

ABSTRACT

Title of Document: PRE-TASK PLANNING TIME AND
WORKING MEMORY AS PREDICTORS OF
ACCURACY, FLUENCY, AND
COMPLEXITY.

Katharine Brown Nielson, Ph.D., 2012

Directed By: Dr. Robert DeKeyser, Professor, Second
Language Acquisition

Working memory, which accounts for the ability to process information in the face of interference, is critical to second language acquisition (SLA) and use. The interaction of working memory capacity (WMC) with specific pedagogical interventions is a logical place for empirical SLA research, both to examine the cognitive processes underpinning second language performance and to identify instructional treatments that may differentiate learners based on their WMC. A good candidate for such an examination is planning time, a pedagogical intervention that has been the subject of extensive empirical research, which has, thus far, been largely unrelated to WMC. The study undertaken here considers WMC along with two different types of pre-task planning time (guided and unguided) as predictors of the accuracy, fluency, and complexity of learners' discourse.

Ninety-two intermediate ESL students from seven classes at a community college participated in this study by completing two different working memory span tasks as well as two different “there-and-then” oral story-telling tasks. The treatment condition of the story-telling tasks was manipulated so that learners’ performance could be considered in terms of provision of pre-task planning (+/- planning), type of planning (guided vs. unguided), and order of planning (planning first or planning second).

The results demonstrate that the relationship among type of planning time, order of planning time, and WMC is complex. Task order had a clear effect on learners’ production, regardless of the provision of planning time. When learners began the series of story-telling tasks under the + planning condition, their output on the subsequent, unplanned task varied according to whether they had first received guided or unguided planning time. In addition, guided planning time and unguided planning time also have very different effects on learners’ production, with guided planning time promoting a focus on accuracy at the expense of complexity and unguided planning time fostering fluency. Finally, this study indicates that task conditions can affect learners with high and low WMC in different ways. Learners with high WMC are more likely to comply with complex story-telling instructions, improving their focus on grammatical form at the expense of fluency, whereas learners with low WMC are more likely to improve their fluency as a result of task repetition, regardless of the task conditions.

PRE-TASK PLANNING TIME AND WORKING MEMORY AS PREDICTORS
OF ACCURACY, FLUENCY, AND COMPLEXITY

By

Katharine Brown Nielson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Robert DeKeyser, Chair
Professor Michael Long
Professor Steven Ross
Dr. Michael Bunting
Professor Stephen Koziol, Dean's Representative

© Copyright by
Katharine B. Nielson
2012

Acknowledgements

I extend my sincerest thanks to my advisor, Dr. Robert DeKeyser, for patiently sticking with me through this long process, which began with a term paper for his class in 2007. His guidance and encouragement over the years have been invaluable. I would also like to thank my committee members: Dr. Michael Long, Dr. Steven Ross, and Dr. Michael Bunting for their guidance, feedback, and assistance with this dissertation as well as with other research over the years. In addition, I would like to thank Dr. Cathy Doughty for her mentorship throughout my doctoral studies, and Dr. Stephen Koziol for agreeing to serve as my Dean's Representative.

I am thankful for the support of my graduate student colleagues, especially Gisela Granena, Megan Masters, and Susan Benson for their help talking through issues with data collection and analysis and for their comments on early drafts. I am also grateful for the CASL Dissertation Fellowship, which allowed me to finish the dissertation in timely fashion. And none of this would have been possible without many ESL students and teachers at Northern Virginia Community College, who volunteered their time to participate in this study.

Finally, I am incredibly thankful for my family. To my husband, Erik, for his unwavering encouragement and the countless hours he spent reading drafts of my work; to my parents and brother and sister, for their willingness to take care of our children while we were in class or conducting research; and to my two sons, Sam and Mac, who have demonstrated more patience and understanding than should be expected of a four- and a six-year-old: thank you all.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iii
List of Tables.....	v
List of Figures.....	vi
Chapter 1: Review of the Literature.....	1
1.1 Working Memory.....	2
1.2 Working Memory and Stages of SLA.....	6
1.3 Working Memory and Task Complexity in SLA.....	8
1.4 Planning and SLA.....	10
1.5 Working Memory and Planning in SLA.....	13
1.6 Manipulating Planning Time.....	18
1.7 Cognition and Task Conditions.....	22
Chapter 2: Purpose of the Study.....	26
Chapter 3: Research Hypotheses.....	30
Chapter 4: Methodology.....	35
4.1 Participants.....	35
4.2 Instruments.....	37
4.3 Procedures.....	40
4.4 Calculation and Operationalization of Variables.....	43
Chapter 5: Results.....	48
5.1 Descriptive statistics and initial analyses of variables.....	48
5.2 Planning Order/Conditions and WMC: Initial Analyses.....	51
5.3 The Order of the Provision of Planning Time.....	56
5.3.1 Descriptive statistics split by planning order.....	56
5.3.2 Between-groups data analysis without the influence of task repetition.....	58
5.3.3 Repeated measures analyses: No planning first.....	61
5.3.4 Repeated measures analyses for participants who began with planning time.....	63
5.3.5 Summary of results from split-group repeated measures analyses.....	66
5.4 Structural Complexity.....	67
5.5 Summary of relationship among variables and planning conditions.....	68
5.6 Working Memory.....	73
5.7 Interrelationship among dependent variables and WMC.....	81
5.7.1 WMC, type of planning time, and planning order.....	82
5.7.2 WMC and No-Planning Conditions.....	86
5.8 Specific Accuracy Measures.....	89
Chapter 6: Discussion and Conclusions.....	93
6.1 The Influence of Task Order on Learner Fluency.....	93
6.2: Guided and Unguided Planning as Predictors of Fluency and Complexity.....	97
6.3 Accuracy of Participants' Performance.....	99
6.4 WMC and learner production.....	102
6.4.1 WMC and accuracy across planning conditions.....	102
6.4.2 WMC and fluency.....	104

6.4.3 WMC and discourse complexity.....	109
6.7 Synthesis of Results by Planning Condition and WMC	111
6.7.1 Guided Planning Time	111
6.7.2 Unguided Planning Time	112
6.7.3 No-Planning	112
6.7.4 Task Repetition	112
6.7.5 Task Repetition and WMC	113
6.7.6 WMC.....	114
6.8 Conclusions and Directions for Future Research.....	114
Appendices.....	119
Appendix A: Language background questionnaire.....	119
Appendix B: Images for story-telling tasks A and B	120
Appendix C: Instructions for guided planning time	121
Appendix D: Instructions for unguided planning time	122
Appendix E: Instructional handout given to participants	123
References.....	124

List of Tables

Table 1: Participants' Native Languages	37
Table 2. Task conditions and task order for both groups.....	42
Table 3. Descriptive statistics for accuracy with and without guided and unguided planning time	48
Table 4. Descriptive statistics for fluency with and without guided and unguided planning time	49
Table 5. Descriptive statistics for complexity with and without guided and unguided planning time	49
Table 6. Dependent variables used in inferential statistical analyses	51
Table 7. Results of initial, repeated-measures analyses of covariance	55
Table 8. Descriptive statistics for participants who had either guided or unguided planning time during the first story-telling condition	56
Table 9: Descriptive statistics for participants who had either guided or unguided planning time during the second story-telling condition	57
Table 10. Results of ANCOVAs for all participants on Task 3 (the first story-telling task) only.....	61
Table 11. Results of repeated measures ANCOVAs divided by the order of the provision of planning time	67
Table 12. Structural complexity (Clauses per T-unit) in terms of planning order and planning conditions.....	68
Table 13. Comparisons of participants' planned and unplanned performances according to planning conditions (guided/unguided) and planning order (first/second)	69
Table 14. Groups according to learners' WMC.....	75
Table 15. Structural complexity (Clauses per T-unit) in terms of WMC and planning conditions.....	81

List of Figures

Figure 1: Corrections per 100 words under guided planning, no planning, and unguided planning conditions	59
Figure 2: Pruned speech rate under guided, unguided and no-planning conditions. ..	60
Figure 3: Differences in the number of self-corrections per 100 words within subjects between planning conditions (planned and unplanned) and planning groups (guided and unguided) when planning time was provided during the first story-telling task..	64
Figure 4: Differences in pruned speech rates within subjects between planning conditions (planned and unplanned) and planning groups (guided and unguided) when planning time was provided during the first story-telling task.	65
Figure 5. Mean z-scores for dependent accuracy, fluency, and complexity variables under guided and unguided planning conditions when planning time was offered first.	70
Figure 6. Mean z-scores for dependent accuracy, fluency, and complexity variables under guided and unguided planning conditions when planning time was offered for the second story-telling task.....	71
Figure 7. Mean z-scores for dependent accuracy, fluency, and complexity variables under no-planning conditions when planning time was offered for the second story-telling task. Learners in both groups completed this unplanned task as their first one in the sequence of telling two different stories.	72
Figure 8. Mean z-scores for dependent accuracy, fluency, and complexity variables under no-planning conditions when planning time was offered for the first story-telling task. Learners in Group 1 completed the story-telling task prior to this unplanned task with guided planning time whereas learners in Group 2 completed the task prior to this one with unguided planning time.....	73
Figure 9. Mean self-corrections per 100 words according to WMC groups under guided planning and no-planning conditions when guided planning was offered first.	76
Figure 10. Mean self-corrections per 100 words according to WMC groups under guided planning and no-planning conditions when guided planning was offered during the second story-telling task.	77
Figure 11. Mean self-corrections per 100 words according to WMC groups under unguided planning and no-planning conditions when unguided planning was offered during the first story-telling task.....	78
Figure 12: Mean self-corrections per 100 words according to WMC groups under unguided planning and no-planning conditions when unguided planning was offered during the second story-telling task.	78
Figure 21, Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under guided conditions when planning time was offered first.....	83
Figure 22. Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under guided conditions when planning time was offered second.....	84

Figure 23. Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under unguided conditions when planning time was offered first.....	85
Figure 24. Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under unguided conditions when planning time was offered second.....	86
Figure 25. Mean z-scores for all dependent measures according to WMC group when learners had no time to plan for the first story-telling task.....	87
Figure 26. Mean z-scores for all dependent measures under no-planning conditions according to WMC group when learners had guided planning for the first story-telling task.....	88
Figure 27. Mean z-scores for all dependent measures under no-planning conditions according to WMC group when learners had unguided planning for the first story-telling task.....	89
Figure 28. Percentage of errors due to lexis, plural nouns, and subject-verb agreement when participants were offered guided planning time.....	90
Figure 29. Percentage of errors with lexis, plural nouns, and subject-verb agreement when learners completed stories with unguided planning time.....	91
Figure 30. Percentage of errors with lexis, plural nouns, and subject-verb agreement when learners had no time to plan.....	92

Chapter 1: Review of the Literature

Working memory, a psychological construct that accounts for how individuals simultaneously process and retrieve information, underpins explanations for complex cognitive behaviors, including second language acquisition (SLA). The executive component of a person's working memory capacity (WMC) measures how much information can be processed at the same time in the face of interference, and high WMC is correlated with the acquisition of syntactic and vocabulary knowledge (Kormos & Sáfár, 2008), performance on second language reading tasks (Harrington & Sawyer, 1992), language aptitude scores (Robinson, 2002), and fluency and lexical complexity during second language speaking tasks (Gilabert & Muñoz, 2010). There is empirical evidence to suggest that learners with higher WMC can benefit more than learners with lower WMC during specific instructional treatments, for example, by demonstrating more noticing during the focus on form technique of recasting (Mackey et al., 2002) or by benefiting more from studying abroad (Sunderman & Kroll, 2009). Working memory is one of the core components comprising language aptitude, central to understanding language acquisition (Dörnyei, 2005; Robinson, 2005c; Sawyer & Ranta, 2001).

And while there is an increasing body of research-based principles for instructed SLA, including, for example, recommendations to focus on form (Doughty & Williams, 1998; Long, 1991) and provide opportunities for output and interaction (Gass, 1997; Izumi, 2002; Long, 1981, 1996; Swain, 1995), there is very little information on how to tailor instructional treatments to aspects of language aptitude

(Robinson, 2005a; Robinson, 2007). For this reason, the interaction of pedagogical interventions with individual difference variables, such as WMC, is a logical place for empirical SLA research, both to investigate the cognitive processes underpinning L2 performance and to identify instructional treatments that may compensate for low WMC. After a brief review of the existing research on WMC and SLA, the remainder of this chapter will discuss a specific instructional strategy that has shown promise in this regard—the provision of pre-task planning time. Research on planning time is part of a large body of empirical inquiry on the effects of manipulating task conditions to promote changes in the accuracy, complexity, and fluency of learner discourse. Hypotheses about SLA and cognition, for example, the Trade-Off Hypothesis (Skehan, 2009) and the Cognition Hypothesis (Robinson, 2001, 2005c, 2011), are generally used to motivate these studies, yet there has been very little research on the interaction of cognitive individual difference factors (such as WMC) with task conditions (but see Guar-Tavares, 2009; Kormos & Trebits, 2011). Previous research on planning time and other methods of influencing the cognitive complexity of second language tasks inform the framework for the empirical study presented in the remainder of the dissertation.

1.1 Working Memory

Most second language research involving working memory relies on Baddeley and Hitch's (1974) model, which originally consisted of three components: the phonological loop, the visuo-spatial sketchpad, and the central executive. There have been a number of SLA studies focusing on the phonological loop, which is the portion of working memory responsible for processing auditory information. For

example, O'Brien et al. (2006) investigated the role phonological memory plays in the SLA processes of adult learners, implicating phonological working memory in various stages of SLA depending upon learners' proficiency levels. Williams and Lovatt (2003) demonstrated that there is a positive correlation between phonological memory and grammatical rule learning, with both Italian and an artificial language. There has also been empirical SLA research on the central executive component of working memory, which "is assumed to be responsible for the attentional control of working memory" (Baddeley 2003, p. 201). Many of these studies have relied upon permutations of the reading span test, originally developed by Daneman and Carpenter (1980), which requires learners to read sets of sentences, determine whether or not they are true or false, and then recall the last word of each sentence in the set. For example, Harrington and Sawyer (1992) found a relationship between the reading span task in English and reading comprehension in English for advanced L1 Japanese learners. Payne and Ross (2005) considered both phonological memory and the central executive component in a study of computer-mediated communication, and found a correlation between phonological memory and L2 gains and performance.

Baddeley (2000) updated his original, three component model to include what is called the "episodic buffer," a mechanism that is

principally concerned with the storage of information rather than with attentional control. It is capable of binding together information from a number of different sources into chunks or episodes, hence the term "episodic"; it is a buffer in the sense of providing a way of combining information from different modalities into a single multi-faceted code. Finally, it is assumed to underpin the capacity for conscious awareness (Baddeley 2003, p. 203)

This component was added in order to explain how input from various sources (e.g., long-term storage and the phonological loop) could be processed simultaneously. In terms of language, the episodic buffer might account for how linguistic input from an interlocutor could be compared with knowledge already stored in long-term memory, which is what happens throughout the various cognitive processes underlying the SLA process (e.g., “noticing”¹). It is also likely that the episodic buffer is necessary for manipulating language, which requires more time than is available in the phonological loop before the information held there begins to decay. For example, the episodic buffer might hold information transferred from the phonological loop and be the staging-point for comparison with previously learned structures or rules. While there has been little empirical research on how the episodic buffer might fit into accounts of SLA, Christoffels (2006) suggests that the ability to transfer linguistic input from the phonological loop to the episodic buffer quickly could account for the high-level language abilities demonstrated by professional interpreters.

Baddeley and Hitch’s multi-component model of working memory (1974) provides the framework for most SLA research to date, but there are other models of working memory that are compatible with investigations of the cognitive processes responsible for language learning. These models tend to focus on processing rather than storage, so they are similar to the central executive component of Baddeley and Hitch’s model. For example, Cowan (1999) interprets working memory as

¹ Noticing was first introduced to SLA in 1990 by Schmidt when he hypothesized that conscious knowledge through noticing was necessary for language learning. In 1995, Robinson updated the definition as “detection with awareness and rehearsal in short-term memory...necessary for learning and the subsequent encoding in long term memory” (1995a, p. 318). Most recently, Schmidt has

overlapping subsets of three components: long-term memory, the portion of long term memory that is “activated,” and the portion of activated memory that is the focus of attention. In this model, WMC is limited in that activated memory is constrained by time, and the focus of attention is limited by cognitive resources (p. 62). Engle, Kane, and Tuholski (1999) also define working memory as a combination of long-term memory, activated long-term memory, and attention, and they specify that WMC “is not really about storage or memory per se, but about *the capacity for controlled, sustained attention in the face of interference or distraction*” (p. 104). These theorists have updated the original Baddeley and Hitch (1974) model of working memory to make it “more relevant to complex cognition in general, focusing directly on the role of the central executive in maintaining task-relevant information” (Williams, 2012, p. 428).

