.12	.11	.23	1.11			
Model 2d					.19	.03	1.02
Age	-.03	.09	-.06	-0.30			
Yrs. Education	.67	.27	.41	2.46^{*}			
Sem 2-back	.06	.06	.19	1.01			
Model 2e					.34	.17	7.82^{**}
Age	.03	.08	.06	0.33			
Yrs. Education	.53	.24	.33	2.18^{*}			
Syn 1-back	.30	.11	.47	2.80^{**}			
Model 2f					.24	.08	3.00
Age	.001	.08	.002	0.01			
Yrs. Education	.57	.26	.36	2.22^{*}			
Syn 2-back	.16	.09	.32	1.73			

Note. Two-tailed regression values are presented. Each *n*-back task was entered in a separate regression model.

$^{*}p < .05$. $^{**}p < .01$.

Chapter 4: Discussion

4.1 Verbal Working Memory and Discourse Comprehension

Our study evaluated whether working memory for phonological, semantic, or syntactic information as assessed by *n*-back tasks was related to comprehension of written discourse in older adults. We found relationships between working memory for phonological and syntactic information and discourse comprehension. *N*-Back and discourse comprehension tasks were presented visually to explore the role of linguistic working memory in reading comprehension. To our knowledge, our study provides the first use of visual stimuli to measure working memory for phonological, semantic, and syntactic information, and the first evaluation of how these working memory measures relate to comprehension of written discourse.

DCT performance was positively related to *PhonoBack* and *SynBack* accuracy, but not *SemBack* accuracy. These results are different from what had been initially hypothesized. Why were phonological and syntactic working memory related to discourse comprehension, but not semantic? Penney (1989) provided evidence that written text stimuli are converted into phonological codes for subvocal rehearsal during short-term memory tasks. This could explain the relationship we found between phonological working memory and discourse comprehension. Participants may have been rehearsing content from the DCT stories to answer the comprehension questions. The relationship between *SynBack* performance and discourse comprehension could be related to a syntactic processing skill central to understanding both sentences and discourse. Discourse comprehension is a complex task. It requires a variety of cognitive processes, including reading skill, attention, and working memory. The small but significant correlations we found between measures of linguistic working memory and discourse comprehension could reflect the fact that working memory comprises only part of the processing necessary. It is important to note that accuracy across all three *n*-back tasks was high, without much variability across participants. This may have reduced the magnitude of the correlations between *PhonoBack* and *SynBack* and DCT. The tasks could have been made more challenging by adding additional levels of processing (i.e., 3-back, 4-back, etc). If the task had been more challenging the magnitude of the correlations might have been greater in all conditions, including *SemBack*. However, it is also possible that the linguistic process being tapped by the *SemBack* task, in which participants judge whether two nouns belong to the same category, is not important to discourse comprehension. Instead of single-word semantics, the use of context and background knowledge appears most important to link ideas across sentences during discourse comprehension (Nicholas & Brookshire, 1995). To make the semantic task more similar to the semantic processing used in discourse comprehension, one could modify the *SemBack* task to require judgments between items related to context instead of category.

N-Back tasks found to be correlated with DCT total scores were also correlated with the four subscores. The main pattern that emerged demonstrated relationships with Detail subscores. These correlations presumably were due to lower overall accuracy on Detail questions (Main Idea subscores were at ceiling). The relationship with Detail subscores could also reflect the higher cognitive demands associated with understanding details in discourse compared to main ideas (Nicholas & Brookshire, 1995). With greater variability in DCT subscore accuracy, a more consistent pattern with *n*-back scores may emerge.

After adding the *n*-back scores to a regression model with demographic variables, our findings were generally the same as our simple correlation results. That is, performance on the *PhonoBack* 2-back and *SynBack* 1-back working memory tasks predicted discourse comprehension beyond the contribution of age and education. These results suggest that the *n*-back tasks were tapping a working memory ability that uniquely affects comprehension of discourse.

N-Back accuracy was higher and reaction time was shorter for 1-back compared to 2-back across all tasks. Given the greater processing requirements of the 2-back level, these results are expected, and conform to patterns found in the *n*-back literature (e.g., Cansino et al., 2013; Daffner et al., 2011; Wright et al., 2007).

