Individual Differences in Working Memory Capacity Influence Spoken Word Recognition

Dissertation

Presented in partial fulfillment of the requirements for the degree Doctor of Philosophy in the graduate school of the Ohio State University

by

Christine Marie Szostak, Ph.D.

Graduate Program in Psychology

The Ohio State University

2013

Dissertation Committee

Mark A. Pitt, Ph.D., Primary Advisor

Per Sederberg, Ph.D., Secondary Advisor

Simon Dennis, Ph.D.

Eric W. Healy, Ph.D.
Copyright by

Christine Marie Szostak

2013
Abstract

Prior work has shown that when speech is unclear, listeners show a greater dependence upon semantic than on acoustic information to aid word identification when distracting stimuli (e.g., other talkers) are present. The current project extended this work to explore whether individual differences in working memory capacity (WMC) would influence the likelihood that listeners will depend on the biasing information when distracted. In five experiments, participants heard sentences that contained an early target word with or without noise at its onset and a subsequent word that was semantically biased in favor of the target word or one of its lexical competitors (e.g., *The wing had an exquisite set of feathers* or *The wing had an exquisite set of diamonds* where *diamonds* would be semantically associated with *ring*). The sentences were presented in the presence of distracters ranging in their degree of signal-similarity to that of the sentence (e.g., another speaker vs. an everyday nonspeech sound). Participants made target word identification and sentence sensibility judgments for each sentence they heard. The findings showed that those with lower WMC were more likely to depend upon biasing than on acoustic signal information, but only when the signal was masked by noise. In contrast, those with higher WMC showed less dependence upon the biasing information than those with lower WMC, even when the signal was masked by noise. Although performance across distracter similarity was not influenced by WMC, the likelihood of being able to
anticipate what distraction would be heard was shown to influence performance as a function of WMC. A discussion of the role of WMC in spoken word recognition, especially during distraction, is provided and the potential mechanisms involved in this process are considered.
Acknowledgements

I first wish to thank God for the strength and endurance to complete this second dissertation and to further grow and develop as a psychologist!

I also thank Carole and Donald Szostak for their continued love and support throughout this process. Also, I wish to thank them for continuing to believe in me through my many years of schooling!

I wish to also sincerely thank my advisor, Mark A. Pitt, Ph.D. for his very generous time and devotion to this process and for agreeing to learn about WMC along with me. His skill, expertise, support, and wealth of knowledge of spoken word recognition can be seen throughout this work!

Sincere thanks are also extended to Per Sederberg for his sincere enthusiasm about this project. His energetic attitude toward this project allowed me to continue to have the confidence necessary to tackle WMC for the first time, especially when the going got tough! Sincere gratitude is given to Eric Healy for being an important part of every step of my training beginning with my Master's Thesis and continuing all the way through helping me meet requirements for completing my doctorate. I also wish to sincerely thank Simon Dennis for taking me under his wing in learning about memory and for providing me with invaluable information, ideas, and thoughts on this project. The four members of my committee have provided the expertise, specialties, and invaluable guidance that have
made this project what it is. Without their mentorship, this project would not have been nearly the scholarly work it has become.

In addition, I wish to thank Sven Mattys for his generous and very helpful discussions and his continued interest in and enthusiasm for this project. It is because of his work in the past few years that this project came into existence. His time and willingness to help troubleshoot, provide materials for earlier versions of this project, and his much appreciated sage advice have genuinely helped this dissertation take on new shape. I also wish to thank Randy Engle for his contagious enthusiasm regarding this project and his willingness to provide the materials used for testing WMC. I also wish to thank Kenny Hicks for providing me with the OSpan materials and his assistance in learning to analyze and interpret OSpan scores. I also wish to thank Troy Smith who truly was my WMC guru. His knowledge and expertise with memory in general and WMC in particular provided me with invaluable directions that helped me when beginning on this journey. Without his guidance, beginning this endeavor would have felt monumentally more challenging.

I do not have the words to truly express the degree of gratitude I have for the very dedicated multimedia team at ODS for their amazing support, unbelievable dedication to my training, never ending enthusiasm, kind words when I needed them most, and even the occasional support for my caffeine addiction! I am especially grateful for the generosity, time, and kindness that Sarah Terry and Nicole Chattah have shown me over the past six years! Without them and their amazing staff's help, I never could have reached the level of success I have had!
Further, I can not thank Sheena Riepenhoff enough for always having the time to answer my many questions. Her enthusiasm, patience, and friendly attitude helped to make my time here at OSU sincerely enjoyable and pleasant!

I also sincerely thank Doug Finlay and Nolan Crabb for the generous assistance with all things IT. Without their efforts, time, and kind patience, I would not have been able to reach this stage in my training!

I also wish to thank the numerous students who have put their blood, sweat, and tears or at least their laughter, time, dedication, motivation, and research assistant hours into various parts of this project, most especially the many hours of performing and coding manual OSpan tasks! These truly dedicated individuals include: Hartman Brawley, Paige Lisco, Karin Luk, Chloe Meyer, Jason Miao, Kristen Peters, and Sara Stanutz. Along with these individuals, I thank the participants for their contribution to this project. Without their help and time, this project would not have been able to succeed.

I also extend sincere gratitude to Lauren Seifert, Ph.D. without whom, I would have likely never caught the cognitive psychology bug. Her dedication to starting me off on the right Christian foot as a psychologist will always remain an important part of my life. I only hope that in my future endeavors, I can learn to exude the same type of kindness and dedication that she has demonstrated with respect to my own future endeavors.

Finally, I wish to extend very special thanks to Armand Parker Szostak for sticking by my side throughout this process. His many enthusiastic tail wags, labbie kisses, and nuzzles provided much of the stress reduction necessary for completing a second dissertation.
Vita

Education

M.A., Cognitive Psychology, Department of Psychology, The Ohio State University, Columbus, Ohio.................................................................2007-2009

Ph.D., Clinical Psychology, Alliant International University, Fresno, California..........................................................................................2004-2007

Predoctoral Intern, Clinical Psychology, Kansas University Medical School, University of Kansas, Kansas City, Kansas.......................2006-2007

M.A., Clinical Psychology, Alliant International University, Fresno, California..........................................................................................2002-2004

B.A., Psychology, Malone College, Canton, Ohio.........................................1995-2000

Duo B.S., Elementary and Special Education with a Concentration in English, Malone College, Canton, Ohio.............................................1995-2000
Publications


Fields of Study

Major Field: Graduate Program in Psychology
Cognitive/Experimental Psychology
Psycholinguistics
Table of Contents

Abstract ........................................................................................................................................... ii
Acknowledgements .................................................................................................................. iv
Vita ................................................................................................................................................ vii
List of Tables ............................................................................................................................... xiv
List of Figures ............................................................................................................................... xv
Chapter 1: Introduction .............................................................................................................. 1
  The Challenge of Identifying Spoken Words .............................................................. 1
  Information Sources Aiding Word Recognition ................................................... 1
  The combined Influence of Acoustic and Lexical Information ..................... 3
  Backward Context Influences on Word Recognition During Distraction .... 5
  Working Memory Capacity ....................................................................................... 6
  WMC and Spoken Word Recognition ................................................................. 9
  Focus of the Current Study .................................................................................... 12
Chapter 2: Experiment 1 ....................................................................................................... 14
  Method ......................................................................................................................... 16
  Participants .................................................................................................................. 16
  Stimuli ......................................................................................................................... 16
  Target Talker Stimuli ............................................................................................... 16
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMC Analysis</td>
<td>60</td>
</tr>
<tr>
<td>Target Word Accuracy Analysis</td>
<td>60</td>
</tr>
<tr>
<td>Biased Word Response Analysis</td>
<td>60</td>
</tr>
<tr>
<td>Summary</td>
<td>61</td>
</tr>
<tr>
<td>Chapter 7: General Discussion</td>
<td>62</td>
</tr>
<tr>
<td>Summary and Overview</td>
<td>62</td>
</tr>
<tr>
<td>Lexical Activation</td>
<td>66</td>
</tr>
<tr>
<td>Attentional Flexibility</td>
<td>67</td>
</tr>
<tr>
<td>Attentional Reallocation</td>
<td>69</td>
</tr>
<tr>
<td>Distracter Influences on Word Recognition</td>
<td>69</td>
</tr>
<tr>
<td>Alternative Causes</td>
<td>73</td>
</tr>
<tr>
<td>Concluding Remarks</td>
<td>76</td>
</tr>
<tr>
<td>References</td>
<td>77</td>
</tr>
<tr>
<td>Endnotes</td>
<td>83</td>
</tr>
<tr>
<td>Appendix A: Target Talker Sentences</td>
<td>85</td>
</tr>
<tr>
<td>Appendix B: Secondary Talker Sentences</td>
<td>87</td>
</tr>
<tr>
<td>Appendix C: Sound Corpus Categories</td>
<td>94</td>
</tr>
<tr>
<td>Appendix D: Table and Figures</td>
<td>96</td>
</tr>
<tr>
<td>Table 1</td>
<td>96</td>
</tr>
<tr>
<td>Figure 1</td>
<td>97</td>
</tr>
<tr>
<td>Figure 2</td>
<td>98</td>
</tr>
<tr>
<td>Figure 3</td>
<td>99</td>
</tr>
<tr>
<td>Figure 4</td>
<td>100</td>
</tr>
</tbody>
</table>
List of Tables

Table 1: Overall mean analysis for WMC, target word accuracy, and biased word responses............................................................................................................................96
List of Figures

Figure 1. The plots depict target word accuracy (1a) and biased word responses (1b) as a function of WMC in the congruent condition from Experiment 1. Target word accuracy $r = .03, p > .85$; biased word accuracy $r = -.08, p > .63$.......................................................97

Figure 2. The plots depict target word accuracy (2a) and biased word responses (2b) as a function of WMC in the incongruent condition from Experiment 1. Target word accuracy $r = .29, p < .08$; biased word response $r = .36, p < .05$..........................................................98

Figure 3. The plots depict target word accuracy (3a) and fricative responses (3b) as a function of WMC in the absent condition from Experiment 2. Target word accuracy $r = .32, p < .06$; fricative response $r = -.39, p < .05$...............................................................99

Figure 4. The plot depicts stop responses as a function of WMC in the absent condition from Experiment 2, $r = .31, p < .06$.................................................................................100

Figure 5. The plot depicts semivowel responses as a function of WMC in the absent condition from Experiment 2, $r = .27, p < .11$.......................................................................................101

Figure 6. The plot depicts nasal responses as a function of WMC in the absent condition from Experiment 2, $r = .11, p > .52$.......................................................................................102

Figure 7. The plots depict target word accuracy (7a) and biased word responses (7b) as a function of WMC in the incongruent condition from Experiment 3. Target word accuracy $r = .09, p > .57$; biased word response $r = -.11, p = .50$..................................................103
Figure 8. The plot depicts sentence sensibility judgments as a function of WMC in the incongruent condition from Experiment 3, \( r = .06, p > .74 \).............................................104

Figure 9. The plots depict target word accuracy (9a) and biased word responses (9b) as a function of WMC in the incongruent condition from Experiment 4. Target word accuracy \( r = .31, p < .07 \); biased word response \( r = -.24, p > .13 \)..................................................105

Figure 10. The graph depicts the mean OSpan scores for the highest 25% of participants (10a: high WMC) and lowest 25% of participants (10b: low WMC) for Experiments 1 (speech condition), 4 (nonspeech condition) and 5 (stable nonspeech condition), \( n = 8 \) per condition..........................................................106

Figure 11a. The graphs depict the mean proportion of target word accuracy for the highest 25% of participants (11a1: high WMC) and lowest 25% of participants (11a2: low WMC) for the incongruent condition of Experiments 1 (speech condition), 4 (nonspeech condition) and 5 (stable nonspeech condition), \( n = 8 \) per condition............107

Figure 11b. The graphs depict the mean proportion of biased word responses for the highest 25% of participants (11b1: high WMC) and lowest 25% of participants (11b2: low WMC) for the incongruent condition of Experiments 1 (speech condition), 4 (nonspeech condition) and 5 (stable nonspeech condition), \( n = 8 \) per condition...........................108

Figure 12. The plots depict target word accuracy (12a) and biased word responses (12b) as a function of WMC in the incongruent condition from Experiment 5. Target word accuracy \( r = .17, p > .31 \); biased word response \( r = -.36, p < .05 \)...................................109
Chapter 1: Introduction

The Challenge of Identifying Spoken Words

Lexical ambiguities occur frequently in speech. For example, cell phone interference may disrupt a critical portion of the speech signal (e.g., =ing where = represents a noise masker and thus the disrupted word could have been one of multiple intended words such as wing, thing, ring, and so on). To make matters worse, such ambiguities are often experienced in the presence of surrounding distracting stimuli (e.g., listening to a talker on the cell phone mentioned above in a noisy airport). Despite these challenges, listeners appear to identify the words intended by the talker easily and with little effort. How is this accomplished?

Information sources aiding word recognition. It has been long established that acoustic information can aid perception when speech is produced in noise (c.f., Mattys, White, & Melhorn, 2005; Rosen, 1992; Samuel, 1981; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995 for some different types of evidence supporting this argument). A second informational source that listeners can attend to during the processing of words when they are spoken in noise is that of biasing semantic information. This information may come from within the word itself (e.g., Ganong, 1980; McClelland & Elman, 1986; Norris, 1994; Norris, McQueen, & Cutler, 2000; Samuel, 1987), or from outside of this word (e.g., Cole, Jakimik, & Cooper, 1980; Connine,

For example, Samuel (1987) presented listeners with words that had no lexical competitors (words differing only in one target phoneme: e.g., *lengthen* where the underlined phoneme is the target) or words that have at least one lexical competitor that shares all but the target phoneme (e.g., *locket*, where *rocket*, *socket* and so on end in *ocket*). Samuel (1987) covered or replaced the target phoneme with noise (e.g., *ocket* vs. *=ocket*). He asked participants to report whether the target phoneme had been present or absent and to make a forced-choice decision as to which of two phonemes (e.g., /l/ or /r/) had been presented. His findings demonstrated that when the word had no competitors, listeners were more likely to report that the phoneme had been present and were more accurate in identifying the target phoneme. Samuel (1987) argued that his findings suggested that a listener's prior knowledge of words and the presence of lexical competitors can influence perception.

Others have found that as with the influence of lexical knowledge (e.g., the fact that *lengthen* but not *rengthen* is a word), semantic information beyond the word itself (e.g., within a phrase or sentence) can also influence word identification when speech is ambiguous or masked by noise. For example, Samuel (1981, 1990) presented target words such as *cavern* or *tavern* that had noise masking a target phoneme (e.g., /k/ or /t/) to his participants. Again, the noise was either added to or had replaced the target phoneme (e.g., *cavern* vs. *=avern*). Each target word was placed in a sentence that had a word biased toward one of the two lexical competitors. The biased word was either presented prior to (Samuel, 1981) or following (Samuel, 1990) the target word (e.g., *We*
heard that there were horrible bats in the tavern where the underlined word is the biasing word). Listeners made phoneme presence/phoneme identification judgments. His findings showed that listeners were more likely to report that the phoneme was present when they reported that the phoneme had been the phoneme that would result in a meaningful sentence (e.g., reporting /k/ when hearing bats). Szostak and Pitt (2012) extended these findings to demonstrate that when the biasing word followed the noise masked word, the biasing word maintained its influence on target word identification, even when the two words were several syllables away from one another (e.g., The =ing had an exquisite set of feathers where the target word and feathers are eight syllables or approximately 1.1 seconds away from one another.