More recently, Unsworth and Engle (2007) have examined individual differences in WMC in terms of a framework that “combines a flexible attentional component with a cue-dependent search mechanism” (p. 104).² Their model relies on the concept of an activated primary memory, with secondary memory providing long-term storage. They claim that because what a person can hold in primary memory at any one moment is limited, there must be a retrieval component to working memory that searches long-term memory, and that “to retrieve task-relevant information, a discrimination process is needed to differentiate between relevant and irrelevant information” (p. 105). In their model, the source of individual differences in WMC stems from the efficiency of the retrieval component; individuals with low WMC

² Unsworth and Engle indicate that their model has more in common with older views of memory, but that it shares theoretical underpinnings with, for example, Baddeley’s notion of the episodic buffer.

have more trouble searching long-term memory effectively, causing them to have trouble with the higher-order cognitive tasks (including language acquisition) at which individuals with higher WMC excel (p. 108).

1.2 Working Memory and Stages of SLA

It is important to note that individuals with low WMC are not uniformly poorer at all tasks than those with high WMC; working memory does not play a role in tasks that are automatic in nature when there is no contextual interference (Kane et al., 2007; Unsworth & Engle, 2007). That is, individual differences in WMC “typically arise in situations where information needs to be actively maintained or when a controlled/strategic search of memory is required to retrieve task-relevant information” (Unsworth & Engle, 2007, p. 123). This distinction is applicable to SLA because while *acquiring* a second language is a higher-order cognitive task that requires both active maintenance of input, as well as strategic searching of long-term memory, language *use* becomes much more automatic after a language has been acquired. In other words, working memory would be critical in the beginning and intermediate stages of acquisition, but might not necessarily play as large a role in fluent speaking in really advanced L2 speakers because the initial cognitive demands in L2 acquisition are significantly greater than what is required for practiced, fluent speech.

Following this logic, one would expect to see WMC contribute less to the performance of higher-proficiency second language learners whose speech is more automatic. This has been discussed to some extent in the literature on WMC and SLA. Mackey et al. (2002) conducted a small-scale study of the relationship between

working memory, noticing, and interlanguage development of ESL students in Japan. The data indicate that learners with higher WMC, as measured by a listening span task in both the L1 and the L2 and a non-word recall task, reported more noticing. Noticing was operationalized as “learner’s articulation of response to the input” (p. 188) and was collected during retrospective interviews as well as exit questionnaires. In addition, the study shows that learners at lower proficiency levels reported more noticing than those who were further along the acquisition process. In other words, the relationship between WMC and acquisition was more pronounced for lower proficiency learners. There is a similar suggestion in O’Brien et al.’s (2007) study comparing adult learners of Spanish in a study abroad context with those in a traditional classroom setting. WMC was measured with a serial non-word recognition task, and high performance on that test was found to correlate with oral proficiency gains. In addition, the researchers suggest that the link between working memory and language learning was related to stage of language development. This correlation suggests that WMC might be more important for early and intermediate stages of language acquisition, but might have less of an impact on language use in more advanced (and, therefore, more proficient) L2 speakers.³

Another study that considered learner proficiency level and WMC is Payne and Whitney’s (2002), which investigated the ability of Spanish students to improve

³ Gilabert and Muñoz (2010) argue the opposite in their study, which found correlations between learner’s oral accuracy, fluency, and complexity measures and WMC for high proficiency learners but not for learners with low proficiency. They argue that for WMC to make a difference in performance, some degree of proficiency must be reached, and at lower proficiency levels, “learners may need to be using considerable attentional and memory resources to access and retrieve words, and to give messages their syntactic and morphophonological shape, to a point that differences in working memory may not matter,” (p. 37). While it is possible that WMC cannot differentiate among complete novices struggling to string words together, this seems unlikely for intermediate-level learners who can communicate with some degree of fluency. Further, it runs counter to the idea that one relies less on WMC for automatized tasks (Unsworth & Engle, 2007).

their oral proficiency skills over the course of a semester through computer-mediated communication (CMC). The authors targeted third-semester (lower level) Spanish students because “the demands placed on working memory by less fluent [second language] speakers may differ qualitatively, and, most likely, quantitatively, from more fluent [second language] speakers” (p. 13). The study was intended to determine whether or not CMC could help compensate for low WMC. Their hypothesis was that since CMC lessens the cognitive burden of communication (because students have more time to retrieve interlanguage knowledge) it would allow them to have better output practice. This output practice would then lead to more fluent oral proficiency at the end of the semester, even for those students with low WMC. The researchers found that individuals with low WMC who participated in the experimental CMC treatment outperformed those with low WMC who participated in the traditional classroom control group on the oral proficiency exit test. In addition, in a follow-up study with the same data, Payne and Ross (2005) found that learners with low WMC produced more output during the CMC sessions than higher WMC learners did, suggesting that the reduced cognitive burden permitted students to generate more language.

1.3 Working Memory and Task Complexity in SLA

The results of Payne and Whitney’s (2002) study are encouraging because, as the authors point out, “the indication that learning environments can, by design, reduce the burden on working memory and thereby produce a facilitating effect for low capacity individuals offers a new perspective on how instruction can meet the

individual needs of learners” (p. 26).⁴ Modifying task conditions to make language learning tasks more or less difficult is not a new idea (Gilabert, 2007a; Robinson, 2001, 2005a; Skehan & Foster, 1999, 2001),⁵ and one promising pedagogical technique is the provision of planning time. In his (2005a) review of aptitude and SLA, Robinson states that providing students with planning time before or during a task is a pedagogical intervention that can affect fluency, and, ultimately automatization (p. 57). This intervention might be addressing the very issue presented here; that is, if planning time positively affects fluency, then perhaps it is because it helps learners with the retrieval of previously stored information. Time to plan before a speech act would give learners time to retrieve structures or concepts before actually needing them in “online” communication, making it especially beneficial for learners with low WMC.⁶ According to Ortega (1999), “planning may lessen the cognitive load of a given task and free up attentional resources at the micro levels of speech production” (p. 110). This pre-task retrieval during planning might prime the structures/concepts in long-term memory so that they can become more easily accessed during the performance of the actual task. Following the models that view variations in WMC as individual differences in executive control (Engle et al., 1999; Kane et al., 2007; Unsworth & Engle, 2007), learners who are less efficient

⁴ There is also a growing body of research examining the ability of cognitive training tasks to increase WMC, and, therefore, improve performance in other cognitive domains (see, e.g., Morrison & Chein, 2011, for a review of the literature).

⁵ It is important to address the term “task” as it has been used in the literature on WMC, as well as how it is used throughout the SLA literature. Cognitive psychologists (e.g., Engle et al., 1999; Kane et al., 2007; Unsworth & Engle, 2007) use the word to refer to both the psychological tests used to measure the effects of a low WMC and to the everyday activities that are affected by an individual’s working memory. With respect to SLA, tasks are the unit of analysis in a specific pedagogical approach, Task-Based Language Teaching (TBLT) (Doughty & Long, 2003; Long & Crookes, 1993). Within the framework of TBLT, second language learner performance on various tasks has been the subject of significant research, e.g., by studying the effect of modifying task conditions to make them more or less difficult (Gilabert, 2007; Robinson, 2001, 2005c).

⁶ See Skehan and Foster (2001, pp. 199 - 204) for further speculation on how planning might affect cognition in a second language.

searchers would be able to find and access the information they needed during the actual language task more quickly after having had time to plan because they would have just finished finding it while planning.

There is evidence from research in L1 contexts that indicates a link between general planning ability and WMC. For example, many studies that have investigated planning ability as measured on psychological planning tasks have found correlations between WMC and planning ability (Gilhooly et al., 2002; Newman et al., 2003). Similar research has found that the ability to plan is negatively affected when WMC is impaired, for example, in individuals with frontal lobe damage (Owen et al., 1990) or in schizophrenic patients with reduced WMC (Badcock, Michie, & Rock, 2005). Because planning ability is something that appears to suffer when WMC is reduced or impaired, it is possible that the provision of pre-task planning time might interact with WMC.

1.4 Planning and SLA

There have been many empirical investigations of the effects of pre-task planning time on learners' accuracy, fluency, and complexity (Kawauchi, 2005; Mehnert, 1998; Ortega, 1999 & 2005; Yuan & Ellis, 2003).⁷ In her 1999 synthesis of pre-task planning studies, Ortega points out the following: "(a) planned output is both more fluent and more syntactically complex than unplanned output, and (b) results for grammatical accuracy are conflicting and inconclusive" (p. 118). That is, planning allows for more fluent speech, which is a predictable result, given that

⁷ See Ellis (2009) for a synthesis of planning studies, including a helpful table (pp. 480 – 490) reviewing studies of different types of pre-task planning, the operationalization of variables in planning studies, and the results.

students have time to organize their interlanguage before using it. The mixed effect of planning time on learner accuracy is possibly because planning time cannot compensate for gaps in competence; in other words, if L2 learners do not have solid mental representations of linguistic structures, pre-task planning time cannot facilitate more accurate use of them. Skehan (1998, 2009) has suggested that if attentional capacity is limited, learners focus on complexity and fluency at the expense of grammatical accuracy, and that accuracy and complexity are in competition with one another for attentional resources.

Learners' L2 proficiency is something that has been considered in the planning literature to some extent, which is important since the model of SLA discussed thus far suggests a greater reliance on the search and retrieval aspects of WMC for lower-level learners who have not yet automatized the target language. One would expect, then, to see lower-level learners benefit more from planning time. Ortega (2005) discussed the results of two different planning experiments. Both experiments considered the effect of open-ended planning time on the performance of a picture-guided story re-telling task. She found that for the low-intermediate students, pre-task planning resulted in greater lexical complexity, but for advanced learners, there was no difference in lexical complexity for students who received time for pre-task planning. These findings fit neatly into the model proposed here. If providing planning time helps compensate for low WMC, and if the advanced learners did not need to rely on working memory to search for lexical items because their retrieval was largely automatized, then the pre-task planning would probably not have been as beneficial for them.

Ortega (2005) also found that advanced learners during the planning condition displayed accuracy improvements while the low-intermediate learners did not. This could be a difference related to level of language ability and the development of interlanguage skills. The more advanced learners could have been developmentally ready to improve their grammatical accuracy because they had acquired more grammatical knowledge, whereas the intermediate learners had not yet reached that point. Another relevant study is Yuan and Ellis (2003), which investigated the results of three different planning conditions (no planning, pre-task planning, and online planning) on accuracy, complexity, and fluency of a picture-guided narrative. Their participants were undergraduate English majors in the International Business Department of a Chinese university. All students had studied English as a foreign language for eight years prior to enrolling at the university, and, as college students, they had six hours of English each week. The students were certainly as advanced in English as the ones Ortega (2005) classified as advanced in Spanish in her experiment. While the focus of this study was not on the interaction of language level with planning treatments, the high proficiency level of the subjects is interesting, given the results. Yuan and Ellis found that the pre-task planners were no more fluent than the ones who did not receive planning time. This conflicts with previous findings on pre-task planning, and the authors suggest that perhaps it is because the students were asked to complete the narrative task under time pressure. This could, of course, be a factor, but it could also be that for higher-level language students, the time to plan before the task was not helpful for fluency. The pre-task planning condition did affect accuracy and syntactic complexity, which is similar to Ortega's

results for the advanced students. Again, this could be because level of language ability interacts with the effects of pre-task planning (Crookes, 1989).

1.5 Working Memory and Planning in SLA

In his (2005c) paper “Cognitive complexity and task sequencing,” Robinson categorizes planning as a resource-dispersing task dimension, meaning that taking planning time away from a task makes the task more difficult because cognitive resources are then dispersed. This theory might explain how planning time could help students organize their mental resources. He goes on to say that

increasing task demands has the effect of gradually removing processing support (such as planning time) for access to current interlanguage, and thus practice along these dimensions requires and should encourage faster and more automatic access and use. (p. 24)

Again, this explanation of how the cognition of task planning might work fits into the model proposed here. If students with low WMC—because of problems with search and retrieval during task performance—have more difficulty in terms of complexity and fluency in the early stages of SLA than their peers with higher WMC, then planning might give them the processing support they need to perform more comparably. Improved performance in the target language should permit the learners with low WMC to move through the language acquisition process more efficiently so that their performance eventually becomes more fluent and automatized.⁸ In other words, learners require a great deal of practice before their second language production becomes automatic and fluent (DeKeyser, 2010, pp. 130 – 131;

⁸ See Skehan (2002, pp. 88 - 89) for a suggested timeline of the stages of second language acquisition, going from noticing to automatizing/achieving fluency. He points out that this is not a universally accepted model of development, but that it draws on modern notions of language and cognition, which are similar to those considered so far.

Segalowitz, 2010, pp. 76 – 77), and if instructional treatments can be tailored to learners' individual differences to maximize the accuracy, complexity, and fluency of their discourse, this process of practice and skill acquisition will be more effective. This consideration of how an individual difference variable (WMC) dovetails with pedagogical treatments intended to manipulate learner performance to facilitate acquisition is a logical extension of the research on task conditions and learner performance, which, as Gilabert (2007a) points out, has “been concerned with how a balanced approach in the three areas of production can potentially lead to more effective language use and acquisition” (p. 45).

Skehan and Foster (1999) suggest the need for research along these lines in the conclusion of their study on the influence of task structure and processing conditions. They state that during pre-task planning, “participants can attempt to foresee what language or content organization will be required for the task. This can then be assembled in working memory and the attempt can be made to draw on it while the task is being done” (115).⁹ Their study does not consider individuals with low WMC as a separate group, but it follows from their explanation that if planning allows time for the pre-assembly of language so that it can be imported into working memory during online processing, it might be especially helpful for those with trouble searching and/or retrieving from long-term memory. The need for more investigation of SLA and specific aspects of working memory is clear. According to Skehan (2002), we should be examining the processes that “are more concerned with storage, and especially retrieval. The other work on memory within

⁹ The authors do not define the model of working memory they are using here, but the notion of “assembling” language in working memory would fit the search and retrieval model, as well as the idea of Baddeley’s episodic buffer as a workspace.

aptitude[...]indicates promise but does not address the area of retrieval of memory, or memory organization in an effective manner” (p. 92).

Though the connection between pre-task planning and overcoming constraints on L2 acquisition from low WMC seems intuitive, the relationship between these two constructs has been the subject of just two empirical investigations. The first was conducted by Guara-Tavares (2009) in order to consider the effect of planning time on the accuracy, fluency, and complexity of 25 intermediate Brazilian EFL learners during two picture-guided narratives. Using a control group design, she found an effect for planning time on accuracy, $F(1, 23) = 25.74, p = .00$ and fluency, $F(1, 23) = 10.371, p = .00$. In addition, she found correlations between working memory and fluency, $r(12) = 0.588, p < .05$ in the unplanned condition, and between working memory and fluency, $r(12) = .585, p < .05$ and accuracy, $r(12) = .916, p < .01$ in the planned condition. However, it is important to note that because Tavares used a variant of the speaking span task (Daneman & Green, 1986) in the learners’ L2, the correlations between WMC and the provision of planning time might be tangled up with learners’ overall proficiency in English. In other words, in addition to (or, potentially instead of) measuring the learners’ WMC, the instrument was likely measuring their L2 speaking proficiency. However, as Tavares points out at the end of her paper, “despite its limitations and lack of conclusive evidence, the findings of the present study may be relevant since they seem to demonstrate that the relationship between WM capacity, planning and L2 speech performance is a complex one which merits future scrutiny” (p. 189).

Nielson (in press) conducted a similar experiment with 46 students enrolled in intermediate ESL classes at a community college in the United States. She hypothesized that pre-task planning time would benefit learners with low WMC more than learners with high WMC. To operationalize WMC, she used a spatial working memory span task so that language proficiency would not affect the working memory measure. The results of the study did not support the original hypothesis, as planning time was equally beneficial (with respect to fluency and complexity) for learners with both high and low WMC. However, despite the lack of interaction between the aptitude variable (WMC) and the pedagogical treatment (the provision of planning time), the study did demonstrate that planning time can help learners regardless of WMC. In addition, this study confirmed the findings of previous research on SLA and working memory, with an effect for WMC on fluency and structural complexity both with and without pre-task planning time.

One unexpected finding in Nielson (in press) was an effect for the *order* of the provision of planning time. As in previous planning studies (Ortega, 1999, 2005; Gilabert, 2007a), that experiment employed a counter-balanced, repeated-measures design, so each student completed a monologic oral story-telling task with pre-task planning time as well as a different story without any pre-task planning time. The study included two different picture-guided narratives, with the order of the provision of planning time reversed for half the group. Nielson (in press) found that learners who had planning time under the first story-telling condition demonstrated better syntactic complexity and fluency under the subsequent, unplanned condition. In

other words, the planning time appeared to have a beneficial carryover effect to the unplanned condition.