Age and *n*-back accuracy were negatively correlated; older participants tended to have lower accuracy on the *n*-back tasks. Using a similar set of *n*-back stimuli, Hayes (2011) found no performance difference between younger (18-30 years) and older (50-90 years) participants. However, the task requirements in Hayes' study differed from ours. Hayes required their participants to make exact matches between items (i.e., indicating when the current item in the sequence was identical to one presented before). In our study, participants made judgments about relationships between items (i.e., whether words rhymed or belonged to the same category, or whether sentences described the same event). Adding the linguistic processing requirement appears to make the tasks more sensitive to age. Studies have shown a continuous decline for processing-intensive tasks, including working memory for verbal information, across the adult lifespan. These declines in working memory have been measured using letter-matching *n*-back (Cansino et al., 2013) and complex span tasks (Park et al., 2002). Our study provides preliminary evidence for an age-related decline in verbal working memory for specific linguistic processes using the *n*-back paradigm. Future research could use our experimental methods to directly compare younger and older adults.

We found negligible differences in accuracy across the phonological, semantic, and syntactic *n*-back tasks. While *SynBack* accuracy was higher overall, this was likely due to a task-related factor: a number of participants reported they completed the task – especially at the 2-back level – by matching individual nouns and verbs found in the sentences instead of processing the entire sentence as a unit. To avoid this, we could modify the task to prevent matching individual words across trials. One way to do so could involve alternating text versions of reversible sentences (e.g., *the girl chased the dog*) with pictures depicting the scenes described in the sentences.

There did not appear to be differences in performance in our sample of non-brain-damaged adults for tasks tapping phonological, semantic, and syntactic working memory. One interpretation of these findings is that they lend support to general processor theories of verbal working memory (e.g., Just & Carpenter, 1992), which suggest the presence of a unitary resource for processing verbal information in the working memory system. Wright et al. (2007) found participants with a variety of different types of aphasia performed better on *SemBack* tasks compared to *PhonoBack* and *SynBack*. It is possible that poorer performance on a particular task reflects a generalized impairment to language functioning in the corresponding area (e.g., phonology) instead of a deficit specific to working memory. Future studies could systematically examine performance patterns on these tasks among individuals with different types of aphasia to better isolate the effects of linguistic working memory deficits in aphasia. Another interpretation of the consistency of working memory performance across conditions is that the *n*-back tasks were not constructed in a way to truly distinguish working memory subsystems for different types of linguistic information. We did not directly examine the degree to which our *n*-back results support general or modular processor theories of working memory. To do so, one might modify the *n*-back task to contain the same stimulus format (e.g., a simple sentence) for each linguistic task and require different judgments across conditions.

Overall DCT performance was high. Accuracy was at ceiling for the Main Ideas questions and marginally lower for the Details questions. Despite high accuracy, there was a greater difference between Stated vs. Implied information in the Details questions compared to Main Ideas. Accuracy was lowest for the Details – Implied subscore. These results are consistent with other studies of the DCT (Nicholas & Brookshire, 1995) and similar discourse comprehension measures (Ferstl, Walther, Guthke, & von Cramon, 2005) in non-brain damaged adults. Similar patterns of comprehension have been found in people with aphasia (Nicholas & Brookshire, 1995), early- and middle-stage Alzheimer's dementia (Welland, Lubinski, & Higginbotham, 2002), traumatic brain injury (Kennedy & Nawrocki, 2003), Huntington's disease and Parkinson's disease (Murray & Stout, 1999). Nicholas and Brookshire (1995) provided rationale for these patterns of performance in healthy and clinical populations. Heuristic processes – the use of context and general knowledge to facilitate comprehension – may be best suited to global information compared to details. With less support from heuristic processes, understanding details in discourse becomes more cognitively demanding. Similarly, constructing inferences to answer questions about implied information requires more cognitive resources than using information that was directly stated in the text. The reader must first recognize that the information needed to construct an inference is not stored in verbatim form in his/her memory. Next, the reader needs to identify relevant elements of information, and establish connections and relationships between these elements, in order to answer the question. As such, people can generally recall the gist of passages better than details, especially for information that was directly presented in the passage.