Collectively, all of the above findings suggest that listeners are not only sensitive to the information within the acoustic signal, but that they can also use their own experiences as well to aid word identification. This includes both knowledge of whether something is a word within the spoken language and knowledge about how the world works or whether something is semantically meaningful.

**The combined influence of acoustic and lexical information.** Given that both acoustic-phonetic information in the signal and semantically biasing information can aid word recognition, a specific interest for researchers has been to identify what types of informational (signal or semantic) sources are primarily relied upon by the processor in different listening situations. For example, work reported by Mattys, Carroll, Li, and Chan (2010) examined whether the information that is depended upon (signal-based or knowledge-based), when listening to unclear speech, will shift when a secondary talker's speech is competing with the talker's speech that is being processed. In their study,
participants were presented with ambiguities such as, *mild option*, where the intended phrase could have been either *mild option* (a word-word phrase) or *mile doption* (a word-pseudoword phrase). Importantly, their stimuli were presented in such a way that acoustic (e.g., glottalization) and lexical (e.g., *option*) cues were pitted against one another. Each phrase was presented alone or in the presence of a competing (secondary) talker. Participants were required to identify whether the first word of the phrase had been *mild* or *mile*. Mattys et al. (2010) found that when the target talker was presented alone, listeners showed greater dependence upon the acoustic-phonetic features in the signal than on lexical information. When the target talker was presented simultaneously with a secondary talker, performance shifted such that the listeners showed more dependence on their lexical knowledge than on the information contained within the acoustic signal. Mattys et al. (2010) concluded that as long as a secondary talker is not present and the signal is clear, listeners will depend more heavily on the signal than on the biasing information. When listeners become distracted by a secondary talker, they will shift their dependence to more salient information (e.g., available semantically biasing information).

Mattys et al. (2010), however, only examined cases where biasing information immediately followed the ambiguity (e.g., *mild option* where the underlined portion represents the ambiguity). Thus, in their study, listeners only needed to maintain information about the word being identified in memory, at most, for several hundred milliseconds before the biasing information (e.g., *option*) became available. However, biasing semantic information does not always immediately follow the word the listener is identifying. For example, in the sentence *The wing had an exquisite set of feathers,*
several intervening words are present between the word being identified (underlined word) and the biasing information (e.g., *feathers*).

**Backward Context Influences on Word Recognition During Distraction**

At present, nothing is known as to whether distant biasing information maintains an influence on word identification when the listener is distracted by a secondary talker. If some type of biasing information does not become available immediately (e.g., *The =ing had feathers*), the lexical representation of the word being identified (e.g., *wing...*) must be maintained in memory while additional information from the talker being attended to is processed and the speech of other talkers is filtered out. Essentially, if several additional words are forthcoming that do not yet help confirm or reject the representation of the word, (e.g., *had an exquisite set of*), the lexical representation must maintain its activation level long enough for the system to continue to evaluate its likelihood of being the word that was intended by the talker.

The backward influence of subsequent biasing information has interesting implications for word recognition when a listener is distracted. If the listener is sufficiently distracted by the secondary talker, the ability to generate a rich lexical representation may not occur, which may be the reason Mattys et al.'s (2010) participants shifted their attention to the biasing information. It should be noted that the participants in their study were strongly encouraged to attend to the signal even when the signal suggested a word-pseudoword (*mile doption*) outcome. Warren & Warren, (1970) suggested that when biasing information follows the word being identified, it may be that the clear portion of the word is stored in memory (e.g., *ing*) and then the later biasing information along with the clear portion of the ambiguous word are synthesized together...
in order to produce a coherent sentence (e.g., \textit{ing} + \textit{feathers} = \textit{wing})\textsuperscript{1}. Regardless of whether a rich or an impoverished representation is stored, something must be maintained in memory over time in order for the biasing information to have an influence on word identification. If the processor must also continue to filter out a secondary talker's speech, such a situation would seem to tax memory even further.

\textbf{Working Memory Capacity}

Conway et al. (2005) argued that it would seem non-optimal to have a processor that could not maintain information in memory while other information is present that is interfering with the information in memory or is simultaneously being processed. For example, if hearing a target talker say something like: \textit{The wing had an exquisite set of feathers} while a secondary talker is producing an unrelated sentence (e.g., \textit{We ordered a pepperoni pizza for dinner last night}), the following must take place. If there is some uncertainty that \textit{wing} was intended by the target talker, the processor would need to maintain this word in memory while continuing to process further information so as to help confirm or reject this word. At the same time the above is taking place, the processor must also actively filter out the unrelated speech so as to ensure the message is properly processed. They noted that a theoretical construct known as working memory capacity (WMC) has been proposed as a measure of an individual's ability to engage in situations such as the above where the listener must process language while filtering out a distracting secondary talker (Conway et al., 2005).

Conway et al.'s (2005) theoretical framework holds that WMC is made up of two components. The first component is domain-general executive control/attention which is used for the processing of information. They define executive attention as the ability to
hold in active memory a memory representation while simultaneously processing other information or experiencing interference. Essentially, this is the amount of attention that an individual can devote to the processing of a stimulus or the maintaining of a stimulus in active memory (c.f., Conway et al., 2005; Engle, 2002). The second component is domain-specific resources such as rehearsal and strategies that are used in the storage of information. Simply put, executive attention can be thought of as the amount of space available to process information and the resources can be thought of as the things that are helping to actively maintain information within the available attentional space. Conway et al. (2005) argued that the purpose of this multi-domain system is to actively maintain information during distraction or while simultaneously processing other stimuli. For example, in the above case where a listener is attempting to process a target talker who is saying *The wing had an exquisite set of feathers* while ignoring a secondary talker, the following might be occurring. As *The wing* is perceived, the processor may identify an initial lexical candidate based on any available information (e.g., formant transitions... leading to *wing*). Once this has been accomplished, the representation is held in executive attentional space. While maintaining this information in memory, the system must also continue to process incoming information, potentially evaluating its ability to help support or refute identification of the word being identified. If the processor must also filter out a distracting secondary talker's speech, attentional capacity is likely going to be required for this task. Conway et al. (2005) argued that in situations where attentional demands are increased, such as when needing to actively ignore a secondary talker, the domain-general executive attention will become more depended upon. This would in tern
leave less attention available that can be devoted to the target talker's speech and thus would likely cause word identification to be impaired in some way.

Findings have suggested that executive attention, but not necessarily the resources of the domain-specific component of WMC, is strongly correlated with general fluid intelligence (c.f., Unsworth, 2010). Executive attention, more so than domain-specific resources, has also been found to be able to reliably predict performance on a number of complex cognitive tasks; including tasks that are both non-linguistic in nature (e.g., DeCaro, Thomas, & Beilock, 2008; Engle, 2002; Watson & Strayer, 2010) and linguistic in nature (e.g., Calvo, 2004; Collflesh & Conway, 2007; Conway, Cowan, & Bunting, 2001; Engle, 2002; Lustig, May, & Hasher, 2001; Otten & Van Berkum, 2009; Rai, Loshcky, Harris, Peck, & Cook, 2010). Conway et al. (2001), for example, presented participants with a passage read by a target talker in one ear and a to-be-ignored passage spoken by a secondary talker in the other ear. Participants were required to shadow, as closely as possible, the target talker's speech. Importantly, without the participant's knowledge, his or her name was inserted into the secondary talker's passage. Conway et al. (2001) found that those with lower WMC were nearly three times as likely to hear their own name as those with higher WMC and showed disruptions in the ability to accurately shadow the talker when the listener's name was spoken. Participants did not however show this same level of distraction when presented with another person’s name. When later queried about what they recalled from the ignored passage, those with lower WMC were reliably more likely to report having heard their name than those with higher WMC (Conway et al., 2001). The idea here is that although individuals with both lower and higher WMC would have been using their executive attention to filter out the
secondary talker, those with lower WMC had reached their capacity and thus were unable to properly completely filter out this irrelevant information. In contrast, those with higher WMC had enough attentional capacity available to adequately filter out the secondary talker's speech. It is this executive attention, therefore, that Conway et al. (2005, also see Engle, 2002; Kane et al., 2004) argued is what is really being measured when investigating WMC.

**WMC and Spoken Word Recognition**

Although a fair amount of work has been conducted to explore WMC, no direct work has examined whether individual differences in WMC are associated with how listeners recognize spoken words. Based on the multi-component theoretical framework of WMC described by Conway et al. (2005), one question of interest with respect to spoken word recognition is whether individual differences in WMC are associated with how spoken words are recognized when the listener is distracted. Some work (e.g., Kane & Engle, 2000; Unsworth, 2010) suggests that listeners with lower WMC are more likely to more quickly reach attentional capacity when engaged in complex cognitive tasks than those with higher WMC. Further, they argued that when such attentional capacity is reached, such individuals will show difficulty in being able to inhibit, or in Kane and Engle's (2000) terms, resist the influence of prior knowledge about semantically biasing information that may be interfering with word recognition when biasing information is present.

To study the issue of inhibition among listeners with lower vs. higher WMC, Kane and Engle (2000) presented participants with multiple lists of to-be-recalled words. Each list of words was selected from a single category (e.g., animals). Importantly, some
categories were used only once while others were repeated across multiple adjacent lists. For example, participants might be required to learn a list of animals followed by two lists of furniture. Immediately after learning a list of words, participants were required to recall as many of the words as possible. Importantly, the participants were required to study the lists and perform the recall task either in the presence of a distracting activity (e.g., performing a complex finger tapping sequence) or with no distraction.

Kane and Engle (2000) were interested in two questions. (1) Would differences in WMC emerge as a function of the presence of a distracter? (2) Would those with lower WMC be more likely to recall items from an earlier same-category list than those with higher WMC, and thus experience more interference from previously studied lists? Their findings indicated that when no distracter was present, those with higher WMC outperformed their lower WMC peers. One of the major differences across the two groups was that those with lower WMC reported more words from prior same-category lists (what they termed intrusions) than those with higher WMC. In contrast, when distracted, they found that those with lower WMC showed little to no change in their performance while those with higher WMC became indistinguishable from their lower WMC peers.

Kane and Engle (2000) argued that the reason for the overall poor performance seen from those with lower WMC, and the poorer performance for those with higher WMC under distraction, was as follows. When information is being actively stored in working memory, the information resonates with semantically related information that has been previously stored in long-term memory (LTM). This resonance causes the information stored in LTM to be activated. Because such information is not part of the current list,
this information must be adequately inhibited (suppressed or resisted against in their terms) in order for the participant to not produce an intrusion during recall. Those with lower WMC do not have the attentional capacity necessary to be able to adequately inhibit such interference in general and when distracted, those with higher WMC reach their attentional capacity and show this same inhibition deficit. Such an inhibition argument makes sense when considering that neuroimaging studies have found that one of the regions involved in WMC is the pre-frontal cortex, including the left inferior frontal gyrus. This region has been found to be important for inhibitory control (c.f., Jonides & Nee, 2006 for an elegant review of the literature). For example, one task, frequently used in imaging work to explore inhibitory control (which appears to be at least one cause for the ability to suppress interference), has been the recent negative probes task. During this task, participants are presented with a list of stimuli (e.g., words) and are then given a stimulus probe. The task for participants is to decide whether the probe was presented during the just studied list of items. Importantly, probes can be from the current list, the immediately prior list, a more distant list, or a stimulus that has never been presented. Participants with lower WMC show more interference when probes from an immediately preceding list are presented than their higher WMC peers (c.f., Jonides & Nee, 2006). Commonly, during this task, and more importantly, when the recent list probe appears (the interference), activation of the left inferior frontal gyrus is observed (Jonides & Nee, 2006).

Based on the fact that those with lower WMC have shown more difficulty with being able to properly inhibit interfering information from LTM, as was shown by Kane and Engle (2000), the following seems plausible with respect to recognizing words in the
presence of subsequent biasing information when the listener is distracted. For an individual with lower WMC, it would be expected that when hearing *The *wing had an exquisite set of feathers*, as *wing* becomes available, *wing* would be activated and stored in working memory. If later biasing information becomes available that is biased in favor of a different word (e.g., *The wing had an exquisite set of diamonds*), the word *diamonds* should cause *ring* to become activated in LTM. Because this individual's limited attentional capacity is already being compromised by the need to filter out the distracter, the ability to properly inhibit *ring* should be reduced. This, in turn, may cause the listener to have difficulty in identifying the intended word.

**Focus of the Current Study**

The overarching purpose of this project was to provide a first-pass exploration of whether individual differences in WMC are associated with the ability to recognize spoken words. More specifically, this work examined whether individual differences in WMC are associated with the likelihood that listeners will become more dependent upon distant subsequent biasing information than on signal-based information when they are distracted. A series of five experiments was performed in order to explore this issue. Experiments 1-3 explored the degree to which listeners became more dependent upon biasing information and the mechanisms (e.g., a failing inhibitory mechanism) that are driving the increased dependence on this information. Experiments 4 and 5 asked the question of whether the type of distraction mattered. In other words, it may be that only when the distraction is highly similar to the target talker's signal will it be able to influence word identification for those with lower WMC. As long as the two signals are relatively different from one another, those with lower WMC may be adequately able to
attend to the signal and thus not depend highly on biasing information. Combined these five experiments will provide a first-pass glimpse into the ways in which executive attention can influence spoken word recognition and the mechanisms involved in this process.
Chapter 2: Experiment 1

The purpose of Experiment 1 was to explore whether individual differences in WMC are associated with the degree to which listeners shift their dependence toward biasing information during distraction. Specifically, this experiment attempted to explore whether those with lower WMC would be more likely to depend upon distant subsequent biasing information than their higher WMC peers when they were distracted. Operationally, this experiment tested the question of whether individual differences in WMC would be correlated with whether subsequent biasing information that is far in proximity from a word that is being identified (e.g., *The wing had an exquisite set of feathers*, where the underlined portions represent the target and biasing words) will be depended upon during word identification when a secondary talker is present.

In order to examine the influence of a distant biasing word on word identification in the presence of a secondary talker, the current methodology was based on the study by Szostak and Pitt (2012, also see Samuel, 1990) described in Chapter 1. In their study, they presented participants with sentences containing a target word such as *wing* that had at least one rhyme competitor (words sharing all but the target word's onset phoneme: e.g., *ing*). Each target word was placed in a sentence with a biasing word that became available several syllables after the target word had been presented. The biasing word was either biased toward the target word (congruent condition: e.g., *The wing had an exquisite set of feathers*) or a competitor of that word (incongruent condition: e.g., *The
wing had an exquisite set of diamonds). In one of their stimulus versions, a noise was added to the phoneme (e.g., The \textit{wing had an exquisite set of feathers}) and in the other it was replaced by the phoneme (e.g., The \textit{=ing had an exquisite set of feathers}).

Because the present experiment had the interest of exploring whether available acoustic information would be attended to by various listeners, only the phoneme present (e.g., The \textit{wing had an exquisite set of feathers}) stimulus versions were used. To explore the influence of a secondary talker on word identification in noise, a secondary talker was presented simultaneously with the target talker. The two sentences were presented dichotically. Participants made both a target word identification and a sentence sensibility judgment for each sentence spoken by the target talker.

The congruent condition in the dichotic listening task described above provides a situation in which target word accuracy and biased response reports can not be teased apart. For example, in the sentence \textit{The wing had an exquisite set of feathers}; wing could both be an accurate response based on acoustic-signal information and a response that was biased by feathers. The reason for including the congruent condition in this experiment, therefore, was to provide sentences that would yield an affirmative sentence sensibility judgment. Based on the fact that there is a confound in the congruent condition, the primary condition of interest is the incongruent condition.