A possible explanation for the carryover effect identified by Nielson (in press) might be that time to plan before producing output likely gave learners a chance to prepare by facilitating the assembly of language and structures for later retrieval (Skehan & Foster, 1999), but it also probably allowed learners to retrieve and rehearse strategies for more fluent speaking (either explicitly or implicitly). This preparation time may have allowed them to notice areas where their planned performance deviated from their actual performance. Then, in the second, unplanned story, the learners were able to improve their performance even more, despite having different narrative content, because they were prepared to improve aspects of the speech with which they were unhappy (e.g., pauses, false starts, filler words, increased subordination, etc.). According to the Output Hypothesis, learners must be pushed to produce output in order to develop L2 competence because it is only through attempting to produce the language that some gaps in interlanguage competence become noticeable (Gass, 1997; Swain, 1995; Izumi, 2002, 2003). It may be the case that the pre-task planning time primed learners to notice the problems with their performance more during production, making them better able to improve upon them, even in the unplanned condition. This account of how planning time may have had a facilitative effect on subsequent, unplanned discourse is in line with explanations for why output practice in general is beneficial in SLA. As Muranoi (2007) states, output practice “promotes major cognitive processes in SLA, including

noticing, hypothesis formation and testing, conscious reflection, and automatization” (p. 76).

1.6 Manipulating Planning Time

Missing from the discussion so far is a consideration of what learners actually *do* with planning time. Based on analyses of the notes learners take during planning sessions, and on retrospective interviews and post-planning questionnaires, Ortega (1999, 2005), Sangarun (2005), and Yuan and Ellis (2003) suggest that the differential effects for accuracy, fluency, and complexity found throughout studies of pre-task planning might be due to learners’ interpretations of the task requirements. Further, as Park (2010) points out, different planning studies have provided different instructions for how to use planning time, potentially creating a confound between the effect of the instructions and the planning time itself (p. 10). For example, some researchers specifically ask learners to plan their narratives in terms of content, organization, and language (Crookes, 1989; Skehan & Foster, 1997; Yuan & Ellis, 2003), while others simply tell participants to use the planning time any way they want (Nielson, in press; Ortega, 1999; Wigglesworth, 1997).

There has been some research on the results of guided or structured planning time when compared to unstructured planning time. The research on guided planning time is generally concerned with using the time to foster accuracy in the L2, which is the aspect of L2 production that is least predictably affected by unstructured pre-task planning time. For example, Foster and Skehan (1996) conducted an early study on the differential effects of both guided and unguided planning time, with guided planning operationalized as suggestions to the learners to think about “how they

might use the 10 minutes to consider the syntax, lexis, content, and organization of what they would say” (p. 307), and unguided planning time operationalized as simply providing learners with time to plan. They found differences in the fluency and complexity of learners’ output between all three planning conditions (guided, unguided, and no-planning), with the guided planning generating the most complex and fluent language. However, as Ortega (1999) points out, the difference between the guided planning condition and the unguided planning condition was minimal, and, in fact, guided planning as operationalized by Foster and Skehan (1996) is nearly identical to the unguided planning operationalized by Crookes (1989) (p. 113).

In a subsequent study, Foster and Skehan (1997) contrasted the effects of teacher-fronted planning, solitary planning, and group planning before a group-based, decision-making task. Unlike in Foster and Skehan (1996), the planning conditions in this study were quite different from one another. Teacher-fronted planning was operationalized as a pre-task, teacher-led discussion of either the content necessary for a task or the specific structures necessary for the task; solitary planning was operationalized as giving learners time to plan alone before completing the group task; and group planning was operationalized as giving students time in groups to plan their decision-making task. The researchers found a facilitative effect for both teacher-fronted planning and solitary planning, with solitary planning producing the most complex and fluent speech, and teacher-fronted planning producing the most accurate speech with complexity similar to that generated by solitary planning. As the authors point out, it is not surprising that teacher-fronted planning produces more accurate speech as the fluent speaking teachers are providing a model for learners to

draw on during their performances. However, the researchers further divided the teacher-fronted planning into two modes (structure-focused or content-focused) and both modes had the same (high) effect on accuracy and complexity. As Foster and Skehan (1997) indicate, during even the content-focused planning sessions, instructors provide models of correct and complex language, which then become available to the learners during their performances (p. 241).

Two more recent studies have examined the effects of guided planning time with English as a Foreign Language (EFL) learners in high schools. Sangarun (2005), using a population of seniors in a Thai high school, considered whether or not the provision of meaning-focused, form-focused, and form-and-meaning-focused planning instructions could affect learner output in terms of accuracy, complexity and fluency on monologic argumentative tasks. Sangarun developed careful guidelines for the operationalization of form- and meaning- focused planning, ultimately concluding that it is possible to manipulate how learners use the planning time through specific instructions, and that form-and-meaning-focused planning resulted in a good balance of accurate and complex speech. In a similar study, Mochizuki and Ortega (2008) investigated whether or not pre-task planning with specific grammatical guidance would benefit lower-level EFL learners in a Japanese high school. They divided learners into three conditions: unguided pre-task planning, guided pre-task planning, and no planning. Unlike previous considerations of guided pre-task planning that used vague instructions (e.g., Foster & Skehan, 1996), teacher-fronted modeling (Foster & Skehan, 1997), or explicit written instructions (Sangarun, 2005) as the guided planning conditions, this study provided learners with a

worksheet explaining how to create sentences with relative clauses (including model sentences) and told them that the grammatical explanations might be useful as they prepared to complete the target story-telling task. The results of this study showed an effect on accuracy for those learners who had the worksheet during the pre-task planning time, who had greater gains in accuracy compared to the no-planners and the unguided planners. In other words, this study confirmed that learners' accuracy during oral, monologic story-telling tasks can be manipulated with explicit, form-focused pre-task planning time.

The studies that showed an effect for guided planning time on learners' accuracy (Foster & Skehan, 1997; Mochizuki & Ortega, 2008; Sangarun, 2005) used very explicit means of promoting accurate performances (outright modeling of the target form, a worksheet with grammatical information on how to form sentences, or carefully developed guidelines for how to use the planning time) while the studies of unguided planning that do not show an effect for accuracy tended to provide absolutely no instructions for how to use planning time. This suggests, then, that while the lack of an effect for accuracy in the more general planning studies may be due, in part, to gaps in learners' linguistic competence, some of the persistent inaccuracy ratings (even with the provision of planning time when complexity and fluency increase) are due to learner-internal decisions about the allocation of attention (Skehan, 1998, 2009). If the default mode for communicative oral tasks is to focus on content over form (as suggested by, for example, Van Patten (1990, 1996)), then further research is needed to determine the most efficient way to manipulate task conditions to promote learners' focus on accuracy (e.g., via detailed planning

instructions and specific parameters for the task performance) without losing the benefits for complexity and meaning.

1.7 Cognition and Task Conditions

It is important to point out that there are two competing theories of cognition and task complexity that have been employed to account for the differential effects of task conditions on learner accuracy, complexity, and fluency. Skehan (1998, 2009) argues, with his Trade-Off Hypothesis, that attentional capacity is limited; fluency, accuracy, and complexity are in competition with one another; and studies that show an effect for both accuracy and complexity simultaneously are able to do so by reducing task complexity either by carefully manipulating planning time, as in Foster and Skehan (1997), or by manipulating task requirements, for example, by requiring learners to incorporate both foreground and background information in a precisely structured narrative task (Skehan, 2009, pp. 521 - 522). On the other hand, Robinson's (2001, 2005c, 2011) Cognition Hypothesis argues that learners can draw on multiple cognitive resources simultaneously and that fluency, accuracy, and complexity are not necessarily in competition with one another. Within Robinson's framework, it is possible to promote learner focus on both accuracy and complexity by *increasing* the cognitive demands of a task along what he refers to as resource-directing dimensions.

The Cognition Hypothesis makes a distinction between the effects of resource-dispersing task conditions (e.g., removing the support of pre-task planning time), which, as discussed throughout, does not demonstrate uniform improvements in the accuracy of learner output, and the effects of resource-directing modifications.

According to Robinson (2011), increasing the cognitive requirements of a task along resource-directing dimensions, by, for example, having learners narrate a story in the past tense without any visual cues, “should push learners to greater accuracy and complexity of L2 production in order to meet the consequently greater functional/communicative demands they place on the learner, while negatively affecting fluency” (p. 18). Support for the Cognition Hypothesis can be found in Robinson’s (1995b) study of “There-and-Then” narrative tasks, which demonstrated that learners displayed both better accuracy and more lexical complexity under the more complex “There-and-Then” task condition than they did under the less complex “Here-and-Now” version of the same narrative task (1995b). Gilabert (2007a) investigated the effects of manipulating task conditions along resource-dispersing dimensions (the provision of pre-task planning time) as well as resource-directing dimensions (narrating tasks in the “Here-and-Now” versus “There-and-Then”). He demonstrated that when learners were provided with unguided pre-task planning time, they improved their fluency and lexical complexity, but not their grammatical accuracy, confirming the results of previous planning studies, as well as an additional claim of the Cognition Hypothesis that planning time will not, in and of itself, direct learner resources to accuracy. This same study also offers support for the resource-directing nature of “There-and-Then” task conditions, demonstrating that learners’ accuracy and complexity improved under the more complex tasks.

Both the Trade-Off Hypothesis and the Cognition Hypothesis claim that fluency is negatively affected when processing demands are raised (whether by imposing resource-dispersing or resource-directing task conditions), and this is

supported by studies of planning time as well as other task modifications.

Researchers within both theoretical frameworks have demonstrated that learners can be pushed to focus on accuracy and complexity simultaneously, whether through pre-task planning guidance or a manipulation of task conditions. And while both Skehan (2002) and Robinson (2005c, 2007, 2011), as well as Ellis (2005, 2009, 2012) and Gilabert (2007a, 2007b), call for studies to examine the interaction between individual differences, such as working memory, and task conditions, there is still very little research in this area.

In addition to the two studies of pre-task planning time and WMC discussed earlier (Guará-Tavares, 2009; Nielson, in press), there has been just one other study investigating the interaction of WMC and task complexity. Kormos and Trebits (2011) conducted research with ESL learners in a Hungarian high school who were asked to take a backward digit span test of WMC and then complete two different oral, monologic narratives under conditions hypothesized to affect task complexity. The researchers found that learners with high WMC were able to produce longer clauses during the less complex narrative task, but that in general WMC did not have a linear relationship with learner output or task complexity. They point out that the two tasks they used were actually quite different from one another in their requirements and that more research is needed on the interaction of WMC with more controlled manipulation of task complexity (p. 281).

It is clear that more research is needed on the relationship between WMC and task conditions. Research along these lines will offer critical information related to how individual difference variables interact with pedagogical treatments, allowing

instructors to make instructional choices to benefit learners according to their WMC. Further, understanding how WMC interacts with both resource-dispersing and resource-directing task conditions could shed some light on the cognitive processes underlying the effects of manipulating task conditions. If, as proposed by Skehan (1998, 2002, 2009), attention is limited in capacity, and a simultaneous focus on accuracy and complexity can only happen by simplifying task conditions, then learners with both high and low WMC should be able to focus on both accuracy and complexity when task conditions are simplified. If, on the other hand, learners can be directed to focus on both accuracy and complexity with tasks made more complex along resource-directing parameters, learners with high WMC should be able to focus simultaneously on both accuracy and complexity better than learners with low WMC, who will have more limited cognitive resources with which to attend to form during more complex task conditions.

Chapter 2: Purpose of the Study

Chapter One outlined the literature motivating the empirical study presented in the remainder of this dissertation. To recap, “working memory span has proved to be a robust predictor of a wide range of complex cognitive skills,” including the skills necessary to acquire a second language (Baddeley 2003, p. 202). Cognitive psychologists researching WMC have suggested that individual differences in WMC reflect the efficiency with which people can retrieve previously stored information (Unsworth & Engle, 2007). Further, there is likely a relationship between a learner’s proficiency level and, therefore, stage of SLA, and the role his/her WMC plays in the acquisition process (Kormos and Sáfár, 2008; Mackey et al., 2002; Payne & Whitney, 2002).

There is also evidence to suggest that changing the conditions under which pedagogical tasks are completed can facilitate learner performance, potentially compensating for the limits of WMC (by allowing more time for search and retrieval) during a time when it is critical to acquisition (Payne & Whitney, 2002; Nielson, in press). Improving learner performance during practice so that different aspects of L2 production are emphasized is one way instructors can help with the proceduralization and automatization of L2 production skills (Muranoi, 2007). However, because these pedagogical treatments may affect language processing, it is important to understand how they interact with WMC (Robinson, 2005a, p. 57). According to Robinson (2007), “as L2 tasks increase in complexity, [individual differences] in cognitive

abilities increasingly differentiate performance,” and there is a need for further research to address the interaction of task complexity and individual difference variables. In addition to Robinson (2007, 2011), other theorists investigating instructed SLA have called for more research that considers the cognitive processes underlying performance during different task conditions (Ellis, 2009; Gilabert, 2007a; Ortega, 2005; Skehan, 2002). The study presented here seeks to expand this specific research agenda.

Pre-task planning time has been shown to improve the fluency and complexity of oral narrative discourse for learners with both high and low WMC when compared to the same monologic tasks performed without planning time (Nielson, in press). However, there has been no research on more complex oral narrative tasks, such as those told in the displaced, past condition, which have been shown to promote better accuracy and complexity than simpler oral, monologic tasks, but which are likely to be more difficult for learners with low WMC given the constraints of WMC on the allocation of attention. Further, there are no studies that investigate the effects of WMC and guided pre-task planning time, which has been demonstrated to positively affect learner accuracy, and which is predicted to benefit learners with low WMC more than those with high WMC, given the additional processing support. Finally, more evidence is needed to confirm that planning time can affect learner performance on subsequent, unplanned tasks.¹⁰

¹⁰ There is a substantial body of literature on task repetition and its effects on learner performance, which has demonstrated that repeating the exact same task after an interval of weeks can have positive effects on learners’ accuracy, fluency, and complexity (Bygate, 2001; Bygate & Samuda, 2005). There is also research that demonstrates that repeating the exact same task at regular intervals has positive effects for language development, but has no carryover effect when the content of the task changes (e.g., Gass, Mackey, Alvares & Fernández-Garcia, 1999). However, to the author’s knowledge, other

By examining the interaction of WMC with two different types of pre-task planning time (guided and unguided) with a complex, there-and-then story-telling task, the study presented here offers more information about how conditions of task complexity interact with individual differences in learners' cognitive abilities. While Nielson (in press) did not find an interaction between WMC and the provision of planning time (instead finding that planning time benefited both high and low WMC individuals equally), it is possible that this was because the task in her study was not complex enough to differentiate between learners with high and low WMC. For example, learners were not given specific instructions for how to perform their narratives (e.g., keep them in the past tense), which would make the task more complex in terms of processing, and, therefore, highlight differences between high and low WMC participants. She also used a simple, monologic narrative told with visual support—task conditions that generally favor fluency over accuracy (Robinson, 1995a; Gilabert, 2007a). It may be the case that a more complicated task would have identified performance differences due to WMC because individual differences in WMC increasingly differentiate learner performance when tasks require learners to focus on multiple areas at once (Engle, 2010; Kane et al., 2007; Robinson, 2007; Unsworth & Engle, 2007).

For this reason, the present study continues the line of inquiry begun with Nielson (in press) and considers learner performance on a more complex, “There-and-Then” story-telling task. The study examines the differential effects of guided and unguided planning time when compared to no-planning conditions, taking into

than Nielson (in press), there has been no investigation into the effects of immediate task repetition with new content as a method of improving accuracy, fluency, and complexity.

consideration learners' WMC as well as the order in which the tasks (planned and guided, planned and unguided, unplanned) are completed. In addition to offering valuable information about how task conditions interact with learners' cognitive individual difference factors, the study also offers practical information about how tasks can be sequenced and offered in classroom settings. Because the research presented here was collected with intact classes under quasi-experimental conditions, the results can easily help inform practices in instructed SLA settings. As Ellis (2012) points out, most research on Robinson's Cognition Hypothesis (e.g., Gilabert, 2007a; Robinson, 2007a, 2007b) was carried out in laboratories rather than in classrooms, and there is a need for research on the influence of task conditions on learner discourse in classroom settings (p. 223). While the focus of this dissertation is a computer-based task carried out in a computer lab, the participants were in formal ESL classes, and the experimental protocol was carried out as part of the regular class activities, offering some degree of ecological validity and a framework for further classroom-based research.

Chapter 3: Research Hypotheses

The purpose of the study was to investigate guided pre-task planning, unguided pre-task planning, and no pre-task planning, as well as WMC, as predictors of learner accuracy, fluency, and complexity. The effectiveness of type of planning in terms of manipulating learner focus on content and form was investigated, along with the effect of WMC on learner performance in the guided planning condition, the unguided planning condition, and the no-planning condition. In addition, this study considered task order with respect to the provision of planning time and whether or not planning time—whether guided or unguided—had any facilitative effects on subsequent, unplanned tasks.