4.2 Limitations

There are a number of limitations to our study. First, our sample was a relatively homogeneous group of educated individuals (mostly women). Education was not found to correlate with *n*-back accuracy in our sample, although it did correlate with DCT total score. A wider range of education level may have created more variability in performance on both working memory and discourse comprehension tasks. In particular, participants with lower education may have lower accuracy on the DCT questions concerning details and implied information. They may also perform worse on the 2-back versions of the working memory measures. Daffner and colleagues (2011) found no difference in years of education between high and low performers on a verbal 2-back working memory task. However, their task, which involved matching letters, may not have been difficult enough to reveal an education effect on performance. Carpenter and Just (1989) emphasized the importance of taking individual differences in cognitive processing, including working memory capacity, into account when investigating its role in language comprehension. A more heterogeneous sample would provide a wider range of individual differences in performance on our measures, and would be more representative of the population at large. In addition, a larger sample size may also lead to stronger correlations between performance on our working memory and discourse comprehension tasks.

Second, we used the DCT to measure discourse comprehension among adults without brain damage. The DCT was developed to identify deficits in processing connected discourse in clinical populations, in listening and reading versions. Performance data were provided for 40 non-brain-damaged adults and 60 brain-damaged adults (aphasia, traumatic brain injury, or right-brain damage) for the listening version. The reading version contained performance data for 20 non-brain-damaged adults only. Our participants' scores were in keeping with those of the normative sample. The small normative sample, especially for the reading version, is a weakness of this test. In his review of the DCT, Marco (1998) describes other limitations. Each test story describes a humorous situation and ends in a punch-line. The humorous tone of the stories may limit the degree to which performance can be generalized to a range of discourse needs in day-to-day life. While measures of content, construct, and criterion-related validity were provided in the test, Marco asserts these data do not show whether the DCT can identify brain-damaged adults with discourse comprehension problems. While this claim is less relevant for our study on adults without brain damage, it should be taken into account for future research using the DCT in clinical populations. We selected this test because it measures comprehension systematically across domains of salience and directness, and to allow future comparisons with data from clinical populations. Unfortunately, few clinical tests are available for assessing discourse. Another option to study discourse comprehension in non-clinical populations is the *Nelson-Denny Reading Test* (Brown, Fishco, & Hanna, 1993), which contains reading comprehension passages ranging up to college level of difficulty. However, these passages would likely be too complex for use in clinical studies. Future research should aim to design measures to assess functional reading comprehension in clinical and non-clinical populations.

Third, our study had some task-based limitations related to the *n*-back measure of working memory. In the *PhonoBack* tasks, some participants self-reported matching words orthographically instead of phonologically. That is, they looked for rhyming pairs based on the last letters of the words instead of mentally sounding them out. Rastle and Brysbaert (2006) investigated the relationship between phonological and orthographic processing of written words. They found words that were preceded by phonologically identical pseudoword primes (e.g., kake-CAKE) were processed faster on lexical decision tasks compared to words preceded by phonologically dissimilar primes (e.g., pake-CAKE). The authors suggested that phonological processing occurs rapidly and automatically during reading. This evidence suggests phonological processing may still have been occurring during the *PhonoBack* tasks, even if participants perceived themselves to be using orthographic cues to complete the tasks. Still, there is some conflation of visual and auditory streams in measuring phonological processing using a visual task. Similarly, in the *SynBack* tasks, some participants self-reported matching individual nouns and verbs in the sentences instead of processing the sentence as a whole. That is, a participant might have received an accurate score for matching the sentences *The banker chased the golfer* and *The golfer was chased by the banker* based on the fact that they both contain the noun *golfer*, instead of recognizing that both sentences describe the same event. Participants reported matching individual nouns and verbs in sentences as a ‘short-cut’ strategy to complete the task, especially for the 2-back condition in which the task demands were highest. As a result, it is difficult to tell whether *SynBack* accuracy data truly reflect the syntactic processing construct we had been trying to measure. However, heuristic strategies have been shown to be effective in memory and comprehension tasks among older adults (Burke & Shafto, 2008). Participants who employ heuristic strategies during the sentence processing task may use similar strategies during discourse comprehension.

4.3 Implications and Future Directions

Age-related declines in language processing can impair communication effectiveness, and can have a negative impact on older adults’ social interactions and psychological well-being (Thornton & Light, 2006). One’s ability to understand discourse could affect functioning in a number of ways, including being able to follow detailed instructions and grasp the nuances of a story. These skills can be especially important to older adults who wish to maintain their independence. Our study demonstrated relationships between working memory for phonological and syntactic information and discourse comprehension in older adults. This research contributes to our growing understanding of the relationship between cognitive domains, like working memory, and broader functioning. Future research may evaluate this relationship more clearly in people with aphasia, who may have selective deficits in one type of working memory.