If individual differences in WMC can predict whether a distant biasing word will be depended upon during word identification in the presence of a secondary talker, as WMC decreases, if the target and biasing words are incongruent with one another, the proportion of biased word responses should increase (e.g., reporting \textit{ring} when hearing the sentence \textit{The wing had an exquisite set of diamonds}). If such an increase in biased
responses occurs, it would also be reasonable to assume that target word accuracy (e.g., reporting *wing* in the above example sentence) will decrease. In other words, a negative correlation should exist between WMC and biased word responses and a positive correlation should exist between WMC and target word accuracy. Because this is the incongruent condition, greater bias would lead to less accuracy since the biasing word would be semantically associated with an alternative competitor word (e.g., *ring*). Combined these findings would suggest that when distracted by a secondary talker, those with lower WMC depend to a greater degree than their higher WMC peers on biasing information when it is far in proximity from the word being identified.

**Method**

**Participants.** Participants were 38 individuals receiving course credit for an introductory psychology course in exchange for their participation. All were native speakers of American English with reported normal hearing. All participants also reported being predominantly right-handed.

**Stimuli. Target talker stimuli.** The target talker stimuli were those used by Szostak and Pitt (2012). Twenty-four target words were selected. Each target word was a monosyllabic noun with at least one rhyme competitor (words sharing all but their onset phonemes: e.g., *wing, ring*..., mean number of rhyme competitors = 8.4). All of the target words began with a glide, liquid, or nasal onset (c.f., Samuel, 1981). The rhyme competitor that was semantically associated with the incongruent biasing word (see below) was also required to be a noun. However, the competitor word was not held to the onset phoneme restriction described above and this word was never presented to participants.
A pair of sentences was constructed for each target word. The final word in one member of the pair was biased toward the target word (congruent condition: e.g., *The wing had an exquisite set of feathers*). The final word of the other member of the sentence pair was biased toward the competitor word and was therefore incongruently associated with the target word (incongruent condition: e.g., *The wing had an exquisite set of diamonds* where *ring* would be semantically associated with *diamonds*). The target word was always the second word in the sentence and there were always six to eight intervening syllables between the target and biasing word (e.g., *The wing had an exquisite set of feathers*). This number of intervening syllables was selected for two reasons. (1) The majority of prior studies exploring distant subsequent biasing information influence on word identification have reported use of six to eight intervening syllables (i.e., Connine et al., 1991; Samuel, 1990; Szostak, 2009; Szostak & Pitt, 2012). (2) This amount is argued to fall just beyond the amount of information that is able to be maintained in the working memory of an individual with high WMC who is being distracted (c.f., Conway et al., 2005; Cowan, 2010).

The placement of the target word was based on work from Warren and Sherman (1974) who found that when asked to identify where a masked phoneme was located in a sentence, their participants regularly displaced its location by approximately three phonemes. By presenting the target phoneme in a consistent early sentence/word position throughout the experiment, it was possible to direct participants' attention to the correct sentence location and thus ensure that the masked word would be identified as the target word rather than another word that was close in proximity to the target word.
The biasing words for each sentence pair were always from the same part of speech, and whenever possible, were equal in their number of syllables. Findings reported by Szostak and Pitt (2012) indicated both that (1) the intervening syllables (e.g., had an exquisite set of) were not biased in favor of either the target or competitor word and (2) the biasing word was properly biased in favor of the intended word (e.g., diamonds was more biased toward ring than toward wing). A full list of the target talker's sentences can be found in Appendix A.

The stimuli were recorded by the present author, a female native speaker of American English with an upper Central Ohio accent. The stimuli were recorded in a sound-dampened booth using high-quality recording equipment at 44 kHz. They were then down-sampled to 22.050 kHz. For each sentence, a clear, naturally produced token was selected for use in the experiment. The final word from the sentence was spliced out of the sentence and discarded. Two copies of the remainder of the sentence (The wing had an exquisite set of) were then saved as separate wave files. The congruent biasing word from another token of this same sentence was then spliced onto one copy of the remaining sentence fragment (e.g., The wing had an exquisite set of + feathers) and an incongruent biasing word from another token of the sentence was spliced onto the other copy (e.g., The wing had an exquisite set of + diamonds). This cross-splicing method ensured that the entire phrase preceding the biasing word was identical across conditions.

The onset phoneme of the target word was then located both auditorily and visually using a spectrogram. The located phoneme included all residual evidence of the phoneme (e.g., any formant transitions into the preceding and following phonemes). This method was used to help ensure that no residual evidence of the target phoneme (e.g., formant
transitions) was available in any part of the preceding or following portions of the sentence. Once this phoneme was located, noise was added to the signal using the method described by Samuel (1987). The noise-adding methodology reported by Samuel (1987) was chosen as it allowed the noise to have the same amplitude envelope (shifts in amplitude over time) as the masked phoneme. This methodology had the added benefit of allowing the noise to have a syllable-like quality (c.f., Samuel, 1987). To ensure the noise-added phoneme was sufficiently difficult to identify, a signal-to-noise ratio of .15 was selected. This was accomplished by randomly flipping the polarity of 85% of the samples within the target phoneme. Each noise-added sentence was then stored as a separate wave file for later presentation in the experiment.

**Secondary talker stimuli.** The speech from twelve female talkers from the Buckeye Corpus of Conversational Speech (Pitt, Johnson, Hume, Kiesling, & Raymond, 2005) served as the stimuli for the secondary (right ear) talkers. This is a time-stamped, transcribed, and phonetically marked corpus comprised of 40 individual talkers from Central Ohio who participated in interviews about everyday topics and current events. Four spoken sentences from each of the 12 talkers (total number of sentences = 48) were selected. The sentences had several constraints placed on them. (1) None of the target words, any of their competitors (whenever possible), the biasing words, or any proper nouns were present in any of the 48 sentences. In the single case where a competitor word was present in the secondary sentence, the researcher ensured that the sentence was not paired with the target talker's sentences that contained the same rhyme. (2) Words or controversial topics likely to elicit personal attitudes or strong emotions (e.g., gun control, violence, partisan politics) were not permitted to be in the sentences. (3) The
sentences could have no noticeable pauses, extraneous noises, or speech from the interviewer present within the selected portion of the recording. (4) Whenever it was possible, the secondary talker's speech was devoid of substantial inflectional changes (e.g., becoming excited, major changes in pitch or amplitude). (5) To the degree that it was possible, all selected items were full sentences. (6) All sentences were equal to or no more than 250-750 msec longer than the paired target talker's sentence (see below). A full set of the secondary talkers' sentences can be found in Appendix B.

**Dichotic listening sentence pairings.** Each target talker sentence was paired with one of the sentences from one of the secondary talkers (see Appendix B for details). The two members of the sentence pair from the target talker were not permitted to be combined with sentences from the same secondary talker (e.g., if the sentence, *The wing had an exquisite set of feathers*, was paired with talker a, then the sentence, *The wing had an exquisite set of diamonds*, could not be paired with this same talker).

The target talker's sentence was always presented to the left ear and the secondary talker's sentence was always presented to the right ear. This ear-presentation choice was made owing to findings which suggest that for the majority of right-handed individuals, there is a right ear advantage when listening to speech. This is based on the theory that the language center is maintained primarily in the left hemisphere of the brain (c.f., Foundas, Corey, Hurley, & Heilman, 2006). Essentially, the idea here is that information that enters the right ear has a direct path to the language center in the brain for most right-handed individuals. In contrast, information that enters the left ear immediately goes to the right hemisphere and must cross the corpus callosum in order to reach the language center (Foundas et al., 2006).
To ensure that the target talker's speech began later than the secondary talker's speech, a 250 msec silent pause was prepended to each sentence spoken by the target talker. The primary reason for the decision to have the secondary talker's speech begin before the target talker's speech was to ensure that the target word was not presented immediately after the trial began. Further, this ensured that the task would be made more difficult so as to bring out WMC differences if such differences exist. The target and secondary talker were combined dichotically. The target talker's speech was always placed in the left channel and the secondary talker's speech was always placed in the right channel. To ensure that the secondary talker's speech always ended at the same time or later than the target talker's speech, the secondary talker's entire sentence ranged in length from 0-500 msec longer than the target talker's sentence (total extra duration = 250-750 msec: e.g., 250 msec prepended pause + target sentence duration + 0-500 msec). This ensured that the length of the sentence ended no more than 500 msec after the offset of the target talker's sentence.

**WMC stimuli.** A measure of WMC was obtained using an instrument known as the Operation Span³ (OSpan) task (c.f., Conway et al., 2005). This measure was selected because it has been used extensively in the WMC literature (e.g., Amir & Bomyea, 2011; Aslan & Bäuml, 2011; Colflesh & Conway, 2007; DeCaro et al., 2008; Kane & Engle, 2000; Kane et al., 2004; Redick et al., 2012)and has been shown to have both strong reliability and validity (c.f., Conway et al., 2005).

Forty-two relatively simple math equations having the format $9 \div 3 - 2$ were used during this portion of the experiment. Each had a correct or incorrect solution and was presented in the form of a question (e.g., *Is $9 \div 3 - 2 = 1$?*, *Is $9 \div 3 - 2 = 5$?*) Each math
equation was paired with a unique monosyllabic word (e.g., Is 9 ÷ 3 - 2 = 1?, dog). Each of the 42 equation/word pairings was presented on its own screen. It was centered and was presented in size 24 in font. The word was presented in upper case type (e.g., DOG).

For the item recall portion of the OSpan task (see below), answer sheets were constructed which contained 12 blank lines.

Procedure. OSpan task. The OSpan task requires participants to learn lists of words while simultaneously solving math problems. After each list of words has been presented, participants are required to recall the full word list in serial order. An individual's WMC is based on the amount of correctly recalled words across lists.

Participants participated in groups of one to two in separate sound-attenuated rooms. They were seated in front of a computer monitor with an answer sheet placed in front of them. An experimenter was seated next to the participant with a computer keyboard in front of him or her. At study, the participant was presented with a math problem with a word printed below it (e.g., Is 9 ÷ 3 - 2 = 1? DOG). During each study trial, participants were required to read the entire equation aloud, respond "yes" if the solution was correct or "no" if it was incorrect, and then read aloud the word. For example, if seeing the stimulus item: Is 9 ÷ 3 - 2 = 1? DOG, an accurate response would be "Is nine divided by three minus two equal to one, yes, dog". Participants were instructed to perform the above task as quickly and as accurately as possible. They were told that they should try not to pause or stumble and should begin reading as soon as the equation/word appeared on the computer screen. Immediately following the participant's full response, the experimenter advanced the computer to the next trial by pressing the space bar. After pressing the space bar, the experimenter recorded whether the participant responded
"yes" or "no" either on a response sheet or on a hand-held computer. Each list of words ranged from two to five items in length based on recommendations by Conway et al. (2005).

Following each studied list, the computer prompted participants to write down the entire list of words. Participants were told to write all of the words they could recall in serial order and were instructed to provide a space marker (e.g., -) to indicate the positions of any words they could not recall. The participants were instructed to write the list of words on one of the 12 lines on the answer sheet in front of them and to move onto the next line for the next studied list of words. Once the participant indicated that they had completed recalling the list of words, the experimenter pressed the space bar to advance to the next trial. The order of lists and items within each list was presented in a fixed pseudorandomized order.

Before beginning the OSpan task, participants were given a set of three practice lists to become comfortable with performing the task. During the practice session, the experimenter provided feedback as needed regarding the participant's pace. Following the OSpan task the participants were offered a short break. They then completed the dichotic listening task described below.

**Dichotic listening task.** During this task, participants sat alone in the same sound-dampened room where the OSpan task had been performed. The participants heard the stimuli during this portion of the experiment over headphones. During each trial, the participants heard one of the 48 target talker's sentences in the left ear and a sentence from one of the secondary talkers in the right ear. The stimuli were presented in two
mixed blocks. Each block contained 24 of the sentences. No target words were repeated within a block. The items were presented in a fixed pseudorandomized order.

On each trial, the participant would hear one stimulus item. While the two talkers were speaking, a beige-colored dialog box would appear on the computer screen. This dialog box would remain beige until both talkers had finished speaking. After this point, the dialog box would turn white. Participants were told that this color change was a cue that would let them know when they could begin typing. The box would remain white, with no forced time limit, until the participant was done performing both of the below tasks. Any responses made before the dialog box turned from beige to white were not recognized by the computer. After the box turned white, participants could see what they were typing inside the box. While the dialog box was white, the participants made two responses for the stimulus item that had just been presented. (1) They reported the target word. The participants were informed that this word would always be the second word in the sentence and that they were to listen closely and type the actual word they heard, not what they thought the talker was trying to say based on the rest of the sentence. It was emphasized that they should attend closely to the information in the signal. This method was used in order to fall in line with the instructions reported by Mattys et al. (2010) and so as to provide a stronger test of biasing word influence. Following this response, participants pressed the space bar. (2) They then made a sentence sensibility judgment by pressing the "s" key on the keyboard if the entire sentence made sense or the "n" key if it was a nonsense sentence. This secondary task was used for two reasons. (a) It ensured that listeners were attending to the whole sentence and (b) it forced them to integrate the target word into the sentence rather than just focusing on it and ignoring the latter portion
of the sentence so as to actively rehearse this word. After performing both tasks, the participant pressed the "Enter" key, the dialog box turned from white to beige, and the computer immediately moved on to the next trial. During each trial, participants were told that they would hear an additional talker in their right ear, and that they should ignore this talker as they would be making no responses based on the right talker's speech.

Participants were offered a second break halfway through the dichotic listening task. Prior to performing this task, participants were given several practice trials to become comfortable with performing the task. Following this task, they were thanked for their time and debriefed. The OSpan and dichotic listening tasks combined lasted approximately 45 minutes.

Results and Discussion

Participants who scored less than 85% correct on the math portion of the OSpan task were excluded from all data analyses. This decision was made based on recommendations from Conway et al. (2005). For all participants who achieved at least 85% accuracy on the math portion of the OSpan Task, the data were analyzed as follows. For the OSpan task, WMC was computed based on partial-credit unit scoring (c.f., Conway et al., 2005). This is a measure that is based on the mean proportion of the number of words that is recalled correctly in serial order across studied lists. For example, if a participant correctly recalled two words on a list containing two words, one word on a list containing three words, and two words on a list containing four words, the resulting capacity would be .61 because the mean of 1.0, .33, and .5 is .61^5.
For the word identification task, two measures were obtained. (a) A measure of the proportion of trials per condition (congruent, incongruent) on which the participant reported the intended target word (target word accuracy: e.g., reporting "wing" for *The wing had an exquisite set of diamonds*) was calculated. (b) A measure of biasing word influence (biased word responses) was obtained. For this measure, any response that was semantically associated with the biasing word was counted. The data were coded as a "1" if the response was semantically associated with the biasing word or "0" for all other responses. For example, for the sentence: *The wing had an exquisite set of diamonds*, responses of "ring", "bird", and so on were counted as semantically biased. Note that pseudoword responses which made up less than 3% of the data were excluded from the analysis.

Although the congruent condition will not be the focus given that responses based on acoustic information and biasing word influences can not be teased apart, a brief summary will be provided for the interested reader. The data from the congruent condition are provided in Figures 1a (proportion target word accuracy as a function of WMC in the congruent condition) and 1b (proportion biased word responses as a function of WMC in the congruent condition). In Figure 1a, it is evident that the task was relatively difficult (mean proportion target word accuracy = .68). Although the average is low, there is little to no evidence in the figure that those with lower WMC are dragging the mean down. A Pearson correlation confirmed this finding (*r* = .03, *p* > .87. For the biased word responses, a similar trend can be seen in Figure 1b. Again, the mean is only moderately high (*m* = .72). The slightly elevated biased word responses over that of the target word accuracy responses is the result of participants occasionally reporting words
that were associated with the biasing word but did not share the rhyme (e.g., reporting "bird" when hearing the sentence *The wing had an exquisite set of feathers*). As with the target word accuracy responses, correlational analyses of biased word responses across WMC suggested that those with lower WMC were not reporting words associated with the biased word reliably more frequently than their higher WMC peers \((r = .08, p > .63)\). Thus, as expected, no differences were observed across WMC in the congruent condition. Given this outcome, only the incongruent condition will be discussed further in this and the subsequent experiments.