The study protocol was designed to address the following hypotheses:

- 1) The provision of planning time will benefit learner fluency under subsequent, unplanned tasks for both guided and unguided planning conditions.
- 2) Participants will produce more fluent and more complex language under the
 - (a) guided planning condition and
 - (b) unguided planning conditionthan under the no-planning condition.
- 3) Participants will produce more accurate language under the guided planning condition than under the unguided planning condition.
- 4) WMC will play a significant role in the accuracy of learners' performance in the
 - (a) guided planning,
 - (b) unguided planning, and
 - (c) no-planning conditions.
- 5) WMC will play a significantly larger role in the fluency of learners' performance under the
 - (a) no-planning condition and
 - (b) unguided planning conditionthan the guided planning condition.

- 6) WMC will play a significantly larger role in the complexity of learners' performance under the
- (a) no- planning condition and
 - (b) the unguided planning condition
- than the guided planning condition.

The first hypothesis predicted a facilitative effect for planning time on subsequent, unplanned tasks. Nielson (in press) found that learners who received pre-task planning time in the first task of a repeated-measures, story-telling experiment produced more fluent speech during the second, unplanned task than they had during the task for which they had received pre-task planning time. The characters, setting, and plot of the two stories were quite different, so she hypothesized that any facilitative effect of planning time on the second story would have had less to do with rehearsal of specific lexical items and narrative structure and more to do with practicing the task of telling a story. The idea that practice improves performance is supported by a vast array of research in cognitive and educational psychology (see DeKeyser, 2007, for a review). In this case, in addition to getting practice by simply telling one story first, the pre-task planning time before the first task might have provided pre-task practice in preparing for a monologic narrative, while also permitting learners to notice gaps between their planned discourse and their actual discourse. Then, on the subsequent, unplanned tasks, learners were able to produce more fluent speech because they were able to improve their performance based on what they had noticed. Hypothesis 1 predicted that these results would be replicated with a larger n and that the facilitative effects of planning time would be consistent across planning conditions (unguided and guided).

The next two hypotheses were concerned with the relative benefits of planning time on participants' accuracy, fluency, and complexity. With respect to hypothesis 2, which predicted that learner discourse would be more fluent and more complex when participants were given time to plan before speaking, previous research on both guided and unguided pre-task planning suggests that all learners should improve fluency and complexity when given time to plan, regardless of what instructions are given for planning time (Gilabert, 2007a; Mehnert, 1998; Ortega, 1999 & 2005). With respect to hypothesis 3, which predicted that specific, detailed guidance for how to use planning time would influence the accuracy of learner discourse, previous research on guided, pre-task planning indicated that both explicit instructions to focus on form as well as the pre-task provision of specific grammatical models can influence the accuracy of learners' performance (Foster & Skehan, 1997; Mochizuki & Ortega, 2008; Sangarun, 2005).

The next three hypotheses were intended to address the interaction of learners' WMC and pre-task planning time in terms of the accuracy, complexity, and fluency of L2 speech. Hypothesis 4 predicted an interaction between the accuracy of learner output and the three planning conditions (guided planning, unguided planning, and no planning). Learners were instructed to attend to content and form during all three story-telling conditions, and they were told to keep the story in the past tense; both of these conditions were designed to make this task more complex than a simple "Here-and-Now" story-telling task. Providing processing support by way of guided planning was designed to improve accuracy; in addition, requiring learners to tell the story in the past and without access to the pictures to guide them was predicted to

promote a shift toward accuracy and complexity (Robinson, 1995a; Gilabert, 2007a). Because WMC plays a greater role in the performance of more cognitively demanding tasks (Kane et al., 2007; Robinson, 2007; Unsworth & Engle, 2007), given the difficult nature of the task itself, WMC was hypothesized to play a role in the extent to which learners were able to attend to both form and meaning under all three planning conditions. Previous research on working memory and planning suggests that learners with high and low WMC can benefit from unguided pre-task planning time in terms of improvement to fluency and complexity (Nielson, in press). For this reason, hypotheses 5 and 6 predicted that WMC would play more of a role in the (most complex) no-planning condition than in the guided or unguided planning conditions. That is, these hypotheses predicted that learners with low WMC would benefit from the processing support provided by both guided and unguided planning in terms of fluency and complexity.

The research hypotheses presented above were designed to investigate guided and unguided planning time, task order, and WMC as predictors of accuracy, fluency and complexity. In addition to identifying the effect of these predictor variables on specific aspects of learners' output, an examination of the relationship between the accuracy, fluency, and complexity measures under each of the conditions (guided planning, unguided planning, no planning) was undertaken to indicate how changing task conditions can manipulate learners' focus on content and form. Clear trade-offs between the various accuracy, fluency, and complexity measures would indicate support for Skehan's Trade-Off Hypothesis (1998, 2009) whereas a more complex relationship between dependent variables would lend support to Robinson's

Cognition Hypothesis (2001, 2005c, 2011). As both of these hypotheses make claims about learners' cognitive processing, examining differences in learner production attributable to WMC was meant to address the overall research question of how WMC and task conditions can predict the quality of learner discourse.

Chapter 4: Methodology

4.1 Participants

Like Nielson (in press), this study was conducted with ESL students in the United States; participants were recruited from intermediate classes at the Annandale, Alexandria, Manassas, and Woodbridge campuses of Northern Virginia Community College, which has standardized curricula and sequences for ESL instruction. Intermediate students were targeted because the interaction of WMC with the SLA process is likely to be more pronounced for intermediate learners (Mackey et al., 2002; O'Brien et al., 2007; Payne & Whitney, 2002) and because the provision of pre-task planning time is likely to be more beneficial for intermediate learners (Ortega, 1999; Yuan & Ellis, 2003). One potential issue with using an ESL population is the wide range of L1 backgrounds, English language learning experiences, and amount of time spent in the U.S. This was partly controlled by using academic ESL classes within a single institution, as participants were at roughly the same intermediate level as determined by their scores on a standardized ESL placement test. The community college chosen for this study has a substantial sequence of ESL classes, and learners must advance through ten different five-credit courses to complete the ESL requirements. Given the number of different ESL courses and the fact that learners are placed into the courses via a standardized proficiency test, it is safe to assume that learners within each course have similar language proficiency. Because one of the aims of this study was to identify ways of manipulating pedagogical interventions that can be administered to groups of students

in formal ESL classes to promote a balanced focus on accuracy, fluency, and complexity, intact classes were recruited for participation.

One hundred students from seven different classes were recruited for participation in the research protocol, 92 students completed the working memory measures and the first story-telling task, and 72 students completed both the first and second story-telling tasks. Twelve of the students who completed Task 2 (the first story-telling task) and did not complete Task 3 (the second story-telling task) left the computer lab after the first task and before the study protocol was finished. Eight of the students had technical difficulties and did not successfully record their second stories. The twenty students who only completed the first story-telling task were distributed evenly over the seven classes tested for this research (two from session 1, four from session 4, two from session 7, and three from each of the other sessions).

Of the 72 students who successfully completed both story-telling tasks, there were 18 different L1s, with substantial groups of native Arabic speakers ($N = 11$), Korean speakers ($N = 11$), Vietnamese speakers ($N = 11$), Chinese speakers ($N = 7$), and Spanish speakers ($N = 6$). See Table 1 for a detailed list of native language and N -size for participants who completed only Story 1 as well as Story 1 and Story 2.

Table 1: Participants' Native Languages

Native Language	<i>N</i> (Story 1)	<i>N</i> (Story 1 and 2)
Arabic	16	11
Korean	15	11
Vietnamese	13	11
Chinese	8	7
Spanish	8	6
Unknown	6	4
Amharic	4	3
Bengali	4	3
Urdu	4	3
Farsi	2	2
French	2	1
Russian	2	2
Tajik Persian	2	2
Bilim	1	1
Nepalese	1	1
Tagalog	1	1
Thai	1	1
Turkish	1	1
Uyghur	1	1
Total	92	72

4.2 Instruments

Participants began the experimental protocol by completing a short language background questionnaire (see Appendix A) to capture factors other than WMC that might affect L2 performance, such as native language, heritage speaker status, and length of time studying English. Learners in the study had been in the United States between six months and seventeen years ($M = 2.63$, $SD = 3.17$) and studying English between six months and fifteen years ($M = 3.7$, $SD = 3.7$). There were no heritage learners.

In order to obtain a rough estimate of English proficiency and control for the possibility of language proficiency interacting with the experimental treatments, participants completed a ten-item English grammar test, adapted from the English proficiency test available online through Transparent Language (<http://www.transparent.com/learn-english/proficiency-test.html>). Because all participants had taken a standardized proficiency test to determine their placement in the ESL program, this instrument was simply intended to identify any outliers who were not yet at the intermediate level. The mean score on this ten-item test was 8.32 ($SD = 1.32$).

Participants' WMC was measured through two online spatial working memory tasks. While most studies of working memory in SLA use some version of the reading span task (Daneman & Carpenter, 1980) or the speaking span task (Daneman & Green, 1986), this might not be the best option. First, some researchers (e.g., Guar-Tavares, 2009) have administered the reading span task in participants' L2, which is a significant potential confound. Because the executive component of working memory is a domain-general construct, it is perhaps more logical, especially for ESL learners who do not have common first languages, to use a non-verbal measure of WMC, such as a spatial span task. Spatial span tasks correlate with other complex span tasks, such as reading span and operation span (Kane et al., 2004), and all of these complex span tasks measure an individual's ability to actively maintain information in the face of processing interference. Participants in this experiment completed two online tasks developed by a team of researchers at the University of Maryland led by Michael Dougherty: *Blockspan*, which requires learners to

remember and reproduce sequences of flashing lights in a grid, and *Shapebuilder*, which requires participants to remember and reproduce a sequence of multi-colored shapes in a grid (Atkins, Harbison, Bunting, Teubner-Rhodes, & Dougherty, 2009; Sprenger, Atkins, Colflesh, Briner, Buchanan, Chavis, Chen, Iannuzzi, Kashtelyan, Dowling, Bolger, Bunting, & Dougherty, in preparation). These tasks are automatically scored, and learners see whether or not they have reproduced the sequences properly. The tasks increase in difficulty as learners complete the targeted sequences.

After finishing the working memory tasks, all students were asked to complete two different story-telling tasks (a picture-guided narrative with planning time in one of the two planning conditions and a different one without planning time). Two different stories with two different sets of pictures of roughly equal complexity were used for these tasks. Story A is a narrative about two monkeys helping a stuck mouse out of a tree. Story B is a narrative about a man taking his pets on a car ride and crashing in the woods. The pictures for Story A are available at <http://terpconnect.umd.edu/~kbrown34/Slide%201/>. The pictures for Story B are available at <http://terpconnect.umd.edu/~kbrown34/Task%20A/taska.html>. Both sets of pictures are included as Appendix B. The narratives were piloted with eight native speakers to ensure that the stories generated from both sets of pictures were relatively similar in length and complexity. In addition, these stories were used with ESL learners in a similar study of WMC and unguided pre-task planning time (Nielson, in press), and none of the statistical analyses indicated an effect for the story with respect to the accuracy, fluency, or complexity of learner output ($p < .01$).

Learners completed the story-telling task with planning time using written prompts under two different planning conditions: form- and content-focused guided planning and unguided planning. The guided planning condition was developed based on Sangarun's (2005) findings in the study of pre-task planning time with specific guidance, which indicated that the most efficient way of providing pre-task guidance was to ask learners to focus on both form and meaning. The handouts that were distributed during pre-task planning time can be found in Appendix C (guided planning) and Appendix D (unguided planning).

Because intact classes were used for this study, and learners completed the study protocol during class, there was limited time for the researcher to offer specific instructions and examples of how to perform the experimental protocol. In order to make the experimental protocol run as smoothly as possible, participants were each given a sheet of paper with very specific instructions for how to complete the experimental tasks. Each sheet of paper offered learners the url for the experimental website, images of the instruments and the order in which they were to be completed, and clear instructions for completing each task. An example of this instructional sheet can be found in Appendix E.

4.3 Procedures

The researcher visited the ESL classes before the experiment to explain the research project, review the consent form, answer questions, and demonstrate how to complete a picture-guided narrative. Then, on the day of each experiment, the researcher met the classes and escorted participants to a computer lab. Before the

experiment, she demonstrated how to perform the picture-guided narrative task and the working memory tasks. She also distributed the English grammar quiz and gave each participant a sheet with instructions for how to complete the research protocol.

The entire research protocol was conducted in a computer lab during a regular class meeting, with individuals completing the various study tasks individually through a web-based interface. After completing the language background questionnaire, participants performed the working memory tests and the two picture-guided narrative tasks. Participant responses on all three tasks were recorded through the experiment website. During each computer lab session, the researcher was in the room to answer questions and monitor performance. In order to counterbalance the experimental conditions in terms of the provision of planning time, as well as the picture-guided narratives, students in the same classes were randomly assigned to one of the two experimental groups, with four conditions each, as shown in Table 2. Students seated near one another in the lab were completing different tasks at different times, so there was little risk that one student's performance would affect that of another.

Table 2. Task conditions and task order for both groups.

		Condition A	Condition B	Condition C	Condition D
Group 1	Task 1	Blockspan	Blockspan	Blockspan	Blockspan
		Shapebuilder	Shapebuilder	Shapebuilder	Shapebuilder
	Task 2	Story A No Planning	Story B No Planning	Story A Guided Planning	Story B Guided Planning
		Story B Guided Planning	Story A Guided Planning	Story B No Planning	Story A No Planning
	Task 3	Blockspan	Blockspan	Blockspan	Blockspan
		Shapebuilder	Shapebuilder	Shapebuilder	Shapebuilder
Group 2	Task 1	Story A No Planning	Story B No Planning	Story A Unguided Planning	Story B Unguided Planning
		Story B Unguided Planning	Story A Unguided Planning	Story B No Planning	Story A No Planning
	Task 2	Blockspan	Blockspan	Blockspan	Blockspan
		Shapebuilder	Shapebuilder	Shapebuilder	Shapebuilder
	Task 3	Story A No Planning	Story B No Planning	Story A Unguided Planning	Story B Unguided Planning
		Story B Unguided Planning	Story A Unguided Planning	Story B No Planning	Story A No Planning

In the no-planning conditions, participants clicked on a link to view a slideshow of the picture prompts. Then, after a brief review of the images, they were immediately instructed to close the window with the images and record themselves telling the story presented in the narrative without any visual support. Audio Dropboxes, which are flash-based mp3 recorders developed by the Center for Language Education and Research at Michigan State University, were created to record students' story telling. The researcher provided the students with the first line of the story (e.g., "Two monkeys were walking in the park..."), and participants were instructed to provide approximately two sentences of the narrative per frame of the story. Participants were given 40 seconds to view the images and no time to plan. They were not permitted to use the pictures to guide them as they told the story in English. The researcher informed participants that they were being recorded in order to assess their spoken English. Before beginning the speaking task, participants were

reminded that they were completing a “There-and-Then” task, so the story must be told in the past; in addition, they were asked to keep their stories as grammatically correct and detailed as possible.

In the guided and unguided planning conditions, participants received a sheet of paper with their planning instructions (Group 1 = guided planning; Group 2 = unguided planning); see Appendices C and D for the sheets that were distributed for each of the planning conditions. Then, as in the no-planning condition, participants clicked on a link to view the pictures and were told to take 10 minutes to plan their stories. The researcher informed students that they were being recorded to assess their spoken English. The participants had access to the planning instructions and writing implements during the planning phase, but they were instructed not to write out their narratives, and the researcher circulated around the room to make sure that learners were compliant with these instructions. The planning instructions and writing implements were collected before the story-telling phase of the experiment. Before beginning the speaking task, participants were told to close the window with the images and they were reminded that they were completing a “There-and-Then” task, so the story must be told in the past; in addition, they were asked to keep their stories as grammatically correct and detailed as possible.

4.4 Scoring and Operationalization of Variables

WMC was measured by scores on Blockspan and Shapebuilder. Learners’ scores on the two measures were combined into a composite WMC rating by averaging z-scores. Operationalizing the variables of accuracy, complexity, and

fluency was less straightforward, as the constructs have been measured in a number of different ways across SLA studies. As Pallotti (2009) points out in his discussion of the different ways researchers have operationalized accuracy, fluency, and complexity across empirical studies, it would be a great benefit to SLA research to standardize a set of measures (p. 599); to this end, this study employs various measures that have been used in other, similar planning studies (e.g., percentage of error-free clauses, Guiraud's Index of Lexical Richness, speech rate, all of which are described below). In order to calculate these measures, all participants' speech samples were transcribed and checked.