References

Abrams, L., & Farrell, M. T. (2011). Language processing in normal aging. In J. Guendouzi, F. Loncke, & M. J. Williams (Eds.), *The handbook of psycholinguistics and cognitive processes: Perspectives in communication disorders* (pp. 49–74). New York: Psychology Press.

Baddeley, A. (1995). Working memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 755–764). Cambridge, MA: MIT Press.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. Annual *Review of Psychology, 63,* 1–29.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory.* (pp. 47–89). New York: Academic Press.

Brookshire, R. H., & Nicholas, L. E. (1993). *Discourse Comprehension Test.* Tuscon, AZ: Communication Skill Builders.

Brown, J. A., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test: Manual for scoring and interpretation, Forms G & H.* Rolling Meadows, IL: Riverside Publishing.

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 373–444). New York: Psychology Press.

Cansino, S., Hernández-Ramos, E., Estrada-Manilla, C., Torres-Trejo, F., Graciela Martínez-Galindo, J., Ayala-Hernández, M., … Rodríguez-Ortiz, M. D. (2013). The decline of verbal and visuospatial working memory across the adult life span. *Age, 35,* 2283–2302.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77–126.

Carpenter, P. A., & Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 31–68). Hillsdale, NJ: Lawrence Erlbaum Associates.

Caspari, I., Parkinson, S. R., Lapointe, L. L., & Katz, R. C. (1998). Working memory and aphasia. *Brain and Cognition, 223*(37), 205–223.

Christensen, S. C., & Wright, H. H. (2010). Verbal and non-verbal working memory in aphasia: What three n-back tasks reveal. *Aphasiology, 24*(6-8), 752–762.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12(*5), 769–786.

Crerar, M. A. (2004). Aphasia rehabilitation and the strange neglect of speed. *Neuropsychological Rehabilitation, 14,* 173–206.

Daffner, K. R., Chong, H., Sun, X., Tarbi, E. C., Riis, J. L., McGinnis, S. M., & Holcomb, P. J. (2011). Mechanisms underlying age- and performance-related differences in working memory. *Journal of Cognitive Neuroscience, 23*(6), 1298–314.

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

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Ferstl, E. C., Walther, K., Guthke, T., & von Cramon, D. Y. (2005). Assessment of story comprehension deficits after brain damage. *Journal of Clinical and Experimental Neuropsychology, 27,* 367–384.

Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language, 86,* 23–39.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 22* (pp. 193–225). New York: Academic Press.

Hayes, R. A. (2011). *The inner workings of working memory: The effects of aging and language impairment on tasks examining verbal working memory* (Unpublished master's thesis). University of Pittsburgh, PA.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by PET 303. *Journal of Cognitive Neuroscience, 9*, 462–475.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Kennedy, M. R. T., & Nawrocki, M. D. (2003). Delayed predictive accuracy of narrative recall after traumatic brain injury: salience and explicitness. *Journal of Speech, Language, and Hearing Research, 46*(1), 98–112.

LaPointe, L. L., & Horner, J. (1979). *Reading Comprehension Battery for Aphasia.* Austin, TX: Pro-Ed.

Love, T., & Oster, E. (2002). On the categorization of aphasic typologies: The SOAP (A test of syntactic complexity). *Journal of Psycholinguistic Research, 31*(5), 503–529.

Marco, G. L. (1998). Review of the Discourse Comprehension Test. In J. Impara & B. Plake (Eds.), *Mental measurements yearbook* (pp. 374–375). Nebraska: Buros Institute of Mental Measurements.

Mayer, J. F., & Murray, L. L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders, 45,* 325–339.

McNeil, M. R., Sung, J. E., Pratt, S. R., Szuminsky, N. J., Kim, A., Ventura, M. . . ., & Musson, N. (2008). Concurrent validation of the Compuerized Revised Token Test (CRTT) and three experimental reading CRTT-R versions in normal elderly individuals and persons with aphasia. In *Clinical Aphasiology Conference*. Teton Village, WY.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 101*(2), 343–352.

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the N-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology, 24*(7), 711–717.

Miyake, A., Just, M. A., & Carpenter, P. A. (1994). Working memory constraints on the resoluion of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language, 33*(2), 175-202.

Murray, L. L., & Stout, J. C. (1999). Discourse comprehension in Huntington's and Parkinson's diseases. *American Journal of Speech-Language Pathology, 8,* 137–148.