The first question that needs to be addressed is: How well were listeners with differing degrees of WMC able to accurately identify the target word. Figure 2a contains proportion target word accuracy as a function of WMC in the incongruent condition. There is a moderate positive trend that can be seen in the figure. Those with lower WMC had more difficulty in being able to identify the target word (e.g., *wing*) accurately than was the case for those with higher WMC.

Although target word accuracy provides an important piece of the story, perhaps the more valuable question is whether the likelihood of reporting the word associated with the biasing word (e.g., *ring*) altered as a function of WMC. Figure 2b shows the proportion of biased word responses as a function of WMC in the incongruent condition. In this figure, the trend is clearly moving in the opposite direction from that of Figure 2a, and the steepness of the trend is slightly larger than that for target word accuracy. The data indicate that the likelihood of reporting a word that was semantically associated with the biasing word (e.g., *ring*) increased as WMC decreased.
Pearson correlations were performed on WMC both with proportion target word accuracy and proportion biased word responses. These analyses indicated that there was a marginally reliable positive correlation between WMC and target word accuracy ($r = .29$, $p < .08$. Moreover, there was a reliable negative correlation between WMC and biased word responses, $r = -.36$, $p < .05$.

The findings from experiment 1 suggest that during distraction, as WMC decreased, listeners appeared to become less dependent upon the acoustic-phonetic information in the signal and tended to become more dependent upon the biasing word in order to aid word identification. Those with higher WMC appeared to depend to a greater degree on the acoustic-phonetic information and less upon the biasing word when identifying the target word. Thus, for those with lower WMC, when the signal is unclear, they are more likely to depend upon information that is more salient regardless of whether this will result in increased word identification accuracy while those with higher WMC are more likely to depend on information (i.e., signal-level) that is most likely to result in accurate word identification. Importantly, the findings suggest that executive attention is necessary for accurate word recognition to occur when a listener is distracted. For those with lower WMC, because of the reduced attention, the remaining executive attention appears to be reallocated to the biasing word, possibly as a way of reducing the amount of resources that must be used to attend to the signal while also trying to ignore the distracting speech.
Chapter 3: Experiment 2

One question that Experiment 1 can not answer is: What caused those with lower WMC to show a greater dependence upon biasing than on signal information? There are two possibilities. The first possible explanation, as was discussed in Chapter 1, is that of an inhibitory deficit (e.g., Kane & Engle, 2000). If WMC has been reached, there will not be enough attentional capacity available for the processor to be able to properly inhibit the level of activation of a semantic associate of the biasing word (e.g., activation of ring when hearing diamonds). Recall that Kane and Engle (2000) found that those with lower WMC were more likely to produce intrusions from a prior same-category list than those with higher WMC when under no distraction. When placed under a distraction, their lower and higher WMC participants showed relatively similar degrees of producing intrusions. The idea is that when attentional capacity is reached, the ability to properly inhibit information that has been activated in LTM will be hindered. As a result, the activated information from LTM will influence what is being maintained in working memory (e.g., the target word).

In Experiment 1 it was assumed that the listener's lexicon would have established a semantic connection between the biasing word and either the target (congruent condition) or competitor (incongruent condition) word (c.f., Collins & Loftus, 1975 for one illustration of this idea). Because of this connection, if an inhibitory deficit explanation is correct, the following would be expected to take place. When hearing a sentence such as
The wing had an exquisite set of diamonds, the word diamonds should cause ring to be activated owing to the semantic connection between these two words. If the listener's attentional capacity has been reached, the ability to inhibit the activation level of ring should be decreased. In such a situation, the listener would become more likely to report hearing ring than wing even if wing had originally been activated based on the acoustic-phonetic information in the signal.

The second possibility is that a compensatory mechanism is at work. If WMC has been reached, listeners may no longer have the attentional capacity necessary to properly attend to the fine-grained acoustic-phonetic information in the signal (e.g., the ability to discriminate information such as formant transitions or subtle changes in amplitude). Because of this, once WMC has been reached, listeners shift their attention to the more salient information in the signal (e.g., the biasing word). As long as the biasing information is congruently associated with the word being identified, such a compensatory mechanism would likely enhance word identification and would possibly help decrease the demands placed on the listener's attention.

There is no way to know which of the two hypotheses (inhibitory deficit or compensatory mechanism) caused those with lower WMC to depend more heavily on the biasing word more so than those with higher WMC in Experiment 1. The purpose of Experiment 2 was therefore to determine whether the data patterns observed for those with lower WMC in Experiment 1 were the result of an inhibitory deficit or a compensatory mechanism. In order to test this, Experiment 2 was identical to Experiment 1 with the exception that the biasing word was removed from the sentence (absent condition⁶: e.g., The wing had an exquisite set of). If an inhibitory deficit hypothesis is
correct, then those with lower WMC should show similarly high accuracy to those with higher WMC. Such an outcome would suggest that those with lower WMC are as equally capable of attending to the signal as their higher WMC peers, but that they simply were unable to inhibit the biasing word’s influence in Experiment 1. If in contrast the compensatory mechanism hypothesis is correct, those with higher WMC should again prove more accurate at identifying the target word than those with lower WMC as those with lower WMC would not have the ability to properly attend to the signal.

Since the absent condition has no biasing word, focus in this experiment will be on the target word accuracy measure alone. If listeners with lower WMC are showing an increased dependence on the biasing word owing to an inability to properly inhibit the influence of this information as a result of reaching attentional capacity, then no differences should emerge across WMC, and all participants should show high accuracy. This would suggest that once the biasing information is no longer present, the acoustic-phonetic information is able to be processed adequately since there is no need to inhibit any biasing information. If, in contrast, listeners with lower WMC are compensating for an inadequate ability to attend to the signal owing to their attentional capacity reaching its limit, a positive correlation between WMC and target word accuracy should emerge.

If the compensatory mechanism hypothesis is correct, it brings up an interesting question. For individuals with lower WMC, to what degree, if any, did the processor encode the fine-grained acoustic-phonetic features (e.g., formant transitions) that were present in the target phoneme (e.g., /w/ in wing)? There are two possible hypotheses. (1) It may be that those with lower WMC are able to process relatively salient acoustic information (e.g., they can easily determine that a phoneme is a liquid) but can not
distinguish more fine-grained acoustic information such as discriminating between the \(F_3\) trajectories of /l/ and /r/. (2) It may be that those with lower WMC may only have the attentional capacity to process the signal at a very shallow (e.g., attending to only the gross-acoustic features) level. In such a case, if a listener with low WMC hears something like \(\text{wing}\), the individual may not be attending to the signal at a level that allows the processor to distinguish between the noise-masker and any discriminable acoustic-phonetic information present in the signal. In such a situation, listeners with lower WMC would likely report hearing a word such as \(\text{sing}\) or \(\text{thing}\) given that the noise-masker in Experiment 1 had a fricative-like quality. The advantage of this very gross level of processing would be that relatively limited executive control would likely be needed to help identify the acoustic-phonetic information present in the signal.

If those with lower WMC are compensating for a signal-encoding deficit, it is possible that those with lower WMC will produce more words that have fricative onsets (e.g., \(\text{thing, sing}\)) than will be observed for those with higher WMC. This would suggest that as WMC is decreased, less depth in signal-processing occurs.

**Method**

**Participants.** Thirty-eight new participants from the same pool meeting the same criteria as those in Experiment 1 participated for course credit.

**Stimuli.** The stimuli were modified versions of the 24 incongruent sentences (e.g., *The wing had an exquisite set of diamonds*) used in Experiment 1. The incongruent condition stimuli were chosen to be modified simply because the incongruent condition had been the condition that was analyzed in Experiment 1.
To create the absent stimulus versions, the final (biasing) word from the sentence was removed (e.g., *The wing had an exquisite set of*). A new secondary talker sentence was then paired with the sentence fragment. The new secondary talker sentences were from the same 12 secondary talkers as had been used in Experiment 1. Use of these new sentences allowed for more generalization across experiments. In other words, if the same effect was observed in this Experiment as had been observed in Experiment 1, it would suggest that the effects were not specific to the particular sentences spoken by the secondary talkers. The same criteria were used when selecting these sentences as was described in Experiment 1. Appendix B contains the stimuli from the secondary talkers.

As with the first experiment, the sentences from the two talkers (target and secondary) were combined dichotically. The target talker was again presented in the left channel and the secondary talker was again presented in the right channel.

In order to ensure that the duration between target word onset and the time at which participants were allowed to make their responses was maintained, the duration of the secondary talker's sentence was nearly identical (max difference = 20 msec) to that of Experiment 1. In other words, the duration of the secondary talker's speech was nearly identical to that of the same incongruent item's secondary talker sentence in Experiment 1. All other aspects of stimulus construction were identical to the first experiment.

**Procedure.** The procedure was identical to that of Experiment 1 with one minor exception. Before beginning the dichotic listening task, the participants were told that they would be hearing incomplete sentences during the experiment and were told that their sensibility judgments would be made based on these incomplete sentences. The entire experiment lasted approximately 30 minutes.
Results and Discussion

Primary analysis. For all participants who achieved at least 85% accuracy on the OSpan, all the data were analyzed identically to that of Experiment 1 with one exception. Because there was no biasing word, only proportion target word accuracy was measured. Biased word responses were not calculated.

Figure 3a provides proportion target word accuracy as a function of WMC. The figure shows a positive-going trend, suggesting that those with lower WMC were less accurate at reporting the target word than their higher WMC counterparts. A Pearson correlation indicated that the trend was marginally reliable, $r = .32, p < .06$. Based on the observed positive trend, the findings suggest that those with lower WMC were having more difficulty than those with higher WMC at identifying the intended word.

Phoneme analysis. One question that remains is: What is the reason for the difficulty that those with lower WMC were experiencing in their target word identification? Recall that earlier, two hypotheses were discussed. (1) It is possible that those with lower WMC were properly distinguishing the signal from the noise but were not able to attend to the very fine-grained features coded in the acoustic signal. (2) It may be that those with lower WMC were attending to the signal at a very shallow level; a level that was too shallow to be able to properly distinguish which features were part of the acoustic-phonetic signal and which features were part of the noise.

To test the above question, the error data were analyzed. The errors were re-coded such that all fricative onsets were scored as a one (e.g., reporting $\text{wing}$ as $\text{sing}$) and all other errors were scored as zeros (e.g., reporting $\text{wing}$ as $\text{ring}$). If those with lower WMC were not properly distinguishing the signal from the noise, then as WMC decreases, the
proportion of fricative responses will increase (e.g., a negative correlation between WMC and frication reports). In contrast, if those with lower WMC were simply making more errors than their higher WMC peers as a result of not being able to distinguish highly confusable acoustic features (e.g., distinguishing between the F3 trajectories of /l/ and /r/), no differences in the rate of reporting frication should emerge across WMC.

Figure 3b contains the proportion of fricative responses as a function of WMC. The results showed a negative-going trend. The size of this trend was moderate. Its trajectory suggests that those with lower WMC were more likely to report a fricative when giving an inaccurate response than were those with higher WMC. A correlation analysis between WMC and fricative responses proved reliable, $r = -.39, p < .05$. The findings from this analysis suggest that those with lower WMC had a higher probability of reporting frication than their higher WMC peers.

To further support the fricative analysis, an analysis of other phoneme types was also performed. Figures 4 (stop responses as a function of WMC), 5 (semivowel responses as a function of WMC), and 6 (nasal responses as a function of WMC) provide these data. The proportion scores were computed in an identical manner to that of the fricative responses, with the single exception that the error data were coded based on the particular phoneme class of interest. In the stop and semivowel cases, there was a slight positive trend suggesting that as WMC increased, the proportion of stop and semivowel responses increased; a direction that is opposite from that of the fricatives. In contrast, there seems to have been little to know trend emerging for the nasal responses. For semivowels the positive trend makes perfect sense given that both (a) Samuel (1981) found semivowels to be the most accurately able to be identified in noise and (b) many of the target words
began with semivowel onsets. For stops, one possibility is that the noise may have for some reason seemed somewhat short in its duration for participants and thus they may have made more stop responses based on this fact (c.f., Miller & Wayland, 1993; Sherman, 1971). For the nasals, the lack of an effect makes sense given that Samuel (1981) found that nasals were less adequately able to be identified in noise than was the case for semivowels. Thus, listeners may have not been able to recognize these phonemes. Correlations between WMC and stop responses \( (r = .31, p < .06) \) and between WMC and semivowel responses \( (r = .27, p < .11) \) proved marginally reliable. The correlation between WMC and nasal responses proved unreliable \( (r < .11) \).

**Summary.** Combined, the phoneme (fricatives, nasals, semivowels, stops) response and target word accuracy data suggest that those with lower WMC were less adequately able to encode the signal than those with higher WMC. Moreover, the findings suggest that this was likely the result of those with lower WMC processing the signal at a level that was too shallow to properly identify crucial information for making an accurate phoneme and thus word identification. The limited amount of attention available to these individuals appears to cause them to not be able to control the level of depth to which they are attending to the signal; an issue that will be addressed further in Chapter 7.
Chapter 4: Experiment 3

The findings from Experiments 1 and 2 clearly demonstrate that an increase in dependence upon the biasing word is related in part to WMC. One remaining question is whether those with lower WMC are always impaired in their ability to process the signal when distracted, or do they only show an impairment when the signal is sufficiently degraded. Experiment 3 provides a control condition to demonstrate that as long as the signal is clear (e.g., hearing *The wing had an exquisite set of diamonds*), those with lower WMC show no impairment in their ability to identify words and are able to do so with equal accuracy to that of their higher WMC peers.

It is likely that those with lower WMC will perform as well as their higher WMC peers as long as the signal is clear. Two pieces of evidence provide support for this hypothesis. First, those with lower WMC in Experiment 1 showed relatively strong likelihoods of reporting a word that was semantically associated with the incongruent biasing word. The only way that these individuals could have done so, is if they had been able to properly process the clear portion of the signal (e.g., *diamonds*). The second piece of evidence comes from the findings reported by Conway et al. (2001). Recall that Conway et al. (2001) presented participants with competing passages; one of which they were instructed to attend to and one of which they were instructed to ignore. The ignored passage contained the listener's name. Although their findings showed that those with lower WMC were more impaired when hearing their own name than those with higher
WMC, little to no differences were observed in their ability to attend to other parts of the
target passage. Had their participants been unable to process the clear signal in general,
then those with lower WMC should have shown a decrement in their performance
throughout the entire passage compared with those having higher WMC.

Method

Participants. Thirty-eight new participants from the same pool meeting the same
criteria as those of Experiment 1 received course credit for their participation.

Stimuli. The stimuli were identical to those of Experiment 1 with one exception. The
target phoneme was presented as clear speech (e.g., The wing had an exquisite set of
feathers or The wing had an exquisite set of diamonds). All other aspects of stimulus
construction were identical to that of Experiment 1.

Procedure. The procedure was identical to that of Experiment 1.