After the speech samples were transcribed and checked, a number of calculations were performed. The number of syllables in each sample was calculated using an online syllable counter (<http://www.syllablecount.com/>), which uses a combination of a U.S. English syllable dictionary as well as a formula-based syllable counter; the formula-based syllable counter is activated when a word cannot be matched in the syllable dictionary. Words that are counted using the formula-based counter are displayed separately, and the number of syllables in each of those words was checked manually. The number of types and tokens in each sample was determined using the Vocab Profile website (<http://www.lex tutor.ca/vp/eng/>). The number of seconds in each speech sample was taken from the mp3 recordings of learners' speech. Clauses, t-units, and errors were calculated by the researcher. Twenty percent of the transcribed samples were randomly selected; the number of clauses, t-units, and errors were coded by an outside rater. Inter-rater reliability was calculated for this subset of transcribed speech samples.

Global accuracy was measured as percentage of error-free clauses (Yuan & Ellis, 2003). All errors in word order, subject-verb agreement, verb tense, article use, preposition use, and pronoun use, as well as omitted words and extraneous words, were considered to be errors. When learners made the exact same error—with the same word in the same context—more than once (e.g., repeating the same phrase with the same omitted word), this was not counted as an additional error. Mispronunciation was not counted as an error. In addition to global error rate, accurate use of specific target structures was also considered. As Mochizuki and Ortega (2008) point out, it is important to consider whether or not learners are able to shift their output to focus on the specific forms being targeted by the guided planning treatment. An analysis of the data in Nielson (in press) revealed that the most frequently occurring errors across performances were subject-verb agreement, tense shifts, appropriate use of the plural form of nouns, and errors with lexis (either errors in word choice or omitted words, not including articles). Because tense shifts are used for stylistic reasons, and it is difficult to determine when learners are generating errors and when they are deliberately shifting tense for emphasis, this was not included in the specific accuracy measures. The number of errors with plural nouns, the number of lexical errors, and the number of errors with subject-verb agreement were calculated, so the overall percentage of each type of error could be calculated. Learners in the guided planning condition were specifically reminded to focus on these areas.

In addition to the accuracy measures just listed, learners' instances of self-repair were also considered as a measure of attempted accuracy. As Gilabert (2007a,

2007b) points out, global measures, such as the number of error-free clauses, and specific measures, such as the number of verbs with correct subject-verb agreement, measure whether or not learners have acquired linguistic accuracy whereas the number of self-repairs addresses their accuracy as it is developing. In the present, self-repairs were calculated as the number of self-repairs per 100 words.

Fluency was measured as both unpruned speech rate, i.e., the number of syllables spoken per minute and pruned speech rate, i.e., the number of syllables produced by each participant after repeated words, fillers, and re-starts were eliminated from the transcript, divided by the total number of seconds of speech produced, and multiplied by 60 (Gilabert, 2007a; Ortega, 1999).

Structural complexity was measured as clauses per t-unit (Kawauchi, 2005; Mehnert, 1998; Yuan & Ellis, 2003). Lexical complexity was measured with Guiraud's Index of Lexical Richness, or the number of types divided by the square root of the number of tokens (Gilabert, 2007a). While traditional type-token ratios have been shown to be very sensitive to differences in text length—i.e., the longer the text, the more likely that types will be repeated, making it difficult to compare texts of different lengths—Guiraud's Index attempts to control for this by taking the square root of the number of tokens rather than the raw number of tokens.

The methodology presented here replicates that of previous planning studies as well as previous studies of WMC. All of the research for this dissertation was conducted over the course of the Spring 2012 semester, with learners tested in February, March, and April. The data were coded and analyzed throughout the data

collection period so that any issues with data collection could be addressed during the semester.

Chapter 5: Results

5.1 Descriptive statistics and initial analyses of variables

Before undertaking any inferential analyses related to the research hypotheses, the descriptive statistics for all of learners' dependent measures of accuracy, attempted accuracy, fluency, and complexity were considered individually. First, the mean scores for each dependent variable for all learners in both groups (Group 1—Guided Planning and Group 2—Unguided Planning) were calculated. See Table 3 for all of the mean, repeated-measures accuracy scores under both guided and unguided conditions.

Table 3. Descriptive statistics for accuracy and attempted accuracy with and without guided and unguided planning time

<u>Variable</u>	<u>Type of Planning</u>	<u>N</u>	<u>Min.</u>	<u>Max.</u>	<u>M</u>	<u>SD</u>
% Verbs with Incorrect Agreement	Guided	34	.00	.75	.14	.15
	None	34	.00	.42	.15	.13
% Error-free Clauses	Guided	34	.13	.91	.55	.21
	None	34	.17	.83	.55	.17
Self-corrections	Guided	34	.00	8.26	3.64	2.36
	None	34	.00	9.38	3.47	2.55
% Verbs with Incorrect Agreement	Unguided	38	.00	.40	.11	.123
	None	38	.00	.67	.14	.13
% Error-free Clauses	Unguided	38	.00	1.00	.63	.21
	None	38	.18	.92	.58	.18
Self-corrections	Unguided	38	.00	9.04	2.53	2.34
	None	38	.00	11.29	3.52	2.89

Table 4 contains the complete descriptive statistics for pruned and unpruned speech rates with and without both types of planning time.

Table 4. Descriptive statistics for fluency with and without guided and unguided planning time

<u>Variable</u>	<u>Type of Planning</u>	<u>N</u>	<u>Min.</u>	<u>Max.</u>	<u>M</u>	<u>SD</u>
Pruned Speech Rate	Guided	34	28.89	172.00	98.29	28.63
	None	34	36.54	173.51	101.36	28.95
Unpruned Speech Rate	Guided	34	34.22	172.00	105.90	27.73
	None	34	46.54	183.24	109.70	29.69
Pruned Speech Rate	Unguided	38	67.06	181.13	117.62	30.39
	None	38	48.95	169.76	105.15	25.51
Unpruned Speech Rate	Unguided	38	67.06	182.26	123.92	29.15
	None	38	61.58	175.61	114.31	26.11

The complete descriptive statistics for both structural and lexical complexity are available in Table 5.

Table 5. Descriptive statistics for complexity with and without guided and unguided planning time

<u>Variable</u>	<u>Type of Planning</u>	<u>N</u>	<u>Min.</u>	<u>Max.</u>	<u>M</u>	<u>SD</u>
Clauses per T-unit	Guided	34	1.00	1.45	1.21	.14
	None	34	1.00	1.60	1.19	.17
Guiraud's Index	Guided	34	4.12	7.21	5.49	.76
	None	34	3.50	6.33	5.16	.66
Clauses per T-unit	Unguided	34	1.00	1.50	1.17	.13
	None	34	1.00	2.00	1.17	.18
Guiraud's Index	Unguided	34	4.36	6.89	5.38	.58
	None	34	3.67	7.76	5.31	.74

Along with the dependent measures of accuracy, fluency, and complexity, learners' WMC was calculated using two different spatial span tasks: Blockspan and Shapebuilder. There was a significant and sizable positive correlation between these two measures, $r = .58$, $p = .00$, $\alpha = .73$, which is in line with the correlation expected when two different WMC span tasks are administered to the same group of participants as each new test unavoidably has its own elements of variance (Conway,

Kane, Bunting, Hambrick, Wilhelm & Engle, 2005, pp 777 – 778).¹¹ Because of this significant, positive correlation, the Blockspan and Shapebuilder scores were averaged into a composite score of WMC ($M = 1440$, $SD = 571.95$).

After calculating the descriptive statistics for the dependent variables and the working memory covariate, the distributions of each dependent variable were visually inspected and evaluated with a Shapiro-Wilk test of normality. With respect to accuracy, the percentage of error-free clauses was normally distributed across the sample, but none of the specific accuracy measures (percentage of verbs with agreement errors, percentage of lexical errors, or percentage of errors with plural nouns) had a normal distribution. In terms of attempted accuracy, the number of self-corrections per 100 words had a normal distribution. For fluency, both unpruned speech rate in syllables/minute and the pruned speech rate in syllables/minute were normally distributed. Of the two complexity measures, only Guiraud's Index, which captures lexical complexity, was normally distributed. Finally, learners' composite working memory score also had a normal distribution.

The normally distributed variables (percentage of error-free clauses, number of self-corrections per 100 words, pruned and unpruned speech rate, and Guiraud's Index), were then considered in separate correlation analyses so that highly correlated variables could be combined and/or removed from the inferential statistical analyses. The measure of accuracy and attempted accuracy (percentage of error-free clauses and number of self-corrections per 100 words) did not have a significant correlation: r

¹¹ For example, in a recent study, researchers comparing the results of Blockspan and Shapebuilder as well as other measures of cognitive ability with a sample size of 117 report a similar correlation between these two spatial span tasks ($r = .62$, $p < .0001$) with a smaller though still significant correlation between Blockspan and Reading Span ($r = .33$, $p < .0001$) and Shapebuilder and Reading Span ($r = .47$, $p < .0001$) (Sprenger et al., in preparation).

= $-.066, p = .588$ with planning time, and $r = -.073, p = .525$ without planning time.

There was a significant, positive relationship between the pruned speech rate and the unpruned speech rate, $r = .972, p = .00$ for the speech rates with planning time and $r = .974, p = .00$ for the speech rates without planning time. While both speech rates are negatively affected by pauses, only the pruned speech takes into consideration dysfluencies, spoken fillers, and repeated words; because the rates were so highly correlated, the pruned speech rate was selected for inclusion in the inferential statistical analyses. As there was only one normally distributed measure of complexity (Guiraud's Index) a Pearson Product-Moment coefficient was not calculated for this variable; the measure of structural complexity was not considered in the inferential statistical analyses. There was no significant relationship between the composite working memory score and any of the dependent variables ($p < .01$). Given the results of the correlational analyses, a subset of the initial dependent variables was selected for inclusion in the parametric, inferential statistical procedures (see Table 6).

Table 6. Dependent variables used in inferential statistical analyses

<u>Construct</u>	<u>Measurement</u>
Accuracy	Percentage of error-free clauses
Attempted Accuracy	Number of self-corrections per 100 words
Fluency	Pruned speech rate in syllables/minute
Complexity	Guiraud's Index of Lexical Complexity
WMC	Composite working memory score

5.2 Planning Order/Conditions and WMC: Initial Analyses

The first inferential analyses were conducted to identify any facilitative effect of planning time on subsequent, unplanned tasks as well as to consider the potential effect of the type of planning time offered and the working memory covariate. Nielson's study (in press), which investigated the use of pre-task planning time to compensate for individual differences in WMC, revealed a significant relationship between the order of the provision of planning time and the fluency of learners' discourse. As in the present study, she used a counterbalanced, repeated-measures design with the order of the provision of planning time (before the first story-telling task or the second) as a between-groups factor. She found that learners who completed the first speaking task under the planning condition significantly improved their subsequent, unplanned performance, when compared to learners for whom the task conditions were reversed (p. 24). In order to test for the possibility of this carryover effect, the data in the present study were first considered in a repeated-measures analysis of covariance, with each of the normally distributed measures of accuracy, attempted accuracy, fluency, and complexity under both planning and no-planning conditions as the repeated-measures dependent variables, and the order of the provision of planning time (first or second) and the type of planning time (guided or unguided) as the between-groups factors. WMC was included as a covariate, operationalized as the composite score of the two working memory measures. In all analyses, an interaction term was initially included for the working memory covariate and the type of planning time in order to test for the assumption of the homogeneity of regression. When this interaction was non-significant, this interaction term was then removed from further analyses.

With respect to the accuracy of learners' discourse, the results of these ANCOVAs revealed that within subjects, there was no main effect of +/- planning on the percentage of error-free clauses $F(1, 67) = .847, p = .361, \eta_p^2 = .012$, nor was there an effect for the interaction of +/- planning with the working memory covariate $F(1, 67) = .380, p = .54, \eta_p^2 = .006$. There were no significant interactions between the between-groups factors (guided planning, unguided planning, planning first, and planning second) on the percentage of error-free clauses.

However, the number of self-corrections per 100 words, which was intended to capture learners' *attempts* at accuracy, yielded different results. The repeated-measures ANCOVA revealed a significant interaction between the type of planning time and WMC with self-corrections as the dependent variable, $F(1, 65) = 4.655, p = .035, \eta_p^2 = .067$, which indicates that the assumption of homogeneity of regression is not met. The significance of this test indicates an aptitude-by-treatment interaction between participants' working memory capacity and whether or not they received either type of planning time. The results of this analysis of covariance require that these data be considered with levels of participants' WMC as a between-groups variable, rather than a continuous covariate, which will be discussed in Section 5.4.

In terms of the fluency of learners' planned and unplanned discourse, the results of these ANCOVAs revealed that within subjects, there was a main effect of +/- planning on pruned speech rate, $F(1, 67) = 5.9, p = .018, \eta_p^2 = .081$. There was no interaction between planning and the working memory covariate, $F(1, 67) = 3.356, p = .056$, although (as with percentage of error-free clauses) this interaction approached statistical significance. There was, however, a significant interaction between +/-

planning and the type of planning time (guided or unguided) on pruned speech rate, $F(1, 67) = 8.921, p = .004, \eta_p^2 = .118$, and there was a significant interaction between +/- planning, the type of planning, and the order of the provision of planning, $F(1, 67) = 4.853, p = .031, \eta_p^2 = .068$. The results of this analysis indicate that in addition to +/- planning, both planning order and type of planning time affect learners' fluency. The implication of this finding will be considered in Section 5.2.3.

When Guiraud's Index of lexical complexity was included in the ANCOVA as the dependent variable, there was a main effect of +/- planning on lexical complexity, $F(1, 67) = 5.059, p = .028, \eta_p^2 = .069$, but no effect for the working memory covariate, $F(1, 67) = 1.541, p = .219$. Nor were there any significant interactions between +/- planning and the order of the provision of planning time, $F(1, 67) = 2.296, p = .134, \eta_p^2 = .033$, or +/- planning and the type of planning time, $F(1, 67) = 2.619, p = .110, .037$.

As the summary of findings in Table 7 shows, the results of the initial analyses of covariance indicate that there is a relationship between the attempted accuracy and lexical complexity of participants' speech samples, the planning conditions under which they generated these speech samples, and their WMC. In addition, there is an interaction between the order of the provision of planning time, the type of planning time (i.e., whether the planning time is guided or unguided) and the +/- planning variable that affects participants' fluency in terms of pruned speech rate. Finally, learners' lexical complexity is significantly different between the planned and unplanned conditions, without an effect for type of planning time or

order of planning time. A summary of the results of the repeated measures analyses of covariance is presented in Table 7, below.

Table 7. Results of initial, repeated-measures analyses of covariance

<u>Construct</u>	<u>Dependent Variable</u>	<u>Effect</u>	<u>Significant?</u>
<u>Accuracy</u>	Percentage of Error-Free Clauses	Interaction between planning time and covariate (WMC)	NO
		Main (+/- Planning)	NO
		Interactions	NO
<u>Attempted Accuracy</u>	Self-corrections/100 words	Interaction between planning time and covariate (WMC)	YES
		Main (+/- Planning)	N/A
		Interactions	N/A
<u>Fluency</u>	Pruned Speech Rate	Interaction between planning time and covariate (WMC)	NO
		Main (+/- Planning)	YES
		Interaction +/- Planning and planning order	YES
		Interaction between +/- Planning and type of planning time	YES
<u>Complexity</u>	Guiraud's Index	Interaction between planning time and covariate (WMC)	NO
		Main (+/- Planning)	YES
		Interactions	NO

Because there was a significant interaction between planning order (regardless of whether the planning was guided or unguided) and the planning condition (+/- planning) with fluency as a dependent variable, the data were split so that participants' performance on planned and unplanned tasks when planning time was provided first could be compared to the planned and unplanned performances of those participants for whom planning time came second.

5.3 The Order of the Provision of Planning Time

5.3.1 Descriptive statistics split by planning order

After the data set was split, the descriptive statistics were considered to ensure a normal distribution in the smaller sample size. See Table 8 for the descriptive statistics for the planned and unplanned conditions for guided and unguided planning separately, when planning time was offered for the first story-telling task.

Table 8. Descriptive statistics for participants who had either guided or unguided planning time during the first story-telling condition

<u>Variable</u>	<u>Type of Planning</u>	<u>Task</u>	<u>N</u>	<u>Min.</u>	<u>Max.</u>	<u>M</u>	<u>SD</u>
Percentage of Error-Free Clauses	Guided	2	17	.34	.91	.61	.19
	None	3	17	.25	.80	.57	.15
Self-corrections/100 words	Guided	2	17	.74	8.26	4.80	2.29
	None	3	17	.00	7.69	3.34	2.33
Pruned Speech Rate	Guided	2	17	28.89	117.05	86.39	20.54
	None	3	17	36.54	161.45	95.86	29.02
Guiraud's Index	Guided	2	17	4.30	7.21	5.35	.66
	None	3	17	4.16	6.16	5.17	.56
Percentage of Error-Free Clauses	Unguided	2	17	.33	1.00	.64	.18
	None	3	17	.18	.92	.57	.20
Self-corrections/100 words	Unguided	2	17	.00	6.82	2.13	1.63
	None	3	17	.76	7.45	2.55	1.81
Pruned Speech Rate	Unguided	2	17	78.00	181.13	128.46	30.68
	None	3	17	67.20	158.28	108.62	23.71
Guiraud's Index	Unguided	2	17	4.36	6.23	5.33	.53
	None	3	17	4.07	6.08	5.39	.52

Table 9 offers the descriptive statistics for all of the normally distributed dependent variables when learners were given planning time during the second story-telling task.