Nasreddine, Z. (2010). *Montreal Cognitive Assessment, Version 7.1.* mocatest.org.

Nicholas, L. E., & Brookshire, R. H. (1995). Comprehension of spoken narrative discourse by adults with aphasia, right-hemisphere brain damage, or traumatic brain injury. *American Journal of Speech-Language Pathology, 4*(3), 69–81.

Papathanasiou, I., & Coppens, P. (2013). Basic concepts and operational definitions. In I. Papathanasiou, P. Coppens, & C. Potagas (Eds.), *Aphasia and related neurogenic communication disorders* (pp. xx–xxii). Burlington, MA: Jones & Bartlett Learning.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17*(2), 299–320.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*(4), 398–422.

Perlstein, W. M., Cole, M. A., Demery, J. A., Seignourel, P. J., Dixit, N. K., Larson, M. J., & Briggs, R. W. (2004). Parametric manipulation of working memory load in traumatic brain injury: behavioral and neural correlates. *Journal of the International Neuropsychological Society, 10*(5), 724–741.

Potagas, C., Kasselimis, D., & Evdokimidis, I. (2011). Short-term and working memory impairments in aphasia. *Neuropsychologia, 49,* 2874–2878.

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology, 53*(2), 97–145.

Rubin, D. L., Hafer, T., & Arata, K. (2000). Reading and listening to oral-based versus literate-based discourse. *Communication Education, 49*(2), 121–133.

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403–428.

Stine-Morrow, E. A. L., Soederberg Miller, L. M., Gagne, D. D., & Hertzog, C. (2008). Self-regulated reading in adulthood. *Psychology and Aging, 23*(1), 131–153.

Sung, J. E., McNeil, M. R., Pratt, S. R., Dickey, M. W., Hula, W. D., Szuminsky, N. J., & Doyle, P. J. (2009). Verbal working memory and its relationship to sentence-level reading and listening comprehension in persons with aphasia. *Aphasiology, 23*(7-8), 1040–1052.

Thornton, R., & Light, L. L. (2006). Language comprehension and production in normal aging. In J. E. Birren & K. Warner Schaie (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 262–288). Burlington, MA: Elsevier.

Ulatowska, H. K., Allard, L., & Chapman, S. B. (1990). Narrative and procedural discourse in aphasia. In *Discourse ability and brain damage: Theoretical and empirical perspectives* (pp. 180–198). New York: Springer-Verlag.

Welland, R. J., Lubinski, R., & Higginbotham, D. J. (2002). Discourse comprehension test performance of elders with dementia of the Alzheimer type. *Journal of Speech, Language, and Hearing Research, 45*(6), 1175–1187.

Wilkinson, G. S., & Robertson, G. S. (2006). *Wide Range Achievement Test 4.* Lutz, FL: Wide Range, Inc.

Wright, H. H., Downey, R. A., Gravier, M., Love, T., & Shapiro, L. P. (2007). Processing distinct linguistic information types in working memory in aphasia. *Aphasiology, 21*(6-8), 802–813.

Wright, H. H., & Fergadiotis, G. (2012). Conceptualizing and measuring working memory and its relationship to aphasia. *Aphasiology, 26*(3-4), 258–278.

Wright, H. H., & Shisler, R. J. (2005). Working memory in aphasia: Theory, measures, and clinical implications. *American Journal of Speech-Language Pathology, 14*(2), 107–118.