Results and Discussion

The OSpan and dichotic listening tasks were analyzed in an identical fashion to that of
Experiment 1. The proportion target word accuracy and biased word response data as a
function of WMC from the incongruent condition can be found in Figures 7a and 7b
respectively. Neither target word accuracy ($r = .09, p > .55$) nor biased word responses ($r
= -.11, p = .50$) showed any reliable correlation with WMC.

The data suggest that as long as the signal was clear, all participants, regardless of
WMC, showed greater reliance on acoustic signal information than on the biasing word
to aid word identification. Across WMC, participants were near ceiling in their
performance (mean target word accuracy = .94). Moreover, the biasing word did little to
influence performance across WMC (mean biased word responses = .04).
The target word accuracy and biased word response findings from this experiment suggest that as long as the target word was clearly perceivable, regardless of WMC, listeners relied most heavily on signal based information and almost never relied on the incongruent biasing word to aid word identification. Even the individual with the lowest WMC (WMC = .30) was almost 90% accurate (target word accuracy = .88). Thus, clearly the biasing word did not have an intractable influence on performance. These findings also fit nicely in line with those of Mattys et al. (2010). Recall that their findings, without exploration of WMC, showed that listeners were more likely to rely on signal- than biasing word-based information when the speech was unambiguous, even in the presence of a secondary talker.

An alternative explanation is that listeners were simply ignoring everything in the sentence beyond the second word. This strategy would result in the exact same data patterns that were reported above. One way to test this question is to examine the sentence sensibility judgments that participants made. If listeners were engaging in such a strategy, then their sentence sensibility judgments in the incongruent condition should be near chance. An examination of the proportion sentence sensibility judgments as a function of WMC can be found in Figure 8. The data indicate that regardless of WMC, participants performed this task relatively equivalently ($r = -.06, p > .70$). Importantly, participant performance proved to be near ceiling ($m = .87$). Thus, the findings clearly suggest that listeners were attending to and properly integrating all information within the sentence.

Combined, the findings from Experiments 1-3 provide a few important insights into the process of identifying words and the influence of executive control/attention in this
process. Those with lower WMC are more likely to depend upon the biasing word than on acoustic information. This is likely the result of a deficit in being able to properly encode the acoustic signal. Based on the fact that listeners with lower WMC in the present experiment showed equally accurate identification of the target word, it is likely that when the signal is clear, listeners with lower WMC have enough attentional capacity to properly identify information such as the formant trajectories that are present in the signal. Thus, on some level, listeners with lower WMC are able to encode the signal; it is just that what appears to be encoded, is only the features that can be clearly and easily recognized. Given that the target talker's speech in the three experiments was articulated clearly, it is evident that as long as the signal is not masked by noise, those with lower WMC can encode meaningful (e.g., formant transitions) acoustic-phonetic features. Had this not been true, those with lower WMC in the current experiment should have performed similarly to those with lower WMC when the signal was unclear in Experiment 1 (e.g., showing poor target word accuracy and high biased word responses).
Chapter 5: Experiment 4

In Experiments 1-3 the question of why listeners with lower WMC were more likely to show more dependence upon biasing information than on acoustic signal information (i.e., the use of a compensatory mechanism to aid in the processing of a poorly encoded signal) was explored. These experiments centered on the direct effects of the target talker's speech. The experiments, however, placed no focus on the distraction. Given that the presence of a distraction has repeatedly been shown to cause disruptions for processing speech (e.g., Conway et al., 2001; Colflesh & Conway, 2007; Mattys and Wiget, 2011; Mattys et al., 2010 also see Styles, 1997) and that distractions in general are often more detrimental to those with lower WMC than those with higher WMC (e.g., Conway et al. 2001; Colflesh & Conway, 2007), an exploration into why the distraction in Experiments 1 and 2 compromised spoken word recognition for those with lower WMC would be of value. The purpose of Experiment 4 was to begin to probe this issue.

It is possible that the signal-encoding impairment experienced by those with lower WMC in the first three experiments was the result of target/distracter stimulus similarity. In other words, it may be that if the distracter had not been speech or at least speech from a highly different talker (e.g., a male secondary talker with a female target talker), those with lower WMC might have maintained enough of their executive attention to be able to properly encode the acoustic information present in the signal. Such a finding might suggest that as long as a distracter is different enough in some way from the target
talker's speech signal (e.g., a nonspeech distraction), less attention is needed in order to
distinguish the distracter from the target talker's voice and therefore properly filter the
distracting stimulus out. In other words, two talkers, especially of the same gender, are
likely to produce signals that are somewhat similar (e.g. they may have similar
fundamental frequencies, they may be saying similar sounding words, and so on). A non-
speech stimulus in contrast may have a very different fundamental frequency, its
amplitude envelope may be quite different, and so on. When nonspeech stimuli are
presented, if they differ in important ways from speech, it would thus be logical to
assume that a nonspeech distracter may be more easily filtered out and thus may require
less executive attention. If less executive attention is needed to filter out nonspeech
distractions, it may be that regardless of WMC, listeners will be more easily able to
identify what the target talker is saying and may show an improvement in signal-
processing and a decrease in the dependence upon the biasing information.

Some, though indirect, evidence favoring this hypothesis comes from work that has
suggested that deficits in word identification may be to some degree impacted by the
specific type of distracter. For example, Toro, Sinnett, and Soto-Faraco (2005) tested
whether listeners would learn an artificial language differently dependent upon what type
of distracter was presented. Participants were asked to listen to an artificial language at
study and then were asked to make word/pseudoword judgments about the language at
test. Importantly, they presented various types of distracters while listeners were learning
the language. Two of their conditions are of most relevance to the present study. In one
condition, an acoustic feature of the talker's speech was occasionally altered (e.g., a
change in frequency). In the other condition, nonspeech distractions were presented (e.g.,
the sound of a door slamming).

Toro et al. (2005) found that when listeners were presented with nonspeech
distractions, they showed higher accuracy in their word/pseudoword judgments than was
observed when they heard the occasional changes in spectral frequency, a mean
difference of approximately 9%. They surmised that when attentional capacity is
compromised, listeners are more likely to be impaired in recognizing words when
distracted by speech-specific distractions than by nonspeech distracters. It is therefore
possible that those with lower WMC will be less impaired by a nonspeech distracter than
they were by the presence of a speech distracter. In such a case, it would be logical to
assume that these individuals would show more dependence upon the signal than on the
biasing information during word identification.

In contrast, evidence from Mattys et al. (2010) suggests that perhaps any type of
distracter may cause similar word identification difficulties to emerge. Recall that Mattys
et al. (2010) presented participants with different types of distracters while requiring
them to attend to a target talker who was producing a phrase such as mild option or mile
doption. In their study they found that both speech and non-speech (e.g., signal correlated
noise) distracters resulted in an increased dependence upon biasing information (e.g.,
option) than was the case when no distracter was present.

To explore the issue of distracter/target talker similarity on word identification, the
same noise-masked congruent (e.g., The wing had an exquisite set of feathers) and
incongruent (e.g., The wing had an exquisite set of diamonds) sentences that were used in
Experiment 1 were again used in this experiment, but this time the distraction was caused
by everyday non-speech sounds such as traffic noises and animal sounds. If the impairment does not occur as long as the distracter is relatively dissimilar from that of the target talker signal (e.g., speech vs. nonspeech), then regardless of WMC, participants should show relatively high target word accuracy and produce relatively few biasing word responses. It is possible that the size of the accuracy and biasing word effects will prove quite similar to that of Experiment 3. Such a finding would suggest that as the two signals become less similar, word identification will require less executive control to be able to filter out and thus will allow for more attention to be able to be devoted to the acoustic signal. Alternatively, if those with lower WMC show the presence of the same type of signal encoding deficit as was seen in Experiment 2 (e.g., not being able to properly attend to the fine-grained acoustic features in the signal), even when the two signals are dissimilar, then, as WMC decreases, target word accuracy should decrease and biasing word responses should increase; a finding similar to that of Experiment 1. Such a finding would suggest that for those with lower WMC, the ability to filter out a distraction that is dissimilar from the target talker signal requires too much executive control for there to be sufficient residual attention available for processing the target talker's speech adequately.

Method

Participants. Thirty-eight new participants from the same pool meeting the same criteria as those in Experiment 1 participated for course credit.

Stimuli. The same 24 congruent (e.g., The wing had an exquisite set of feathers) and 24 incongruent (e.g., The wing had an exquisite set of diamonds) sentences that were used in Experiment 1 were again used in the present experiment.
In order to construct the nonspeech distracters, a corpus of 208 everyday sounds was created. These included sounds such as animal sounds (e.g., a dog drinking water), household noises (e.g., air conditioner, blender, phone ringing), traffic noises, zipping sounds (e.g., zipping jackets), and entertainment (e.g., refereeing whistle at a sports event or cards being shuffled). For a full list of the sound categories that comprised the corpus, see Appendix C. These types of sounds were chosen as they were somewhat similar to those reported by Toro et al. (2005).

The length of each sound ranged from approximately 100 msec to nearly 1,000 msec. If a lengthy sound needed to be shortened, it was restricted to a length where the sound source could still be identified. Although no two sounds were permitted to be identical, two portions of the same sound source were used in the corpus as long as they were deemed noticeably different by the present researcher. For example, the dial tone and subsequent tones that resulted from depressing the buttons on the key pad of a cell phone could both be used as different sound files. This criterion was permitted owing to the number of sound sources necessary for this experiment. To ensure that the sounds were non-speech like, sounds such as singing, coughing, sneezing, laughing, cheering, and clapping were excluded from the corpus.

After the corpus was created, a group of sound files from the corpus (mean number of files = 4.3, range = 3-6) were spliced together with no intervening silences. Files were paired with two goals in mind. (1) The resulting sound-chain was close to the duration of the secondary talker sentence duration it would be replacing from Experiment 1 (max difference between secondary sentence and sound-chain = 19 msec). (2) As much as
possible, the noises that a given sound-chain was composed of were relatively dissimilar from one another, though it should be noted that this goal was not always able to be met.

Once the sound-chains were constructed, their mean amplitudes were set to 70 dB in line with the secondary talker sentences. They were then combined dichotically with the target talker sentences\(^7\). As with the prior experiments, the target talker's speech was placed on the left channel and the sound-chain was placed on the right channel. The reason for choosing to use sound chains as described above, was that speech has been shown to be highly noninvariant (c.f., Klatt, 1980). Use of sound-chains ensured that a relatively dynamic distracter signal would be provided. In order to provide a more stringent test of the question under study, it was necessary to provide a relatively noninvariant distracter.

All other aspects of stimulus construction were identical to that of Experiment 1.

**Procedure.** The procedure was identical to Experiment 1 with one exception. The participants were told that they would be hearing everyday noises in their right ear and were told to ignore these sounds. The experiment lasted approximately 45 minutes.

**Results and Discussion**

**Primary analysis.** The OSpan and dichotic listening data were analyzed in an identical fashion to that of Experiment 1. The incongruent condition data are found in Figures 9a (proportion target word accuracy as a function of WMC) and 9b (proportion biased word responses as a function of WMC). An observable positive trend was found between WMC and target word accuracy. Also, a negative-going trend was seen between biasing word responses and WMC. Correlation analyses however indicated that neither
trend was reliable (target word accuracy $r = .31, p > .06$; biased word responses $r = .24, p > .13$).

The findings suggest that those with lower WMC are not showing a difference in their ability to identify the target word from that of their higher WMC peers. This leads to an important and interesting question. Is it that those with lower WMC are becoming more like their higher WMC peers (e.g., does a less speech-like distracter require less attention to filter out), or is it that those with higher WMC are becoming more like their lower WMC peers? If those with higher WMC are becoming more like their lower WMC peers, it would suggest that for some reason when listening to speech in an environment with a lot of nonspeech distractions, more attention is required to filter out the distracting stimuli.

**Signal similarity analysis.** One simple way to test the question of whether those with lower WMC are improving or whether those with higher WMC are beginning to experience an impairment similar to that of their lower WMC peers is to compare the present data to those of Experiment 1. These two experiments differ only in the fact that one had a nonspeech distracter (Experiment 4: nonspeech distracter condition) and one had a speech distracter (experiment 1: speech distracter condition). Essentially, the question of interest across these two experiments, is whether an interaction is present. It is possible that those with lower WMC will show a marked improvement in performance when the distracter becomes less like the target talker's speech signal while no such change in improvement is seen for those with higher WMC. Alternatively, it is possible that those with lower WMC will remain relatively poor in their ability to accurately
identify the target word, regardless of distracter similarity while those with higher WMC become more impaired when the distracter becomes less speech-like.

To compare the data from these two experiments, an extreme groups design was used. Redick et al. (2012) argued that the use of extreme groups designs can be highly informative as they provide a situation that allows for a more discrete comparison of low and high WMC.

The data from Experiments 1 and 4 were split into quartiles based on WMC. The upper (25% of participants with the highest WMC per experiment: high WMC condition) and lower (25% of participants with the lowest WMC per experiment: low WMC condition) most quartiles (n = 9 participants per quartile per experiment, total N = 36) were then analyzed. These quartiles were selected so as to provide a clear distinction between high and low WMC within each of the two distracter-typed conditions (speech and nonspeech). The extreme groups comparison was used rather than a comparison based on a median split of the entire data set owing to recommendations from Conway et al. (2005). Conway et al. (2005) argued that one major drawback of using median splits with WMC data is that participants with WMC scores falling around the median are likely to be misclassified (e.g., an individual who is capable of performing like a high WMC individual may have a WMC that falls just below the median). Extreme groups designs protect against the likelihood of misclassifying participants, as those participants with mid-range WMC scores are not included in the analysis.

WMC analysis. To ensure that across the two distracter types, those in the high WMC condition were not statistically different from one another, nor were those in the low WMC condition statistically different from one another, paired comparison analyses were
performed on the data. Data for the below analysis can be found in Figure 10. This figure provides mean WMC scores for each of the analyzed quartiles. The analyses indicated that there was no difference between the mean WMC score for those with low WMC in the speech vs. the nonspeech condition, $t(16) = 0.54, p > .29$. Likewise, there was no difference between the mean WMC score for those with high WMC in the speech vs. the nonspeech condition, $t(16) = 0.02, p > .49$.

Further, to ensure that those in the low WMC condition and high WMC condition were in fact statistically different from one another, an additional analysis was performed. Because the high WMC groups (speech vs. nonspeech condition) were not reliably different from one another and because the same was true for the low WMC groups, the data were collapsed over distracter-type in order to explore whether the low and high WMC conditions were statistically different from one another. The analysis resulted in a reliable difference between lower and higher WMC, $t(34) = 14.2, p < .001$. Combined the WMC analyses suggest that: (a) across speech type the two low WMC conditions are highly similar and the same is true for the two high WMC conditions and (b) those in the low WMC condition have reliably lower WMC than those in the high WMC condition. The mean WMC scores used in the present analysis are relatively similar to those reported by others (c.f., Redick et al., 2012).

**Target word accuracy analysis.** The mean proportion target word accuracy data can be found in Figure 11a. This figure presents the nonspeech and speech data for the low and high WMC participants. In terms of target word accuracy, those with low WMC in the nonspeech condition were more accurate than those in the speech condition. Likewise, those with high WMC in the nonspeech condition proved more accurate than
those with high WMC in the speech condition. A two-factor Analysis of Variance (ANOVA) with distracter type and WMC as the two factors and proportion target word accuracy as the dependent measure indicated both a reliable main effect of distracter type, $F(1,32) = 13.2, p < .005$, and a main effect of WMC, $F(1,32) = 4.8, p < .05$. The interaction proved unreliable ($F < 0.3$).