Table 9: Descriptive statistics for participants who had either guided or unguided planning time during the second story-telling condition

<u>Variable</u>	<u>Type of Planning</u>	<u>Task</u>	<u>N</u>	<u>Min.</u>	<u>Max.</u>	<u>M</u>	<u>SD</u>
Percentage of Error-Free Clauses	None	2	17	.17	.83	.54	.20
	Guided	3	17	.13	.88	.50	.23
Self-corrections/100 words	None	2	17	.00	9.38	3.60	2.79
	Guided	3	17	.00	5.71	2.47	1.84
Pruned Speech Rate	None	2	17	51.27	173.51	106.86	28.67
	Guided	3	17	56.56	172.00	110.05	31.16
Guiraud's Index	None	2	17	3.50	6.33	5.15	.77
	Guided	3	17	4.12	6.92	5.64	.85
Percentage of Error-Free Clauses	None	2	21	.33	.92	.59	.17
	Unguided	3	21	.00	.92	.63	.23
Self-corrections/100 words	None	2	21	.00	11.29	4.31	3.37
	Unguided	3	21	.00	9.04	2.86	2.79
Pruned Speech Rate	None	2	21	48.95	169.76	102.33	27.13
	Unguided	3	21	67.06	162.86	108.84	27.83
Guiraud's Index	None	2	21	3.67	7.76	5.25	.88
	Unguided	3	21	4.45	6.89	5.41	.625

The analyses of covariance presented in Section 5.2 indicate that the type of planning time that learners receive, as well as whether they receive it for their first story-telling task or their second, affect the accuracy, fluency, and complexity of their discourse. This pattern is clearly visible in the new descriptive statistics presented in Tables 8 and 9. Because the effect of repeating the story-telling task appears to interact with the effect of planning time (supporting Nielson's (in press) previous findings), the new groups of data were considered three different ways: with a multivariate analysis of covariance only for the first story-telling task (see Table 2 for the complete list of tasks, groups, and conditions) and then with two separate repeated-measures analyses of covariance (one for participants in Condition A and B,

who completed this task without planning time, and one for participants in Condition C and D, who completed it with planning time).

5.3.2 Between-groups data analysis without the influence of task repetition

Because there was an interaction between the order of the provision of planning time and the type of planning time offered, the data were split so that they could be compared in a MANCOVA. For this analysis, only the first of the two planning tasks (Task 2, per Table 2) was considered, and learners were divided into three groups according to whether or not they received planning time (and the type of planning time they received). All of the normally distributed dependent variables in Table 6 were considered in this analysis, and the composite working memory score was included as a covariate. Wilks' statistic showed a significant effect of type of planning time (guided, unguided, or none) on the dependent variables of accuracy, fluency, and complexity, $\lambda = .825$, $F(8, 86) = 2.019$, $p = .047$. Given the statistical significance of the MANCOVA, separate univariate ANCOVAs for each of the outcome variables were undertaken, and they revealed significant differences between groups (guided planning, no planning, and unguided planning) for some of the dependent variables.

With respect to accuracy as measured by the percentage of error-free clauses, a one-way analysis of covariance did not reveal any significant main effects for either the planning group (guided, unguided, or none), $F(2, 86) = .034$, $p = .990$, $\eta_p^2 = .024$ or for an interaction between +/- planning and the working memory covariate, $F(1,86) = .004$, $p = .103$, $\eta_p^2 = .001$. However, there was a significant main effect for planning group on the number of self-corrections per 100 words, $F(2, 86) = 3.901$, p

= .024, $\eta_p^2 = .128$. There was no significant effect for an interaction between +/- planning and the working memory covariate, $F(1, 86) = 1.609, p = .208, \eta_p^2 = .010$ on number of self-corrections. Post-hoc tests revealed that having guided planning resulted in more self-corrections per 100 words than unguided planning time, $t(86) = 2.235, p = .028$; having no planning time also resulted in more self-corrections per 100 words than having unguided planning time, $t(86) = 2.776, p = .007$. These results suggest differential effects for planning conditions in terms of fostering a focus on learner accuracy. See Figure 1, below, for the estimated marginal means of the number of corrections per 100 words under each of the planning conditions.

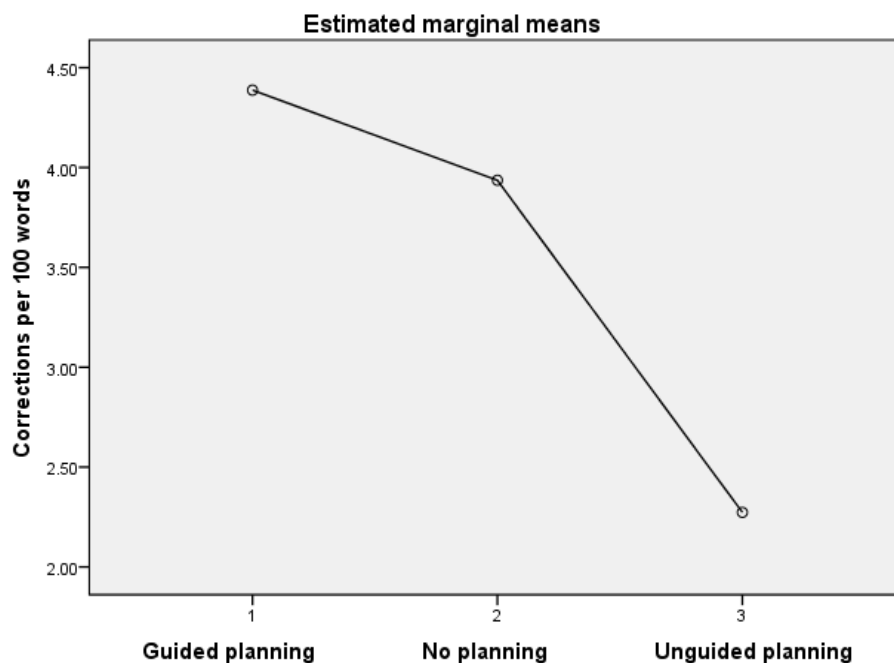


Figure 1: Corrections per 100 words under guided planning, no planning, and unguided planning conditions

With respect to the fluency of participants' discourse, the results of the ANCOVA indicated a significant main effect for planning conditions on participants' pruned speech rate, $F(2, 86) = 5.995, p = .004, \eta_p^2 = .124$, but the effect for an

interaction with the working memory covariate was not significant, $F(1, 86) = .139, p = .719, \eta_p^2 = .001$. The post-hoc tests indicated that participants who were given unguided planning time were more fluent than learners who were given both guided planning time, $t(86) = -3.238, p = .002$ and learners who were given no planning time, $t(86) = -.2938, p = .004$ (see Figure 2, below).

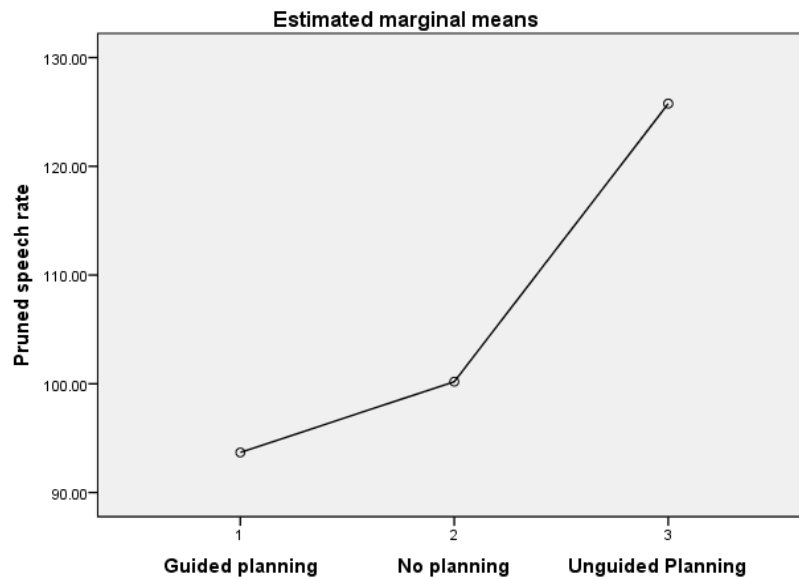


Figure 2: Pruned speech rate under guided, unguided and no-planning conditions.

Finally, with respect to lexical complexity, the results of the ANCOVA with Guiraud's Index as the dependent variable did not reveal any significant main effects for either planning group (guided or unguided), $F(2,86) = .523, p = .595, \eta_p^2 = .010$ or for an interaction between +/- planning and the working memory covariate, $F(1,86) = 1.023, p = .315, \eta_p^2 = .001$. The results of the between-groups analyses of covariance are presented in Table 10, below.

Table 10. Results of ANCOVAs for all participants on Task 3 (the first story-telling task) only

<u>Construct</u>	<u>Dependent Variable</u>	<u>Effect</u>	<u>Significant?</u>
<u>Accuracy</u>	Percentage of Error-Free Clauses	Main (type of planning)	NO
		Interaction between type of planning and WMC	NO
<u>Attempted Accuracy</u>	Self-corrections/100 words	Main (type of planning)	YES
		Interaction between type of planning and WMC	NO
<u>Fluency</u>	Pruned Speech Rate	Main (type of planning)	YES
		Interaction between type of planning and WMC	NO
<u>Complexity</u>	Guiraud's Index	Main (type of planning)	NO
		Interaction between type of planning and WMC	NO

5.3.3 Repeated measures analyses: No planning first

The between-groups MANCOVA and subsequent ANCOVAs presented in section 5.3.2 removed the order of the provision of planning time from consideration by eliminating the repeated measure, and they were used to simply analyze learners' performances under the randomly assigned conditions of guided planning, unguided planning, and no planning on the first story-telling task (Task 2). While the results from this analysis demonstrate that different task conditions (guided planning, unguided planning, no planning) resulted in different mean scores between subjects, they do not offer any information about what happens within subjects when the task conditions are changed. Because the order in which the participants completed the two story-telling tasks interacted with the planning condition to influence their performance, the data were split again so that a repeated-measures analysis of only learners who went from completing Task 2 (the first story-telling task) without planning time to completing Task 3 (the second story-telling task) with either guided

or unguided planning time could be considered (Group 1 and 2 in Conditions A and B, per Table 1). The following subsection of the data demonstrates what happens when learners complete the first story-telling task without any time to plan and then complete the subsequent story-telling task with either guided or unguided planning time.

For self-corrections per 100 words, the results of the analysis from this subgroup revealed an interaction between the planning condition (guided or unguided) and the working memory covariate, $F(1, 32) = 4.873, p = .035, \eta_p^2 = .142$. This confirms the earlier finding of an aptitude-treatment interaction between WMC and the type of planning time (see Section 5.3.2); the implications of this interaction will be considered during the analyses in Section 5.4. There was no main effect for the type of planning on the number of error-free clauses, $F(1, 33) = .244, p = .625, \eta_p^2 = .007$ or for the working memory covariate, $F(1, 32) = .461, p = .502, \eta_p^2 = .014$.

With respect to fluency (operationalized as pruned speech rate), the repeated measures ANCOVA with this subgroup revealed a significant main effect for planning type on fluency within subjects, $F(1, 32) = 5.329, p = .028, \eta_p^2 = .143$, as well as a significant interaction between the type of planning time and the working memory covariate, $F(1, 32) = 4.619, p = .039, \eta_p^2 = .126$.

With respect to lexical complexity, there was no main effect for planning type on Guiraud's Index, $F(1, 32) = .447, p = .509, \eta_p^2 = .014$; nor was there an interaction between planning type and the working memory covariate, $F(1, 32) = .000, p = .992, \eta_p^2 = .000$.

5.3.4 Repeated measures analyses for participants who began with planning time

The third subgroup considered in this series of analyses was comprised of participants who began the story-telling tasks with planning time and then completed the final story-telling task without any time to plan. The repeated-measures analysis of covariance revealed no main effect for +/- planning on attempted accuracy in terms of self-corrections per 100 words, $F(1,31) = .014, p = .908, \eta_p^2 = .000$ or for an interaction between +/- planning and the working memory covariate, $F(1,31) = .572, p = .455, \eta_p^2 = .018$. However, there was a significant interaction between +/- planning and planning condition (guided or unguided) on the number of self-corrections per 100 words, $F(1,31) = 7.817, p = .009, \eta_p^2 = .201$. See Figure 3, below, for a graph of the mean number of self-corrections per 100 words in both the planned and unplanned conditions divided into groups based upon who received guided and unguided planning time.

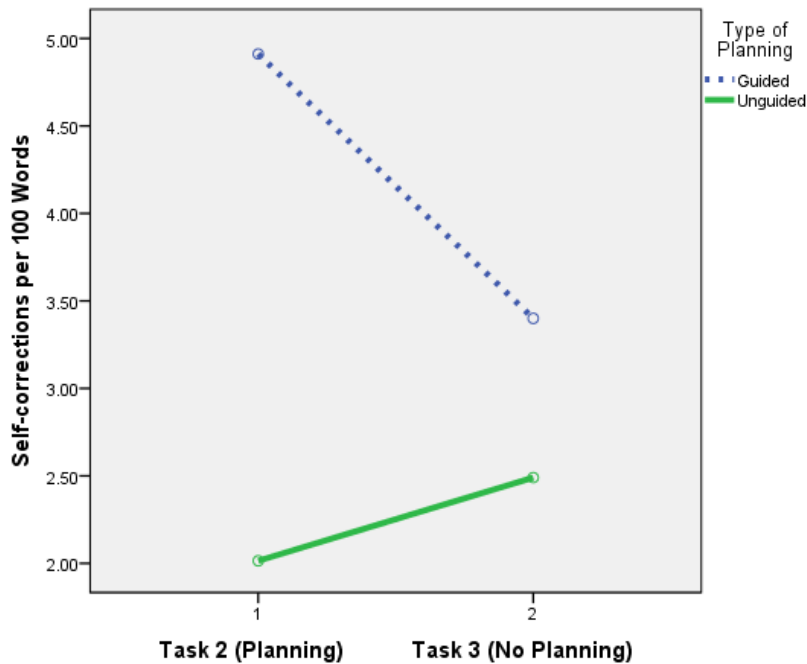


Figure 3: Differences in the number of self-corrections per 100 words within subjects between planning conditions (planned and unplanned) and planning groups (guided and unguided) when planning time was provided during the first story-telling task.

In terms of error-free clauses, the ANCOVA revealed no main effect for +/- planning, $F(1, 31) = .067, p = .797, \eta_p^2 = .002$, no effect for an interaction between +/- planning and the working memory covariate, $F(1, 31) = .139, p = .712, \eta_p^2 = .004$, and no interaction between +/--planning and planning condition (guided or unguided), $F(1, 31) = .114, p = .738, \eta_p^2 = .004$.

Within subjects, there was no significant main effect for +/- planning on fluency, $F(1, 31) = .767, p = .388, \eta_p^2 = .024$ nor was there an interaction between +/- planning and the working memory covariate, $F(1, 31) = .288, p = .668, \eta_p^2 = .006$. However, there was a significant interaction between +/- planning and planning condition (guided and unguided) in terms of participants' fluency, $F(1, 31) = 12.044, p = .002, \eta_p^2 = .280$ (see Figure 4, below).

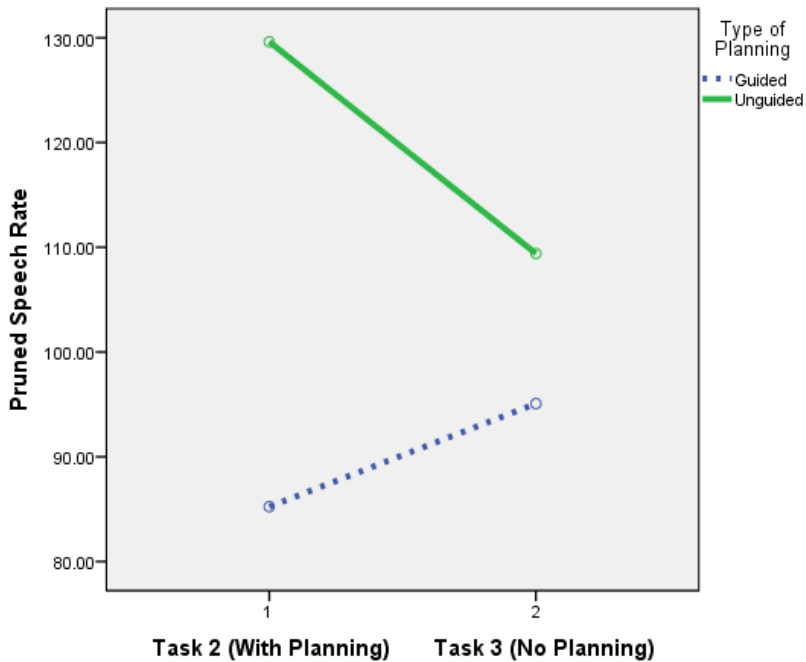


Figure 4: Differences in pruned speech rates within subjects between planning conditions (planned and unplanned) and planning groups (guided and unguided) when planning time was provided during the first story-telling task.