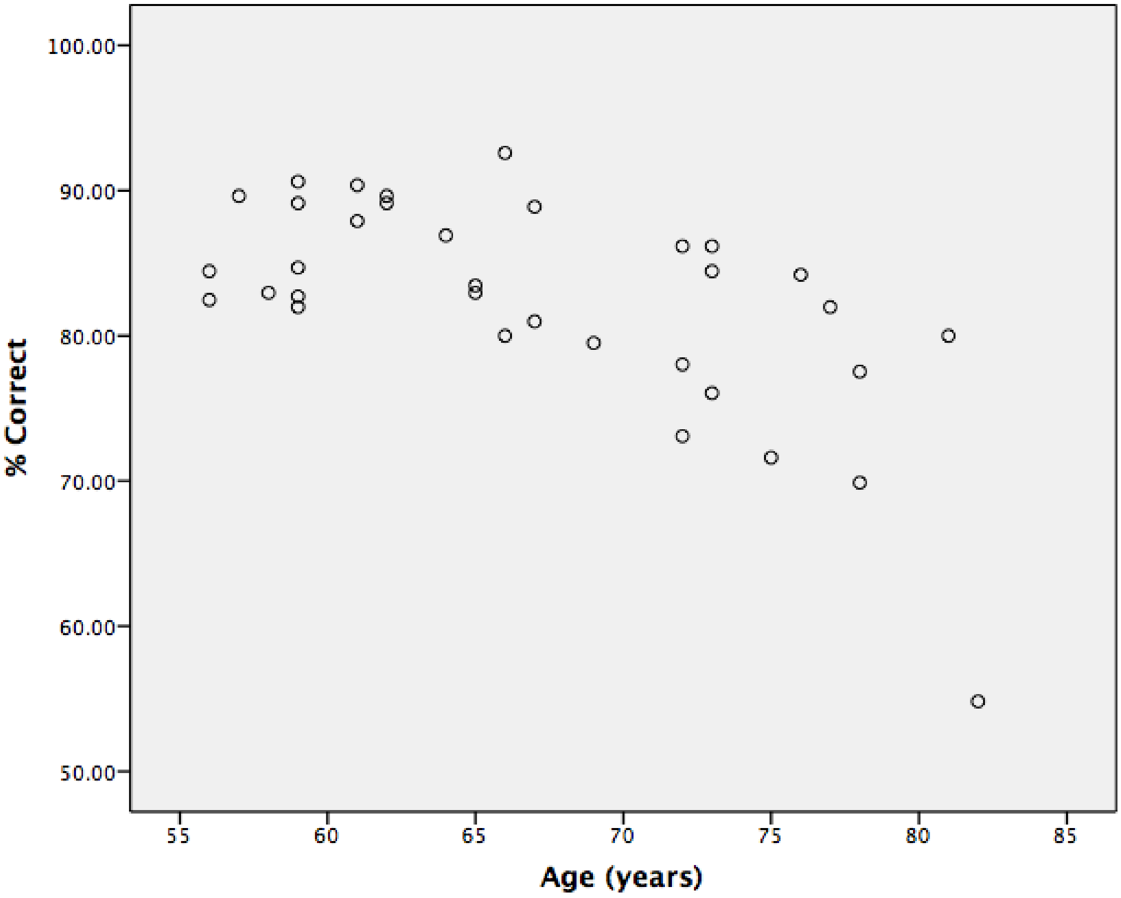
Appendix A: *N*-Back Stimuli

Stimuli for *PhonoBack, SemBack,* and *SynBack* tasks

***PhonoBack* words**	***SemBack* words**	***SynBack* words**
-at: bat, hat, rat, pat, mat *-ap:* tap, cap, map, nap, sap *-in:* pin, fin, bin, tin, kin *-ig:* wig, rig, jig, pig, fig - *ill:* till, hill, pill, sill, gill	*Animals:* wolf, cat, snake, bird, rabbit *Furniture:* chair, desk, dresser, couch, stool *Clothes:* shirt, hat, jacket, blouse, pants *Tools:* hammer, drill, pliers, hatchet, axe *Fruit:* apple, orange, lemon, grape, lime	*Verbs:* pushed, called, punched, kicked, thanked, blames, teased, kissed, chased, hugged [1] *Noun phrases:* actor, golfer, doctor, banker, singer, teacher, lawyer, baker, jogger, mayor [2]

Reproduced with permission from Wright et al., 2007. [1] All verbs used in passive and active tenses; [2] All noun phrases used as subjects and objects.

Appendix B: Relationship Between *N*-Back Accuracy and Age



Note. Accuracy is collapsed across information type and *n*-back level.

Appendix C: *N*-Back Accuracy Breakdown Across Conditions

	PhonoBack		*SemBack*		*SynBack*	
	1-back	2-back	1-back	2-back	1-back	2-back
Correct Hit	27.94	19.17	28.89	19.13	27.80	20.16
Correct Rejection	63.08	51.23	64.04	51.83	66.35	60.56
False Alarm	3.27	16.07	3.63	15.12	1.72	11.31
False Rejection	3.27	7.82	2.68	7.14	3.85	7.54
Missed	2.44	5.71	0.77	6.79	0.27	0.44

Note. Values represent % of total responses per condition and level.