The target word accuracy data suggest that a nonspeech distracter does not cause as much of an impairment in processing the signal as does a speech distracter for those with lower WMC. Although those with lower WMC were still less accurate than their high WMC peers in the nonspeech condition, two important findings should be pointed out. First, those with low WMC in the nonspeech condition were almost twice as accurate as their low WMC peers in the speech condition. Second, these same individuals were seven percent more accurate than those with high WMC in the speech condition (e.g., mean nonspeech condition accuracy for low WMC = .48, mean speech condition accuracy for high WMC = .41).

**Biased word response analysis.** Figure 11b provides the data for the biased word responses. Those with lower WMC were less likely to report the biasing word in the nonspeech condition than when in the speech condition. Those with high WMC were also less likely to report the biasing word in the nonspeech condition than in the speech condition. An ANOVA indicated that both the main effect of distracter type, $F(1,32) = 20.4, p < .001$ and the main effect of WMC, $F(1,32) = 6.6, p < .05$ were reliable. The interaction proved unreliable, $F < 0.9$.

The biased word response data suggest that those with lower WMC were less likely to depend upon the biasing word when listening to speech in the presence of a nonspeech
distracter than when in the presence of a speech distracter. Further, they suggest that when a nonspeech distracter was presented, those with lower WMC actually were slightly more accurately able to identify the target word than those with high WMC who were distracted by speech.

**Summary.** The extreme groups findings suggest that regardless of WMC, listeners were less impaired by the presence of a nonspeech distracter than by a speech distracter when attempting to identify speech when it is unclear. Further, they suggest that those with lower WMC are less likely to attend to the biasing word when the distraction is nonspeech than when it is speech. In other words, those with lower WMC begin to depend more upon information that will result in greater accuracy when a nonspeech distracter is present than when a speech distracter is present.

The extreme groups analysis unlike the correlation suggests that those with lower WMC were more impaired in their ability to accurately identify the target word when a nonspeech distracter was present than those with higher WMC in this same situation. One likely reason for the discrepancy between these two analyses is that the distribution in this experiment showed a heavy cluster of WMC scores around the .55-.60 range. A reliable correlation may have been observed if the data had been more evenly distributed across WMC.

Importantly, the findings suggest that the level of similarity of the distraction seems to have increased demands on attention in a way that is proportionally similar across WMC. In other words, regardless of how WMC changes, as the signal becomes less similar, the same proportion of resources seems to become freed up. Had this not been the case, the improvement in target word accuracy should have been more pronounced in the high
WMC condition than in the low WMC condition when listeners were presented with a nonspeech distracter (i.e., an interaction between WMC and distracter type should have been seen in the extreme groups analysis).
Chapter 6: Experiment 5

In Experiment 4 it was found that when the target and distracter signals were relatively dissimilar, those with lower WMC continued to show an impairment in their ability to properly process the acoustic signal, though they showed a marked improvement in their ability to identify the target word than was the case when the target talker's sentence was produced in the presence of a speech distracter. In Experiment 4, the distracter was produced in such a way that it was made relatively noninvariant and thus was likely not highly predictable. Thus, the purpose of Experiment 5 was to examine whether listeners with lower WMC would continue to show an impairment in signal encoding if the distracter was stable and highly predictable.

It may be that when unpredictable or relatively dynamic distractions are present (e.g., a hospital or casino type of setting where many different and possibly unpredictable noises may occur one right after another), they monopolize executive attention and therefore cause listeners with lower WMC to reach their attentional capacity quickly. If the above hypothesis is accurate, it would be logical to assume that a relatively steady and predictable distracter would not require the employment of a significant amount of attentional capacity and should thus allow for adequate attention to be available for proper signal processing to take place, even when WMC is low. It would therefore follow that those with lower WMC would likely rely less upon biasing information than on the acoustic signal when identifying a word.
To test this hypothesis, the same target talker sentences (e.g., *The wing had an exquisite set of feathers* and *The wing had an exquisite set of diamonds*) as were used in Experiment 4 were again used in the present experiment. However, in this experiment, two important deviations from Experiment 4 were made. (1) Each sentence was paired with one highly steady sound (a sound with a relatively stable amplitude and frequency such as an air compressor). (2) Only a small number of sounds were used as distracters in order to produce stability. These sounds were recycled throughout the experiment in the same and therefore highly predictable order. Use of this methodology would allow listeners to become familiar with the distracters and thus the sounds should become quite predictable. If filtering out a distracter that is highly stable and predictable requires less attention than is the case for a more noninvariant sound such as those presented in Experiment 4, then regardless of WMC, participants should show relatively high target word accuracy and should make relatively few biasing word responses; a set of findings similar to that of the data patterns observed in Experiment 3. Such a finding would suggest that when distracting sound sources are highly predictable and stable, those with lower WMC are able to maintain enough of their capacity to properly attend to the acoustic signal. In contrast, if any distraction, no matter how stable or predictable it may be, is enough of a distraction to require substantial attentional capacity, it may cause listeners with lower WMC to experience a signal encoding impairment. In such a situation, the data patterns should prove relatively similar to those of Experiments 1 and 4. Namely, as WMC decreases, target word accuracy should decrease and biasing word responses should increase. Such a finding would suggest that even when a distraction is
relatively stable and predictable, it takes up a substantial amount of available executive attention.

**Method**

**Participants.** Thirty-eight new participants from the same pool meeting the same criteria as those in Experiment 1 participated for course credit.

**Stimuli.** The same 48 congruent and incongruent sentences (e.g., *The wing had an exquisite set of feathers* and *The wing had an exquisite set of diamonds* respectively) as were used in Experiment 1 were again used in the present experiment. To produce the distracting sounds, four of the 208 sounds from Experiment 4 were selected. The selected sounds all had fairly stable amplitudes and frequencies. The length of each file was matched to the duration of the secondary talker sentence (for the given target talker sentence) from Experiment 1. The amplitudes of two of the sounds were set to 72 dB owing to their consisting of lower frequencies which caused their amplitudes to appear lower perceptually. In contrast, the other two amplitudes were set to 68 dB for the opposite reason. All other aspects of stimulus construction were identical to that of Experiment 1.

**Procedure.** The procedure was identical to that of Experiment 4 with one exception. The presentation order of the distracter sounds was held constant. To accomplish this, the same four sounds were replayed throughout the experiment in the same order. In other words, if sound a was paired with target talker sentence 1, it would also be paired with sentences 5, 9 and so on. This was done to increase predictability of the distracters. As with the other experiments, the order of the target talker sentences was pseudo randomized.
Results and Discussion

**Primary analysis.** The OSpan and dichotic listening tasks were analyzed in an identical fashion to that of Experiment 1. The incongruent data are shown in Figures 12a (proportion target word accuracy as a function of WMC) and 12b (proportion biased word responses as a function of WMC). For target word accuracy, a small but noticeable positive-going trend can be seen in Figure 10a. In contrast, for biased word responses, Figure 10b shows a negative-going trend that is similar to that of experiment 1. Although this trend is likely heavily the result of four participants whose proportion biased word reports fell at or above .5, the trend is still present without these four individuals being included. Correlation analyses showed that only biased word responses ($r = -.36, p < .05$) proved reliable. The correlation between target word accuracy and WMC ($r = .17$) proved unreliable.

The findings suggest that although those with lower WMC may not be more impaired in their ability to encode the signal than those with higher WMC, they are still heavily relying on the biasing word.

**Signal similarity analysis.** One question that the above analysis can not fully address is whether listeners are less impaired by a highly stable nonspeech distraction. In other words, what is needed is a test of whether those in Experiment 5 (stable nonspeech condition), even those with lower WMC, are performing more accurately than individuals from Experiments 1 (speech) and 4 (nonspeech). In other words, is an interaction occurring such that when the distracter is highly stable, those with lower WMC appear similar to their high WMC peers while those with higher WMC outperform their lower WMC peers when the distraction is unpredictable. To accomplish this, the extreme
groups analysis reported in Experiment 4 was expanded to include the data from the present experiment. As with the prior analysis, data from the stable nonspeech condition were split into four quartiles. The nine (25%) highest WMC scorers (high WMC condition) and the nine (25%) lowest WMC scorers (low WMC condition) were analyzed.

**WMC analysis.** As with the extreme groups analysis in Experiment 4, it was first necessary to ensure that WMC was statistically similar (statistically unreliable) across distracter types for the low WMC condition and for the high WMC condition, and to ensure that the low WMC condition was reliably different from the high WMC condition. The data can be found in Figure 10. The data indicate a relatively large difference between those with low and high WMC. When comparing the stable nonspeech condition with those of the speech and nonspeech conditions, it is clear that the participants with low WMC in the stable nonspeech condition had a slightly higher WMC than those with low WMC in the other two distracter conditions. A two-factor ANOVA with WMC (low, high) and distracter type (speech, nonspeech, stable nonspeech) as the two factors and OSpan score as the dependent measure was performed. Only the main effect of WMC proved reliable, $F (1,48) = 349.7, p < .001$. The main effect of distracter type and the interaction between WMC and distracter type proved unreliable, both $F$s < 2.1. The findings indicated that regardless of distracter type, the low WMC participants were all relatively similar in their OSpan scores and the same was true for those with high WMC. Moreover, the findings revealed that those with lower WMC overall had reliably lower OSpan scores than their high WMC counterparts.
Target word accuracy analysis. The target word accuracy data for the present experiment are in Figure 11a. The findings indicated that as the distracter became less similar to the target talkers speech and became more predictable, listeners with lower WMC showed marked improvements in their performance. In contrast, for those with high WMC, performance plateaued once the two signals became less similar. Perhaps more importantly, those with lower WMC in the stable nonspeech condition seem to have improved to the level of their high WMC peers in this same condition. A two-factor ANOVA was performed with WMC and distracter-type as the two factors and proportion target word accuracy as the dependent measure. Only the main effect of distracter-type proved reliable, $F(2,48) = 11.3, p < .001$. The main effect of WMC was marginally reliable, $F(1,48) = 2.9, p < .10$. The interaction proved unreliable, $F < 1.0$. The findings from this analysis suggest that as the distracter becomes less like the target word and is more stable, listeners with lower WMC begin to be more able to attend to the signal. Moreover, those with lower WMC appear to be able to allocate an equal amount of attention to the signal to that of their higher WMC peers as long as the distraction can be anticipated with high accuracy.

Biased word response analysis. The biased word response data can be found in Figure 11b alongside the data from Experiments 1 and 4. For those with low WMC, a large decrease in biased word responses can be seen between the speech and nonspeech condition indicating that those in the nonspeech condition were less likely to report a word associated with the biasing word when the distracter was nonspeech than when it was speech. Those with lower WMC, however, only showed a slight decrease in their biased word responses in the stable nonspeech condition compared with the nonspeech
condition. In contrast, for those in the high WMC condition, a relatively steady and noticeable decrease in biased word responses can be seen across all three conditions, with the speech condition showing the highest proportion of biased word responses and the stable nonspeech condition showing the lowest proportion of biased word responses. Moreover, across all three distracter types, a clear difference between low and high WMC is present with high WMC participants being less likely to report the word associated with the biasing word (e.g., reporting *ring* when hearing *diamonds*). A two-factor ANOVA with biased word responses as the dependent measure was performed. The findings indicated that both the main effect of WMC \( (F(1,48) = 7.3, p < .01) \) and the main effect of distracter type \( (F(2,48) = 15.4, p < .001) \) were reliable. The interaction proved unreliable, \( F < 0.5 \). The findings suggest that those with lower WMC were more likely to endorse the word that was semantically associated with the target word. Moreover, regardless of WMC, listeners proved more likely to report this word when the distraction was unstable and more so when it was also similar to the target talker's speech.

**Full groups analysis.** Given the fact that there was a main effect of distracter type but no interaction in the extreme groups design, an analysis of the full data set for Experiments 1, 4, and 5 with WMC collapsed was performed. This analysis has the advantage of having more statistical power and also provides a general effect size of distracter on performance. This analysis was performed by simply computing the means of all participants in Experiment 1, in experiment 4, and in Experiment 5 for both target word accuracy and biased word responses. These data can be found in Table 1. This table
also includes mean OSpan scores per distracter type (speech, nonspeech, stable nonspeech conditions respectively).

**WMC analysis.** Before examining the target word accuracy and biased word response data, the mean OSpan scores for the three experiments were first compared to ensure that, across the three experiments, the scores were statistically similar to one another. The figure shows a slight increase in the mean OSpan score over the experiments with the stable nonspeech condition showing the highest mean OSpan score and the speech condition showing the lowest score. A single-factor ANOVA with distracter-type as the factor and OSpan score as the dependent measure proved unreliable, $F < 1.7$. 

**Target word accuracy analysis.** For the target word accuracy data, there is a large degree of improvement in performance in the nonspeech condition over that of the speech condition. In contrast, almost no improvement is seen between the nonspeech and stable nonspeech conditions. A single-factor ANOVA with proportion target word accuracy as the dependent measure showed a reliable main effect, $F(2,111) = 23.8, p < .001$. Planned comparison $t$-tests indicated that all three conditions were reliably different from one another: speech vs. nonspeech, $t(74) = 4.4, p < .001$; speech vs. stable nonspeech, $t(74) = 6.3, p < .001$; nonspeech vs. stable nonspeech, $t(74) = 2.5, p < .01$.

**Biased word response analysis.** The biased word response data indicated that the highest proportion of biased word responses occurred in the speech condition and the lowest proportion occurred in the stable nonspeech condition. A single-factor ANOVA with proportion biased word responses as the dependent measure proved reliable, $F(2,111) = 25.8, p < .001$. Planned comparison $t$-tests indicated that all three distracter types were reliably different from one another: speech vs. nonspeech condition, $t(74) =$
4.4, \( p < .001 \); speech vs. stable nonspeech condition, \( t (74) = 6.7, p < .001 \); nonspeech vs. stable nonspeech condition, \( t (74) = 2.9, p < .005 \).

**Summary.** Combined the correlational, extreme groups, and mean target word accuracy/biased word response data provide some interesting, and somewhat unexpected outcomes. First, overall, when a distracter is relatively dissimilar to a speech signal that the listener is trying to process, and the distracter is also highly stable, it impairs the listener's ability to identify the words the talker is saying to a lesser degree than if it is either similar or dissimilar and can not be highly predicted. Next, those with lower WMC are more impacted than their higher WMC peers in attending to the signal but only when a distraction is unpredictable or unstable. Further, and surprisingly, even though those with lower WMC are able to attend to the signal to a similar degree to that seen by their higher WMC peers given that the means are nearly identical across WMC, they are still more likely to depend upon biasing information than their higher WMC peers to aid word identification. It is as though these individuals are aware of the fact that they are sometimes less able to properly identify what the talker is saying when distracted and thus compensate by placing more weight on biasing information than is seen for those with higher WMC.
Chapter 7: General Discussion

Summary and Overview

Five experiments were performed in order to explore the following question: Do individual differences in WMC influence word identification when a listener is distracted? Experiment 1 (e.g., *The wing had an exquisite set of diamonds*) demonstrated that as WMC decreased, target word accuracy (e.g., reporting *wing* when *wing* was spoken) decreased and biasing word responses (e.g., reporting *ring* when *diamonds* was produced) increased. Experiment 2 demonstrated that those with lower WMC had likely been providing more biased word responses in Experiment 1 than those with higher WMC owing to the fact that they were unable to properly distinguish the signal from the noise. Experiment 3 suggested that as long as the signal was clear, listeners with lower WMC were able to perform equivalently with their higher WMC peers.