When participants' lexical complexity was considered in the repeated measures ANCOVA, there was no significant main effect for +/- planning on Guiraud's Index, $F(1, 31) = 1.016, p = .321, \eta_p^2 = .032$, nor was there a significant interaction between +/- planning and the working memory covariate, $F(1, 31) = 1.880, p = .180, \eta_p^2 = .057$, or a significant interaction between +/- planning and planning condition (guided or unguided) with respect to lexical complexity, $F(1, 31) = 2.149, p = .153, \eta_p^2 = .065$.

5.3.5 Summary of results from split-group repeated measures analyses

To summarize, in this set of analyses, the effects of type of planning, time of planning and WMC were considered in two subsets of the entire participant population. After a significant MANCOVA, a series of ANCOVAs was conducted with all participants on Task 2 only, with main effects for both +/- planning and the type of planning (guided and unguided) in terms of the number of self-corrections per 100 words as well as pruned speech rate. Then, to follow this analysis, two different sets of repeated-measures ANCOVAs were conducted with a subset of the participants so that the effect of task repetition could be considered separately. Half of the participants began with the planning condition for the first story-telling task and then completed the second story-telling task without time to plan first while the other half of the participants told the stories under reversed conditions. The two groups were considered separately to compare the effect of going from a planning to a no-planning condition to the effect of going from a no-planning to a planning condition. The results of these analyses are summarized in Table 11.

Table 11. Results of repeated measures ANCOVAs divided by the order of the provision of planning time

<u>Construct</u>	<u>Dependent Variable</u>	<u>Effect</u>	<u>Sig. with NO Planning 1st</u>	<u>Sig. with Planning 1st</u>
<u>Accuracy</u>	Self-corrections/100 words	Interaction between independent variable(s) and covariate (WMC)	YES	NO
		Main (+/- planning)	N/A	NO
		Interaction between planning time and planning condition	N/A	YES
<u>Attempted Accuracy</u>	Percentage of Error-Free Clauses	Interaction between independent variable(s) and covariate (WMC)	No	NO
		Main (+/- planning)	No	NO
		Interactions	No	NO
<u>Fluency</u>	Pruned Speech Rate	Interaction between independent variable(s) and covariate (WMC)	NO	NO
		Main (+/- planning)	YES	NO
		Interaction with WMC	YES	NO
		Interaction between planning time and planning condition	NO	YES
<u>Complexity</u>	Guiraud's Index	Interaction between independent variable(s) and covariate (WMC)	NO	NO
		Main (+/- planning)	NO	NO
		Interaction with WMC	NO	NO

5.4 Structural Complexity

Because the measure of structural complexity was not normally distributed, it was not considered in the inferential statistical analyses (the descriptive statistics are presented in Table 5). However, given the statistically significant effects of task repetition and type of planning time on the normally distributed dependent variables, the structural complexity data were also divided into two groups so that the descriptive statistics for the dependent measure of clauses per T-unit could be considered in light of task conditions. In terms of task order (planning first or planning second) as well as planning conditions (guided planning or unguided

planning), there was almost no difference within groups for structural complexity (see Table 12, below).

Table 12. Structural complexity (Clauses per T-unit) in terms of planning order and planning conditions

<u>Type of Planning</u>	<u>Task Order</u>	<u>N</u>	<u>Min.</u>	<u>Max.</u>	<u>M</u>	<u>SD</u>
Guided	2	17	1.00	1.45	1.23	.14
No Planning	3	17	1.00	1.50	1.22	.14
No Planning	2	17	1.00	1.60	1.16	.20
Guided	3	17	1.00	1.40	1.18	.13
Unguided	2	17	1.00	1.36	1.16	.11
No Planning	3	17	1.00	1.38	1.15	.11
No Planning	2	21	1.00	1.44	1.16	.14
Unguided	3	21	1.00	1.50	1.18	.15

5.5 Summary of relationship among variables and planning conditions

While Table 11 in Section 5.3.5 offers an overview of whether or not the various analyses discussed so far yielded statistically significant results, it does not capture the relationship between the dependent variables under the various conditions. Before moving on to the analyses of WMC, it would be useful, for each of the planning conditions, to indicate the relationship between the variables. In terms of attempted accuracy, as operationalized by learners' self-corrections per 100/words, learners produced more self-corrections when they told their stories under the unplanned condition, except for when they received guided planning time under the first planning condition. The opposite pattern is found with the percentage of error-free clauses: learners produced more error-free clauses under all of the planning conditions, except for when they had guided planning under the second story-telling condition, in which case their first, unplanned stories had more error-free clauses.

Learners' pruned speech rate was faster under all of the planning conditions, except for when they had guided planning time under the first story-telling condition, which was slower than the subsequent, unplanned condition. Guiraud's Index was always greater with guided planning time, regardless of whether it came first or second. For learners in Group 2, who had unguided planning time, Guiraud's Index was always better during the second task, regardless of whether it was planned or unplanned. See Table 13 for a summary of the comparisons of learner performance between planning conditions and planning order.

Table 13. Comparisons of participants' planned and unplanned performances according to planning conditions (guided/unguided) and planning order (first/second)

<u>Construct</u>	<u>Dependent Variable</u>	<u>Planning Condition</u>	<u>Task Order Condition</u>	<u>Better with Planning</u>	<u>Better without Planning</u>	
<u>Accuracy</u>	Self-corrections/100 words	Guided	1 st	x		
			2 nd		x	
	Error-free clauses	Unguided	1 st		x	
			2 nd	x		
	<u>Fluency</u>	Pruned Speech Rate	Guided	1 st		x
				2 nd	x	
Unguided		1 st	x			
		2 nd	x			
<u>Complexity</u>	Guiraud's Index	Guided	1 st	x		
			2 nd	x		
	Unguided	1 st		x		
		2 nd	x			

A visual depiction of the variables in relation to one another helps to illustrate the relative effects of planning conditions and the order of the provision of planning time. The z-scores for all of the dependent measures under planning and no-planning conditions were calculated and then plotted against one another in a between-groups comparison of each repeated measure, organized by the order of the provision of

planning time (see Figure 5, below). When unguided planning time was offered to learners for the first story-telling task, they produced more fluent speech with fewer self-corrections than the group that had guided planning time. The guided planners produced speech that was less fluent and had more self-corrections, and as indicated in Figures 1 and 2, these mean differences in learner output between planning conditions were statistically significant for both these variables. The mean lexical complexity between conditions was equal.

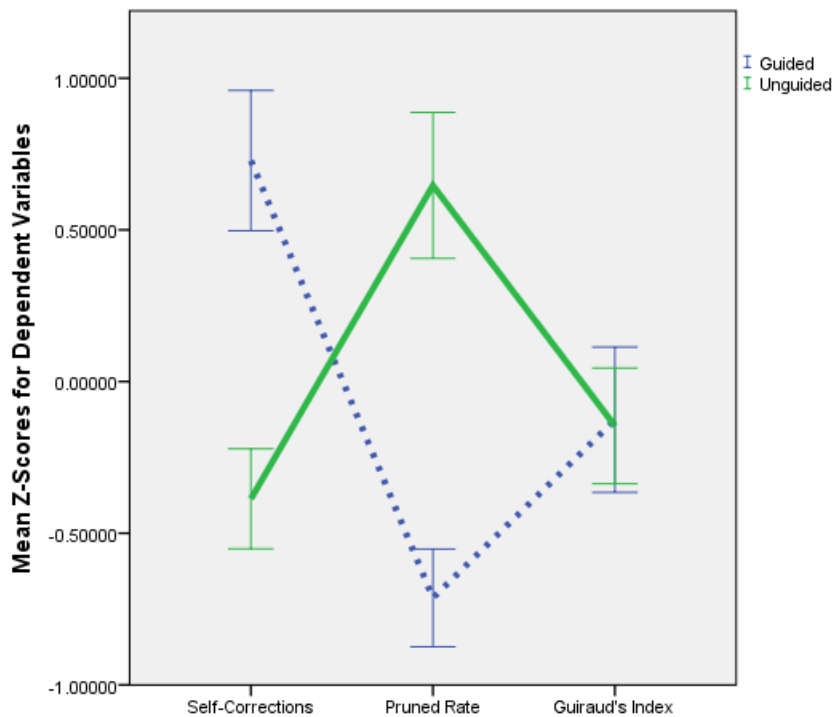


Figure 5. Mean z-scores for dependent accuracy, fluency, and complexity variables under guided and unguided planning conditions when planning time was offered first.

When learners completed the second story-telling task under the + planning condition, the pattern illustrating the relationship among the variables is different; this is likely because the effect of task repetition is also influencing learner performance. Learners with unguided planning time had relatively similar z-scores for the accuracy, fluency, and complexity variables; as presented in Table 11, the differences between

learners who had different types of planning time were not statistically significant (though there were significant interactions with WMC, which will be discussed in Section 5.6). Learners with unguided planning time had fewer self-corrections than the guided planners, a nearly identical speech rate, and a higher score on Guiraud's Index. It appears as though the guided planners prioritized lexical complexity over self-corrections when planning time was offered during the second story-telling task (see Figure 6, below).

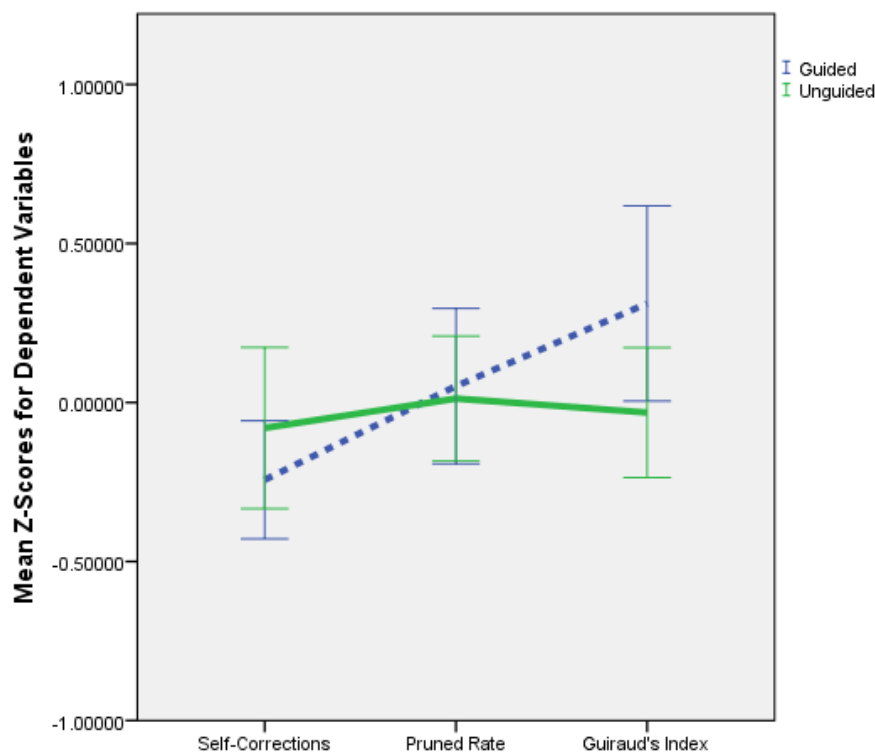


Figure 6. Mean z-scores for dependent accuracy, fluency, and complexity variables under guided and unguided planning conditions when planning time was offered for the second story-telling task.

Because the design of this experimental protocol was mixed, with learners in both Group 1 (guided planning) and Group 2 (unguided planning) varying only in the type of planning offered during the planning condition, learners in both Group 1 and

Group 2 completed the unplanned condition. However, these two groups will be considered separately because of the possibility that the type of planning offered during the planned condition affected learner performance on the subsequent, unplanned story-telling task. For the first task (the one which both groups performed without any planning time and which came before the planned stories), the relationship between the mean z-scores for the various dependent variables is very similar. Learners in Group 1 (guided planning) had slightly higher fluency at the expense of accuracy and lexical complexity while the reverse was true of learners in Group 2 (unguided planning); there were no significant differences between the means of the groups under the no-planning condition. Figure 7 offers a line graph with these relationships.

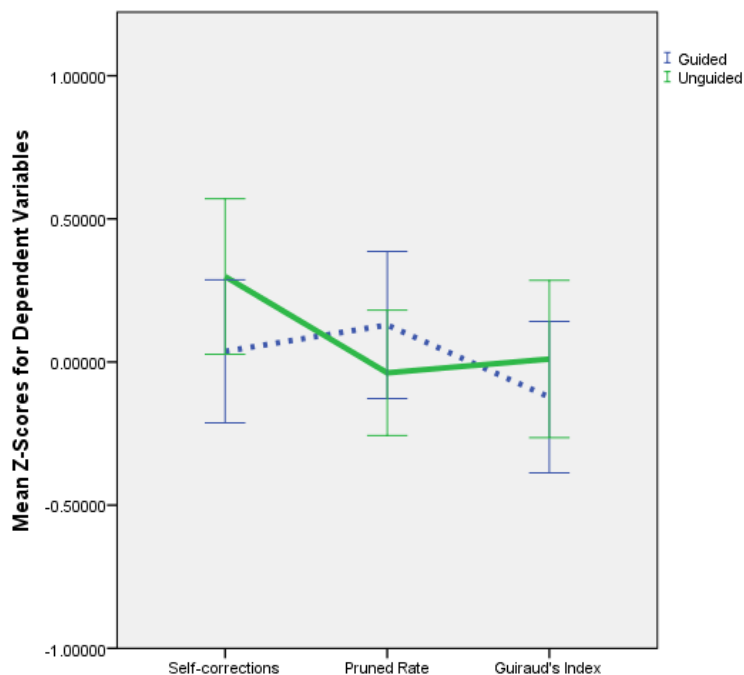


Figure 7. Mean z-scores for dependent accuracy, fluency, and complexity variables under no-planning conditions when planning time was offered for the second story-telling task. Learners in both groups completed this unplanned task as their first one in the sequence of telling two different stories.

There was more variation in the performance of the two experimental groups when they completed the unplanned task immediately after a task with planning time. Group 1 (the group that had guided planning for the first task) had more self-corrections and a lower pruned speech rate while Group 2 (which had unguided planning time prior to the unplanned story telling condition) had a higher pruned speech rate and greater lexical complexity (see Figure 8, below); the differences between the two groups in terms of the number of self-corrections were statistically significant.

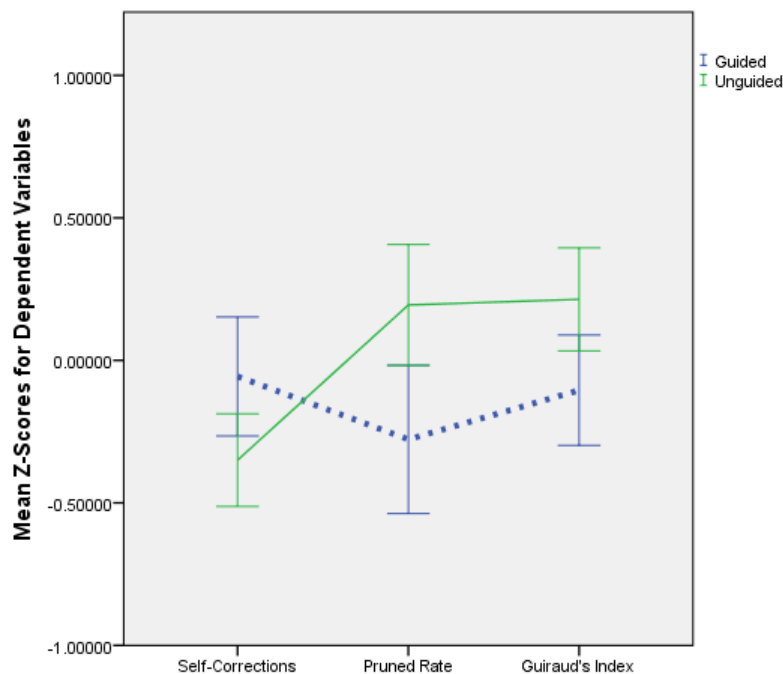


Figure 8. Mean z-scores for dependent accuracy, fluency, and complexity variables under no-planning conditions when planning time was offered for the first story-telling task. Learners in Group 1 completed the story-telling task prior to this unplanned task with guided planning time whereas learners in Group 2 completed the task prior to this one with unguided planning time.