Experiments 4 and 5 examined the contribution of the distracter. Experiment 4 showed that those with lower WMC were less impaired by a distracter that was nonspeech than one that was speech. Moreover, those with higher WMC also showed an improvement in their performance when the distracter was nonspeech than when it was speech. Experiment 5 indicated that even when the distraction was highly predictable nonspeech, listeners with lower WMC continued to show more of an impairment than those with higher WMC in their performance, but importantly, this was only true for biased word responses. Those with lower WMC proved similar to that of their higher WMC peers in
their degree of target word accuracy. These findings suggest that the amount of executive attention a listener has will influence what information the individual will attend to during word identification. Further, they begin to provide reasons why listeners shift their attention to more biasing information when their attention is decreased.

The fact that WMC was able to account for at least some of the variance in Experiments 1, 2, 4, and 5, suggests that the ability to process words in the presence of distracting stimuli involves some degree of executive attention. Had this not been the case, no differences should have been observed across these experiments. The phoneme response analyses of Experiment 2 (e.g., fricative, stop, semivowel, and nasal,) suggest that when the listener is distracted, an adequate amount of executive attention must be available in order for the processor to be able to identify information such as the presence/absence of formants in addition to the frication being produced by the noise. If this hypothesis were not accurate, then those with lower WMC should not have shown any differences in the types of errors made from that of their higher WMC peers in Experiment 2.

The lack of any WMC effects in Experiment 3 is interesting in that it suggests that the ability to detect the phoneme in the presence of noise requires a relatively large amount of executive control. One way to illustrate this noise effect, is to examine the mean difference between target word accuracy across the congruent conditions in Experiments 1 and 3 which differed only in that the noise covered the target phoneme in Experiment 1 while this phoneme was presented as clear speech in Experiment 3 (e.g., The wing had an exquisite set of feathers vs. The wing had an exquisite set of feathers). If a large amount of executive control is being used to discriminate the phoneme from the noise and to
identify the formants of that phoneme, than the difference between the congruent condition of these two experiments is more likely to be large. The mean target word accuracy difference between these two conditions (clear minus noisy speech with the biasing word feathers) was .28, an effect that proved reliable $t (74) = 13.2, p < .001$.

Because the ability to discriminate the phoneme from the signal and the ability to identify the intended phoneme/word appear to require sufficient executive attention to be available, it makes sense that those with lower WMC were fairing worse in Experiments 1, 2, and 4 than those with higher WMC in target word accuracy and were showing an increased dependence on the biasing information as a way to compensate for this limitation.

The influence provided by the presence of noise (e.g., /=/ from =ing) alone is not the entire story, however. One final issue that needs to be discussed in order to more adequately understand the influence of executive attention on word identification under distraction, is the distracter itself. The extreme groups design in Experiment 5 provides some very important (and at least one unexpected) findings with respect to WMC. Recall that this analysis explored low vs. high WMC listeners in the speech, nonspeech, and stable nonspeech conditions. One important finding from this analysis, was that regardless of WMC, performance improved. Thus it would seem that for all listeners, filtering out a speech distracter places more demands on the processor than is the case for a nonspeech distracter. This would seem to suggest that the degree of similarity between the target talker and the distracter signals does not have a direct influence on WMC, but rather signal similarity influences word identification in general. In contrast to signal similarity, however, WMC was influenced by distraction predictability. When a distracter
changed without warning, listeners with lower WMC were more unable to accurately attend to the target talker's speech signal than when they could anticipate what distraction they would be hearing along with the speech. The findings provide important and different consequences for word identification for those with lower vs. higher WMC.

Those with high WMC are able to relatively easily filter out a distraction as long as it is relatively different from that of the target talker's speech. The stability or predictiveness of the distraction does not matter as long as the signal is dissimilar from that of the target talker's speech signal. For those with lower WMC, only when they can anticipate what specific distraction will be present during the target talker's sentence were they able to adequately filter the distraction out while also attending to the words spoken by the target talker. Recall that earlier it was argued that the ability to discriminate the phoneme from the noise and properly identify the target phoneme and thus accurately recognize the word, requires executive attention. It is, therefore, likely that because more executive attention is needed to identify the word when the noise is present, too little attention can be devoted to properly filtering out the distraction unless it can be highly anticipated. If listeners with lower WMC are able to adequately filter out the distraction when it is anticipated, this would potentially free up some of their remaining executive attention to be devoted to the noisy portion of the signal.

The one truly unexpected outcome, was that even though those with lower WMC were more accurate in the stable nonspeech condition than they were in the nonspeech condition, they were still more biased than their higher WMC peers, though they were less so than those with lower WMC in the nonspeech condition. It is possible that these
individuals realize that they have difficulty in processing speech when they are distracted and thus place more of their attention on any available biasing than acoustic information.

Below are sections that examine separately some of the issues described above. Further, these sections attempt to add new insights into what may be taking place throughout the word recognition process when executive attention is low or high and the listener is distracted.

**Lexical Activation**

One specific and important finding that is evident in Experiment 2 (when the biasing word was absent: e.g., *The wing had an exquisite set of*) is that when WMC is low, and a listener is distracted by a secondary talker, the ability to identify the target word is not only impaired, but the ability to properly distinguish extraneous noise from the talker's speech appears to be nearly impossible to accomplish. Had this not been the case, in Experiment 2 those with lower WMC should have produced fewer fricative onsets when reporting the target word.

It is likely, that as was argued in Experiment 2, those with lower WMC are experiencing a deficit in their ability to properly encode the signal when it is unclear. This would likely result in an impoverished lexical representation being activated. Essentially, in this case, these individuals are not able to attend closely enough to the signal to adequately distinguish it from the noise and thus a rich representation would not be able to be acquired. This is because presumably too little executive attention is left over after the distraction has been filtered out to allow the processor to attend at an adequate level to the target signal.
Attentional Flexibility

If the signal encoding deficit hypothesis is correct, one question that arises is: Why is it that individuals with lower WMC are having such difficulty properly encoding the signal? One possible answer is that listeners with lower WMC may lack the executive control necessary to adjust their focus on the signal. In other words, these individuals may not be able to focus closely enough to the signal when distracted to be able to properly identify the more fine-grained acoustic information.

Some evidence supporting the idea of limited attentional flexibility comes from a study by Colflesh and Conway (2007). They predicted that listeners with lower WMC would have less flexibility in controlling to what degree stimuli were attended to than their higher WMC peers. To test the above hypothesis, they presented participants with two passages simultaneously in dichotic form. Participants were asked to shadow one of the two passages while listening for their name to be spoken in the other passage. They found that those with higher WMC were more likely to recall hearing their own name than those with lower WMC while maintaining accurate shadowing. Colflesh and Conway (2007) noted that the above finding could provide an important piece of information about WMC when combined with a finding reported by Conway et al. (2001) described in Chapter 1.

Recall that Conway et al. (2001) found that when listeners were asked to ignore a passage containing their name while shadowing another passage, those with lower WMC were more likely to become distracted when hearing their own name spoken in the unattended passage than those with higher WMC. Colflesh and Conway (2007) argued that the two findings combined demonstrate that those with higher WMC have the ability
to properly control what is being attended to more adequately than those with lower WMC. They further argued that those with lower WMC have much more immutable attentional control when distracted than those with higher WMC.

The findings from Colflesh and Conway (2007) and Conway et al. (2001) suggest that it is likely that those with lower WMC in the present set of experiments were not adequately able to shift their attention as needed (e.g., they were not able to increase their focus on the signal when it became less clear: e.g., wing vs. $\text{w}$ing) so as to distinguish its features adequately from those of the noise masker).

One further piece of evidence that provides some, though not conclusive support for this argument comes from what participants were told during the instructions in the experiments. In all five experiments, in order to provide the strongest possible test of biasing word influence on word identification, participants were told that when making their word judgments, they should be as accurate as possible. They were further told that in order to accomplish the above, they would need to listen closely to the "sounds that made up the word" that the target talker produced and not use other information in the sentence to make this judgment. This emphasis fits in line with instructions reported by Mattys et al. (2010). The participants were also made aware, multiple times, that there would be noise along with the target phoneme (in all but Experiment 3). Further, they were given several examples of what a noise and phoneme together would sound like and were given feedback on a couple of demos during the instructions in order to ensure that they understood the task properly. Participants should only have produced a large proportion of fricative target word onsets (e.g., thing in Experiment 2 if they could not flexibly focus their attention at the level necessary for proper signal processing.
Attentional Reallocation

It is clear that regardless of why those with lower WMC were experiencing a signal-encoding deficit, these individuals are compensating to overcome this deficit. To accomplish this, these individuals are relying on the biasing word. The finding suggest that during distraction, as attention decreases and the signal becomes less clear, attention is reallocated in a way that will likely help maximize word identification while reducing the demands placed on the system. To accomplish this, as the signal becomes less clear, listeners with lower WMC begin to place greater attention on biasing information. It may be that the system titrates the weighting of various types of information such that as more attentional resources are given to other tasks (e.g., filtering out the distraction), weighting of the signal is decreased and weighting of the biasing information is increased (c.f., Norris and McQueen, 2008 for a possible method of implementing this idea using a Bayesian processor).

Distracter Influences on Word Recognition

Despite how the above process may be taking place, it is clear that the distracter is playing some role in word processing that affects individuals differentially based on the degree of attentional capacity that is available to them. Interestingly, the distraction seemed to only impair listeners when the signal became unclear. Regardless of how low WMC proved to be, listeners were able to identify the target word and properly integrate this word into the sentence (as can be seen by the high accurate sentence sensibility judgments observed across WMC in Experiment 3: e.g., *The wing had an exquisite set of diamonds*). It is possible that even when the target talker's speech was clear, the distracter still produced a cost, it just proved not to be enough of a cost to actually impair word
identification accuracy. If those with lower WMC really are attending to the signal at a very shallow level, even if those with lower WMC were doing so in the clear speech experiment, these individuals would have been able to identify the target word accurately and make their sentence sensibility judgments with relatively high accuracy. Thus, it is entirely possible that for those with lower WMC, the distracter took up a significant amount of attentional capacity and thus caused the same reduced attention to be paid to the signal. This situation would still lead to a processing cost, but the cost would not be sufficient enough to impair accurate word identification. When the signal was unclear in contrast, the processor cost became evident as those with lower WMC showed a noticeable impairment in their ability to identify the target word (e.g., target word accuracy in Experiments 1, 2, and 4 and the increased fricative reports in Experiment 2).

One important finding with respect to the distracter as noted earlier, is that not all distracters provide the same degree of impact on participants' ability to recognize words. Had listeners been equally distracted as long as the signal was unclear and a distracter had been present, then Experiments 1 (speech distraction) and 4 (nonspeech distraction) should have resulted in relatively identical data patterns. Given that listeners, regardless of WMC showed greater accuracy in attending to the signal and less dependence upon the biasing word in the presence of a nonspeech distracter than in the presence of a speech distracter suggests that the two types of distracters are not equally disruptive to the processor.

It may be that the reason for the increased amount of distraction in the presence of a speech distracter, is that the two signals (target and secondary talker) would be more similar to one another than would likely be the case for a nonspeech distraction (i.e.,
If this is accurate, then it may be that as the two signals become more similar, more attention is needed to distinguish them from one another so as to ensure the correct signal is being filtered out.

Alternatively, it may be that because the distracter was coming into the right ear, the right-ear signal was reaching the language center first (Foundas et al., 2006) and thus when this distraction was speech, it was beginning to be processed before the left-ear signal had a chance to reach the language center. In such a case, it is possible that the target talker's speech was less adequately able to be processed because other information in the same center within the brain was already taking up valuable executive control. Because those with lower WMC have less executive control to spare, it would make sense that they would be less able to identify *wing* when distracted by speech than by nonspeech. Moreover, because those with higher WMC would presumably need to use some of their executive control to deal with the competing speech scenario, they too would be expected to show somewhat of a deficit in the presence of a speech distracter compared with a nonspeech distracter. Such an explanation would fit with the extreme groups analysis reported in Experiment 4 where all participants, regardless of WMC showed less accuracy in identifying the target word in the presence of a speech distracter than in the presence of a nonspeech distracter.

One way to more thoroughly test the above hypothesis, would be to present the target talker to the right ear and the secondary talker to the left ear. If the above hypothesis is correct, namely, that the speech entering the right ear is reaching the language center first and thus is using up valuable attentional capacity, then switching the ear presentation should eliminate much of the observed deficit. It is likely, however, that even if this
hypothesis is correct, those with lower WMC will still be at a disadvantage compared with their higher WMC peers. One reason is that even a nonspeech distracter impaired those with lower WMC to a greater degree than those with higher WMC. This should not be the case if the entire deficit could be explained by the fact that the secondary talker's speech is reaching the language center first and is thus using up too much available executive attention.

Experiment 5 (e.g., stable nonspeech distracter) resulted in an interesting and somewhat surprising set of findings that may need a little further explanation. Recall that in this experiment, those with lower WMC performed equivalently with their higher WMC peers in terms of their target word accuracy, but showed greater dependence upon the biasing word than their higher WMC peers. These two findings may seem a bit incongruous. However, one possible explanation may be that the findings are an indicator that those with lower WMC, may be aware, on some level, that they possess a signal encoding deficit. If they have some awareness of this deficit, they may be more likely to compensate for this shortcoming once they are put into a situation where any type of distraction is present, even if the distraction is highly predictable. In such a situation, these individuals may be willing to rely more heavily on biasing information when certainty of the intended word is low (as when a noise is partially masking the word being identified). In contrast, in this same low-distraction situation, those with higher WMC may continue to rely more heavily on the acoustic signal.

Regardless of how the distracter is having its effect on word identification, an important result is apparent from the present set of experiments. The ability to process words in the presence of subsequent biasing information is not fully automatic. Had this
been a truly automatic process, the distracter should have had no influence on word identification. A truly automatic process should have been able to take place similarly regardless of whether the listener had been distracted or not (c.f., Styles, 1997 for a thorough review of automatic vs. controlled attention).

Alternative Causes

The discussion thus far has been based on the working assumption that individual differences in WMC led to the data patterns observed across the five reported experiments. This assumption begs for a discussion of whether WMC was in fact the real contributing factor or whether some other process could have been masquerading as WMC and thus produced the observed effects. One argument in favor of an alternative process producing the observed effects is the somewhat weak trends observed in the experiments. Although most of the trends observed across the five experiments range in strength from $r = .29-.39$, they only account for approximately 9-16% of the total variance in the data. So if WMC is not the primary contributing factor, what may be able to account for more of the total variance? At least three alternative possibilities might be able to account for more of the variance in the data.

One alternative account that may be able to explain the data patterns observed across the experiments is that of general fluid intelligence. General fluid intelligence (e.g., the ability to reason, to solve problems, and to use new information) has been found to be highly correlated with WMC (c.f., Conway et al., 2005; Engle, Laughlin, Tuholski, & Conway, 1999; Kane et al., 2004). Standard tests of general fluid intelligence have been found to be highly correlated with the domain-general component of WMC (Conway et al., 2005). It may therefore be that had general fluid intelligence been tested, more of the
variance would have been accounted for. Because of how highly correlated these two constructs have been found to be (c.f. Conway et al., 2005), it may not be possible to easily tease the two constructs apart, especially if the two constructs account for some of the same involved processes.

A second possibility is that attention in general was the contributing factor to the data patterns. Recall that the model of WMC that was investigated in the present work assumes that WMC is made up of two components; specific resources and general executive attention. It was argued that the effects were based on the attentional part of WMC. It may, however, be that some more general attentional store, not directly linked to WMC produced the effects observed across the experiments. This possibility is difficult to test as it would be quite difficult to determine at what point attention is no longer part of WMC. To this author's awareness, there is no test at present that adequately can distinguish the attentional resources of WMC from a more general attentional store.

A final possibility is that those who received lower OSpan scores were simply less motivated. In such a case, an individual with lower motivation, for whatever reason (e.g., tiredness, lack of interest, focus on unrelated factors...), might not have been performing to their best ability on the OSpan, or at least the storage (recall) portion of the OSpan task. Such a situation would likely lead to reduced OSpan scores. In applying this to the dichotic listening task, it would be logical to assume that someone with lower motivation would simply use the most easily accessed material to drive their responses. In other words, a listener with lower motivation may find it takes less of their effort to simply respond with a word that is semantically related to the biasing word (e.g., feathers) rather than listening closely to the signal, an outcome that would prove similar to the observed
data patterns. One piece of evidence that provides some support against a purely motivational explanation is that of the sentence sensibility judgments in Experiment 3. If those with lower OSpan scores simply had lower motivation rather than poorer WMC, then their sentence sensibility judgments in Experiment 3 should have also been low given that their would be no reason for them to perform this task well. In other words, there was no reward given for good sentence sensibility performance. Thus, there should have been no extrinsic motivational drive for good performance. Given that despite differences in WMC listeners in Experiment 3 performed quite well on this task ($m = .87$) and there were no differences across WMC ($r = -.06$), suggests that those with lower WMC were putting a similar degree of effort into this task as their higher WMC peers. There is no logical reason why listeners with lower motivation should choose to perform better on this task than on the target word identification task as the sentence sensibility task requires not only understanding the target word, but integrating it with the rest of the sentence.

The present experiments provide no way to directly test whether WMC is in fact leading to the observed trends. What is needed at this point is a study that measures both WMC and general fluid intelligence, along with a test of listener motivation. If WMC is in fact having the expected influence, then it should be able to account for a significant amount of unique variance that the other two factors do not account for. Such a test would provide stronger evidence either favoring or disfavoring the WMC influence argued to be driving the data trends in the current set of experiments.
Concluding Remarks

Overall, the present findings provide new insights into spoken word recognition. Importantly, they suggest that WMC, most notably executive attention, plays an important role in spoken word recognition when subsequent biasing information is present and the listener is distracted. The findings demonstrate that as WMC decreases, listeners become more dependent upon biasing information and this is most likely due to a deficit in signal processing.

Further research is necessary to better understand how these processes work and the underlying mechanisms that evoke them. As WMC is a construct that individuals have been found to vary widely on (c.f., Conway et al., 2005; Kane et al., 2004; Redick et al., 2012 for demonstrations across differing populations) and because WMC was shown to influence performance in the present set of experiments, further exploration into the influence of WMC on spoken word recognition is necessary.
References


80


Szostak, C. M. (2009). Identifying the *eel on the table: An examination of processes that
aid spoken word ambiguity. *Unpublished Master's Thesis*, The Ohio State University.


Endnotes

1. Although Warren and Warren (1970) were referring to a situation where lexical knowledge has a direct influence on phoneme perception, no specific claims are being made here as to whether lexical information has a direct or indirect influence on perception.

2. Because the stimuli were produced at a rate that was deemed by the present and another phonetically trained researcher to be unnaturally slow, the stimuli were time compressed to approximately 78% of their original speech rate.

3. This task was used in this and all subsequent experiments.

4. Redick et al. (2012) recommended using an automated version of the OSpan task wherein the timing of studied word lists is based on a relative constant rate. Pilot tests of the automated OSpan task performed by the present researcher proved uninformative regarding WMC. In these tests, participants appear to have begun learning how much time was available and appear to have been using strategies such as rehearsing studied lists between trials (e.g., solving the math problem and then using additional time to rehearse the list before moving on to the next trial). Because WMC is a construct which deals with the simultaneous ability to store and process separate stimuli according to Conway et al. (2005), use of this test was unlikely to provide a good measure of the construct of interest.
5. This exclusionary and scoring method was used for this and all subsequent experiments.

6. The biasing word was never presented in this experiment. Thus, the use of the term "absent condition" is solely for clarification when comparing this condition with conditions in other experiments.

7. Owing to human error, a single nonspeech sound clip was included in the chains for both versions of the match sentence. Because these two sentences were not adjacent to one another, the sound clip in question was not located in the same position within the two chains, and given the large number of sound clips ($n = 208$) used in this experiment, it was highly unlikely that this single duplicate presentation would have any impact on the data. Thus, to ensure that statistical power remained as strong as possible, this item was analyzed in this experiment.
Appendix A: Target Talker Sentences

The target talker stimuli that were used in Experiments 1-5 are provided below. The italicized word was the word congruent with the incongruent biasing word. Note that the italicized word was never spoken. The last two words in all sentences are the biasing (congruent vs. incongruent) words. These two words were spliced out of the sentences in Experiment 2.

The knock/lock was not quite able to be heard/snapped.
The leak/beak was unexpectedly fixed/sharp.
The lip/rip was very noticeably chapped/stitched.
The list/wrist was found to be thoroughly numbered/injured.
The log/dog was very actively rotting/barking.
The look/book was not going to be modeled/written.
The match/latch was readily able to be struck/closed.
The math/bath had included a lot of fractions/bubbles.
The mole/role was not adequately digging/assigned.
The mug/rug was going to be completely filled/swept.
The nap/wrap was unfortunately not restful/tasty.
The nest/guest had some very colorful sparrows/luggage.
The niece/lease was finally going to be married/cosigned.
The night/light was not actually starless/shining.
The roast/ghost was considered to be very salty/spooky.
The rum/gum was not going to be poured/chewed.
The well/bell was unexpectedly dry/rung.
The wick/nick was insufficiently burning/repaired.
The wife/knife was most definitely pregnant/pointed.
The wig/rig was not going to be shampooed/driven.
The wing/ring had an exquisite set of feathers/diamonds.
The wish/dish was unexpectedly granted/shattered.
The wool/bull was not going to be knitted/angered.
The yarn/barn was in the first stage of being woven/painted.
Appendix B: Secondary Talker Sentences

The secondary talkers' sentences that were used in Experiments 1-3 are provided below.

The word above each sentence set indicates the target word associated with the sentences. The letters at the beginning of each sentence indicate whether it was paired with the congruent (C), incongruent (I), or absent (A: Experiment 2) condition.

knock

C  My mother was always there for me I would just come

I  They have only come to my home three times in twenty years

A  That's the responsibility I took on as being a parent

leak

C  She's three so she's still puppy but she's getting there

I  I don't know what it was like really here I can imagine

A  I've threatened to move on more than one occasion

lip

C  This one guy that's been doing stuff for our house for a long time

I  I saw that the first time in 1992

A  If you put a discount store in there I would love it
C There not always doing this and there not bothering people
I They have two elementaries one middle school and one high school
A You go in the morning and your a senior you go in the afternoon

C I'm doing a function tonight for 300 people
I You can get on this and you can get on that
A Cause I rent and I always get very upset

C The first chance that I've had in probably twenty years
I I've been here for a little over a year
A I never liked meat my parents always made me eat it

C Question him about where he's been or anything like that
I I don't understand why he can't do that you know I've actually called them
A They were splitting the kids up and putting them in other classrooms
C I was marathon training and I only got up to about nine miles
I And the fact that they were like driven to that point
A I changed my name when I went to college I just decided I'm out of here

mole
C She doesn't really drive very much which is good
I You just keep after them you keep giving them information
A and I'm like so up in the air about what I want to do

mug
C I don't have a like a regular size class
I I would have struggled with it but it was not my own job
A Now its such a open market and they know that they have to

nap
C I just wasn't interested in that when I was in high school at all
I My mom can be getting paid from the government
A Come from a completely different you know background
nest

C There's a main place where they park all their trucks and then he just goes
I It is harder on him because he is demanding
A You come from a situation like that that doesn't mean you have to

niece

C He lived around us he never called us you know
I I would have had a whole lot more kids you know I would have just kept having them
A Because I had about six years experience they had to pay

night

C A company my sister had worked for wanted her to come
I You know they shouldn't know where the keys are if there were keys
A Cash it in get a video or video game

roast

C I guess its true you just have to wait and see once you get there
I You don't want them you know working actively against you
A I could kind of keep in touch with what they were supposed to be learning
C You don't want them you know working actively against you
I That's what I like about it is the ambiance
A It's the fact that I have a tremendous respect for my kids

well
C One of them had four children one of them had like six you know
I I normally wake up at six anyway so it doesn't matter
A We don't really have a set schedule right now

wick
C I mean a year would come up and I would say ok here we go
I For the convenience of the hours
A There's so much volunteering to be done

wife
C It was just I never end up fitting in with everybody
I If I had five or six thousand dollars out of it I was happy
A Oh exactly I'm planning on just staying in the house
C The value of this is very important
I Think that the parents should be held responsible
A I mean in those situations they still want them to

C I mean there's enough room where people can come and sit down there's like a fold
I There both fairly wooden neither one of them seems comfortable
A Oh I have a few kids from time to time that said that they want to

C And they only work for about forty percent of the population
I The apartments are just very very plain
A I was never upset about the fact that they weren't together

C I just thought it was very strange that he would have his whole family
I Its not a long period of time I'm only there till ten
A We figured out that it was just the time period
yarn

C  I went to her and I said it was nice working with you I'm going home

I  You know I think those are still very important

A  When I was fourteen I didn't care what anybody told me
# Appendix C: Sound Corpus Categories

Categories that made up the sound corpus from Experiment 4 are provided below. Counts are based on the number of items that fell into that category.

<table>
<thead>
<tr>
<th>Count</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>animals (e.g., birds chirping, forest sounds, sheep)</td>
</tr>
<tr>
<td>8</td>
<td>body movement (e.g., walking, crunching)</td>
</tr>
<tr>
<td>3</td>
<td>cleaning (e.g., brushing objects, dusting, picking up spilled objects)</td>
</tr>
<tr>
<td>5</td>
<td>destruction (e.g., breaking glass, hitting things)</td>
</tr>
<tr>
<td>16</td>
<td>entertainment (e.g., card shuffling, race starter)</td>
</tr>
<tr>
<td>7</td>
<td>grooming and dressing (e.g., electric shaver, zipper, hair drier)</td>
</tr>
<tr>
<td>19</td>
<td>indoor household (e.g., air conditioner, appliances)</td>
</tr>
<tr>
<td>17</td>
<td>musical instruments and tones (e.g., bells, harp, whistles)</td>
</tr>
<tr>
<td>4</td>
<td>large tools/machinery (e.g., jackhammer, metal saw, air compressor)</td>
</tr>
<tr>
<td>3</td>
<td>liquid (e.g., pop fizzing, poring a drink)</td>
</tr>
<tr>
<td>7</td>
<td>miscellaneous (e.g., cranking an object, shaking a tablet bottle)</td>
</tr>
<tr>
<td>4</td>
<td>money (e.g., counting, coin dropping)</td>
</tr>
<tr>
<td>12</td>
<td>non-animal nature (e.g., river, waves, water droplets, weather)</td>
</tr>
<tr>
<td>7</td>
<td>office (e.g., typing, printer)</td>
</tr>
<tr>
<td>8</td>
<td>outdoor household (e.g., wind chimes, lawn mower, grass trimmer)</td>
</tr>
<tr>
<td>8</td>
<td>paper (e.g., rustling, crumbling, ripping)</td>
</tr>
<tr>
<td></td>
<td>Category</td>
</tr>
<tr>
<td>---</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>5</td>
<td>scifi (e.g., laser beam, space ship)</td>
</tr>
<tr>
<td>6</td>
<td>tools (e.g., saw, hammer, drill)</td>
</tr>
<tr>
<td>25</td>
<td>transportation (e.g., engines, airplanes, jingling keys, seatbelt latching)</td>
</tr>
<tr>
<td>208</td>
<td>total</td>
</tr>
</tbody>
</table>
Appendix D: Table and Figures

The table and figures for Experiments 1-5 are provided on the next several pages.

Table 1: Overall mean analysis for WMC, target word accuracy, and biased word responses.

<table>
<thead>
<tr>
<th>Condition</th>
<th>WMC</th>
<th>Accuracy</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech</td>
<td>.56 (.02)</td>
<td>.34 (.03)</td>
<td>.46 (.03)</td>
</tr>
<tr>
<td>nonspeech</td>
<td>.58 (.02)</td>
<td>.51 (.02)</td>
<td>.31 (.02)</td>
</tr>
<tr>
<td>stable nonspeech</td>
<td>.61 (.02)</td>
<td>.59 (.02)</td>
<td>.22 (.02)</td>
</tr>
</tbody>
</table>

*note.* The value in parentheses is the standard error of the mean.

*note.* Each cell contains *n* = 38 participants.

a. The conditions refer to Experiments 1 (speech), 4 (nonspeech) and 5 (stable nonspeech).

b. WMC = working memory capacity. Data in this column are the mean OSpan scores.

c. Accuracy= target word accuracy (data are mean proportions).

d. Bias = biased word responses (data are mean proportions).
The plots depict target word accuracy (1a) and biased word responses (1b) as a function of WMC in the congruent condition from Experiment 1. Target word accuracy $r = .03, p > .85$; biased word accuracy $r = -.08, p > .63$. 

97
The plots depict target word accuracy (2a) and biased word responses (2b) as a function of WMC in the incongruent condition from Experiment 1. Target word accuracy $r = .29$, $p < .08$; biased word response $r = .36$, $p < .05$. 

98
The plots depict target word accuracy (3a) and fricative responses (3b) as a function of WMC in the absent condition from Experiment 2. Target word accuracy $r = .32, p < .06$; fricative response $r = -.39, p < .05$.  

99
Figure 4.

Absent Stop Responses by Working Memory Capacity (Experiment 2)

The plot depicts stop responses as a function of WMC in the absent condition from Experiment 2, \( r = .31, p < .06 \).
The plot depicts semivowel responses as a function of WMC in the absent condition from Experiment 2, $r = .27, p < .11$. 

Figure 5.

Absent Semivowel Responses by Working Memory Capacity
(Experiment 2)
The plot depicts nasal responses as a function of WMC in the absent condition from Experiment 2, $r = .11, p > .52$. 
The plots depict target word accuracy (7a) and biased word responses (7b) as a function of WMC in the incongruent condition from Experiment 3. Target word accuracy $r = .09$, $p > .57$; biased word response $r = -.11$, $p = .50$. 

103
The plot depicts sentence sensibility judgments as a function of WMC in the incongruent condition from Experiment 3, $r = .06, p > .74$. 
The plots depict target word accuracy (9a) and biased word responses (9b) as a function of WMC in the incongruent condition from Experiment 4. Target word accuracy $r = .31$, $p < .07$; biased word response $r = -.24$, $p > .13$. 

105
The graph depicts the mean OSpan scores for the highest 25% of participants (10a: high WMC) and lowest 25% of participants (10b: low WMC) for Experiments 1 (speech condition), 4 (nonspeech condition) and 5 (stable nonspeech condition), $n = 8$ per condition.
The graphs depict the mean proportion of target word accuracy for the highest 25% of participants (11a₁: high WMC) and lowest 25% of participants (11a₂: low WMC) for the incongruent condition of Experiments 1 (speech condition), 4 (nonspeech condition) and 5 (stable nonspeech condition), n = 8 per condition.
The graphs depict the mean proportion of biased word responses for the highest 25% of participants (11b₁: high WMC) and lowest 25% of participants (11b₂: low WMC) for the incongruent condition of Experiments 1 (speech condition), 4 (nonspeech condition) and 5 (stable nonspeech condition), $n = 8$ per condition.
The plots depict target word accuracy (12a) and biased word responses (12b) as a function of WMC in the incongruent condition from Experiment 5. Target word accuracy $r = .17, p > .31$; biased word response $r = -.36, p < .05$. 

109