5.6 Working Memory

Because learners' WMC was calculated as a measure of aptitude, it was included in the initial analyses as a continuous covariate. There was a significant

interaction between the covariate and the independent variable of +/- planning in the initial analyses of learner accuracy, suggesting an aptitude-treatment interaction (ATI) between the number of self-corrections per 100 words and WMC. In addition, the interaction between +/- planning and WMC with respect to fluency approached statistical significance ($p = .056$). In order to examine the ATI as well as determine how WMC might be used to divide learners into groups in an educational context, WMC was converted to a grouping variable so that learners' performance within groups categorized via WMC could be compared to one another.¹² The transformation of a continuous covariate to a blocking variable to examine differences in extreme scores is consistent with the literature on WMC and SLA (e.g., Gilabert & Muñoz, 2010; Mackey et al., 2002; Payne & Ross, 2005). The mean composite score on the working memory measures was 1418.47. In order to compare extreme groups while simultaneously maintaining an n-size large enough for between-groups comparisons, the sample size was divided into thirds. The learners with scores clustered around the mean were considered as a separate group and removed from the analysis, creating high- and low-WMC groups, each with a range of approximately 950 points between the cut scores for each group and a standard deviation of about 274 points (see Table 14).

¹² Another potential approach is to regress the dependent measures of fluency and accuracy on WMC for each of the treatment conditions (guided planning, unguided planning, no planning), with a Johnson-Neyman analysis to identify the composite WMC scores at which an ATI takes place. The results of this analysis for the significant interaction of WMC with type of planning time in terms of self-corrections per 100 words revealed that the conditional effect of the dependent variable of planning time was significant at composite working memory scores of 245 to 1510. Because the intention with this research was to identify groups of learners according to their WMC for instructional purposes, WMC was converted to a grouping variable even though using extreme groups can inflate effect size.

Table 14. Groups according to learners' WMC

<u>WMC Group</u>	<u>N</u>	<u>Minimum</u>	<u>Maximum</u>	<u>Mean</u>	<u>Std. Deviation</u>
High	20	1710	2560	2114.74	273.91
Mid	27	1210	1690	1481.92	142.38
Low	25	245	1175	777.62	273.54

The initial repeated-measures analyses of covariance were repeated with the new working memory variable as a between-groups factor. With respect to participants' attempted accuracy, there was a significant interaction between +/- planning and WMC on the number of self-corrections per 100 words, $F(2, 37) = 5.375, p = .026, \eta_p^2 = .127$; a significant interaction between WMC, type of planning time, and the order of the provision of planning, $F(2, 37) = 3.435, p = .018, \eta_p^2 = .270$; and a significant interaction between WMC and the type of planning, $F(2, 37) = 5.709, p = .026, \eta_p^2 = .134$. In the group that had guided planning first, the performance of individuals with high and low WMC was consistent within groups; with planning time, learners with high WMC generated more self-corrections per 100 words than learners with low-WMC; when the same individuals completed the task without planning time, they generated fewer self-corrections. There was a greater difference between performances for the high WMC than for the low WMC (see Figure 9, below).

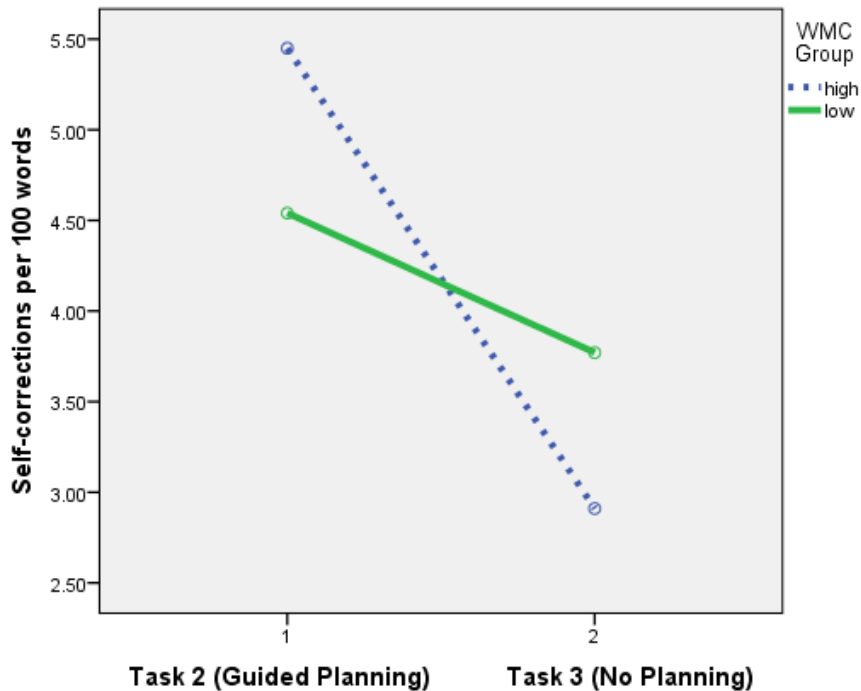


Figure 9. Mean self-corrections per 100 words according to WMC groups under guided planning and no-planning conditions when guided planning was offered first.

However, the pattern was different when learners had guided planning time for the second story-telling task; learners with high WMC followed the same pattern regardless of the order of the provision of planning time (more self-corrections with guided planning time), but learners with low and mid WMC scores produced more self-corrections during the first, unplanned task than they did under the second task with planning time (See Figure 10, below).

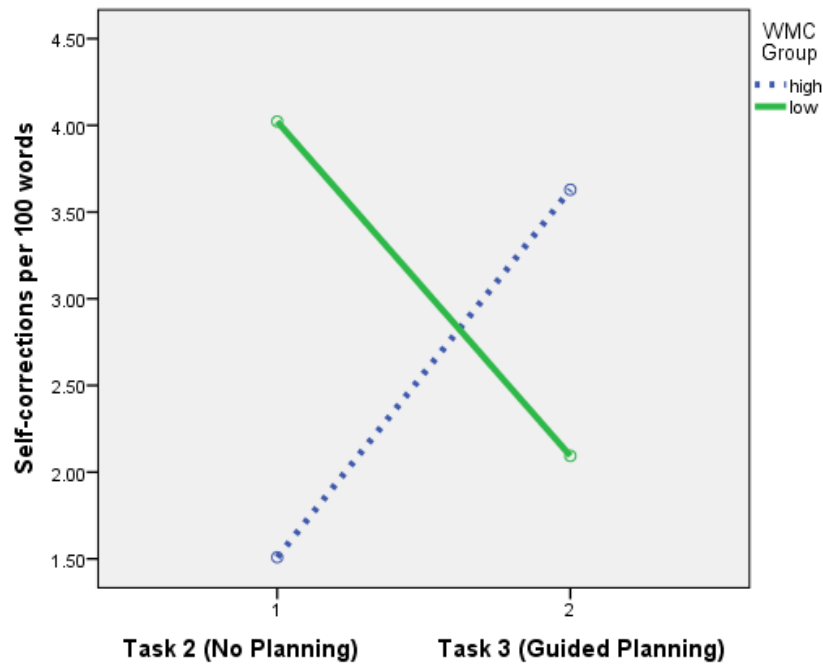


Figure 10. Mean self-corrections per 100 words according to WMC groups under guided planning and no-planning conditions when guided planning was offered during the second story-telling task.

With unguided planning time, the order of the provision of planning time did not appear to make a difference. Learners in all three WMC groups produced more self-corrections under the no-planning condition than they did with unguided planning time, and learners in the high WMC produced more self-corrections under both +/- planning conditions than learners in either the low or mid group, who were closer in terms of mean number of self-corrections per 100 words (see Figures 11 and 12, below).

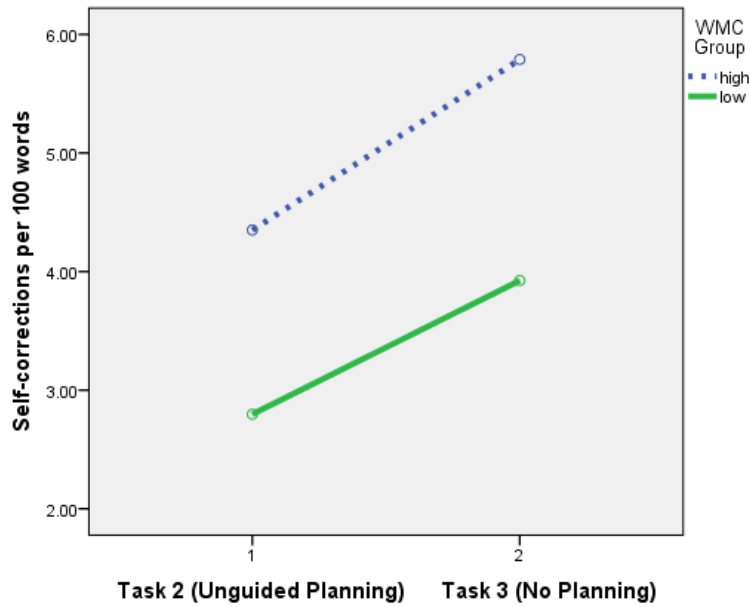


Figure 11. Mean self-corrections per 100 words according to WMC groups under unguided planning and no-planning conditions when unguided planning was offered during the first story-telling task.

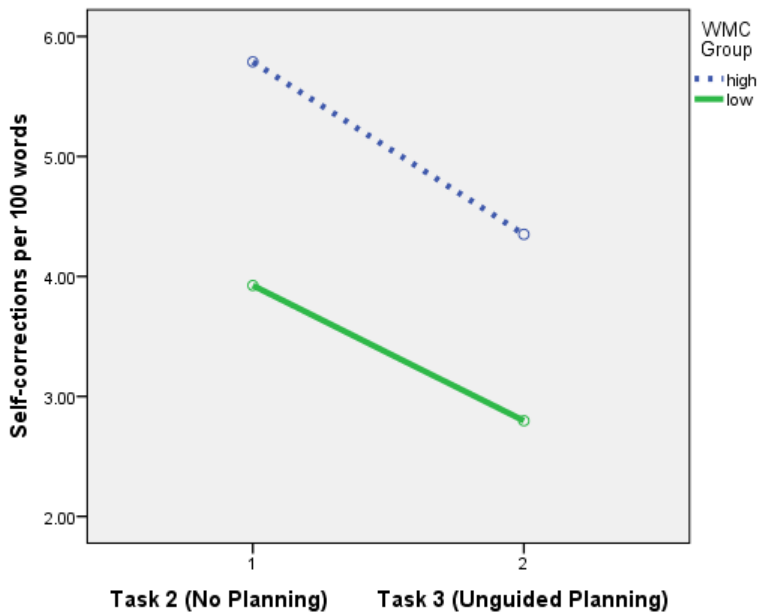


Figure 12: Mean self-corrections per 100 words according to WMC groups under unguided planning and no-planning conditions when unguided planning was offered during the second story-telling task.

There was a significant interaction between +/- planning and WMC group with respect to fluency, $F(1, 37) = 4.765, p = .035, \eta_p^2 = .114$. There was no significant interaction between +/- planning, WMC group, and type of planning time, $F(1, 37) = 3.681, p = .063, \eta_p^2 = .090$ or between +/- planning, WMC group, and the order of provision of planning, $F(1, 37) = .038, p = .847, \eta_p^2 = .001$. Without planning first, learners with high- and low-WMC had nearly identical mean pruned speech rates; however, when they completed their stories under either guided or unguided planning conditions, the learners with high WMC were less fluent than the learners with low WMC (see Figure 13, below).

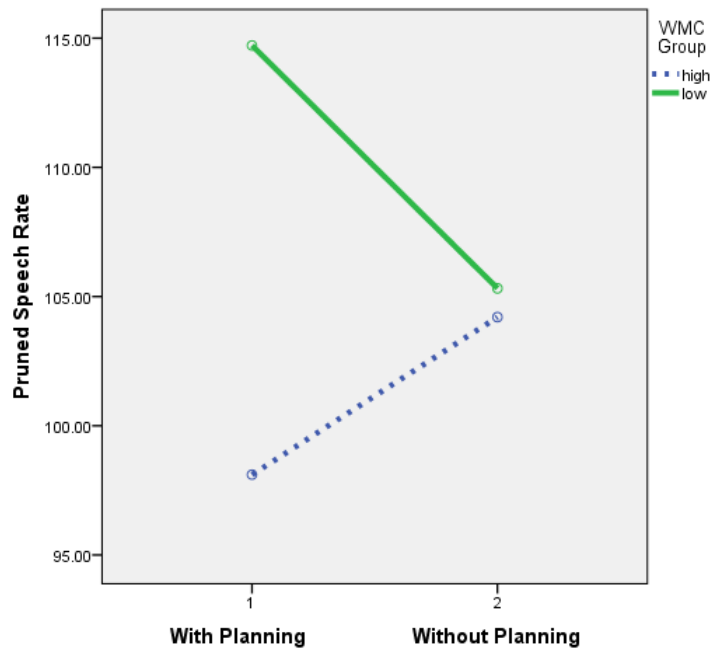


Figure 13: Mean pruned speech rate according to WMC groups under with planning and no-planning conditions.

In terms of lexical complexity as measured by Guiraud's Index, there was no significant interaction between +/- planning and WMC, $F(1, 37) = .009, p = .925, \eta_p^2 = .000$.

As with the analyses presented throughout Section 5.3, given the lack of a normal distribution of the structural complexity variable (clauses per T-unit), these data were not considered in the inferential, between-groups analyses with WMC. As there was no difference in learners' structural complexity when they offered planning time for the first task or the second task, working memory and structural complexity were considered simply in terms of type of planning (guided, unguided, or no planning). There is no clear pattern to learners' structural complexity and their WMC. Within groups, learners tended to perform nearly identically on both planned and unplanned tasks; however, in the guided planning group, learners with high WMC had more complex structures with guided planning than no planning, and in the unguided planning condition, learners with low WMC had more complex structures with unguided planning than under the no planning condition. Table 15 offers the complete descriptive statistics.

Table 15. Structural complexity (Clauses per T-unit) in terms of WMC and planning conditions

WMC Group	Type of Planning	N	Min.	Max.	<i>M</i>	<i>SD</i>
high	Guided	9	1.00	1.45	1.20	.19
	No Planning	9	1.00	1.50	1.13	.17
	Unguided	11	1.00	1.36	1.15	.13
	No Planning	11	1.00	1.36	1.16	.11
mid	Guided	11	1.00	1.40	1.21	.12
	No Planning	11	1.00	1.60	1.20	.17
	Unguided	16	1.00	1.33	1.14	.11
	No Planning	16	1.00	1.44	1.16	.14
low	Guided	14	1.00	1.38	1.20	.11
	No Planning	14	1.00	1.56	1.21	.18
	Unguided	11	1.00	1.50	1.24	.14
	No Planning	11	1.00	1.38	1.14	.13

5.7 Interrelationship among dependent variables and WMC

Robinson’s Cognition Hypothesis (2001, 2005c, 2011) states that when tasks are made complex along resource-directing conditions (e.g., by displacing time as with a “There-and-Then” story-telling task) learners can be directed to focus on both accuracy and complexity at the expense of fluency. Skehan’s Trade-Off Hypothesis (1998, 2009), on the other hand, states that accuracy, fluency, and complexity are in competition with one another, and that when accuracy increases, fluency and complexity will decrease. In addition to determining how task conditions and WMC influence learner accuracy, fluency, and complexity, the research undertaken here was intended to investigate the relationship between the dependent variables under various task conditions (guided planning, unguided planning, and no planning) for learners with different levels of WMC.

The presentation of the results so far has revealed that learners with high and low WMC display different patterns with respect to self-corrections and fluency, depending upon the type of planning time offered and the order of the two story-telling tasks. The next step is to consider those dependent variables in concert for each of the working memory groups.¹³

5.7.1 WMC, type of planning time, and planning order

When learners completed the first story-telling task with guided planning time, both WMC groups followed the same general pattern with respect to the relationship among accuracy, attempted accuracy, fluency, and complexity variables. However, while performance on the lexical complexity measure was nearly identical among the WMC groups, the results of the ANCOVAs presented in Section 5.6 indicate that learners with high WMC have more self-corrections and a lower pruned speech rate than learners in the low WMC group. The mean z-scores for the dependent variables for the two WMC groups are depicted in Figure 14.

¹³ Section 5.3.5 presented the standardized values of each of the dependent variables on a graph for each planning condition (guided, unguided, and no-planning) under both conditions of task order (planning first or planning second). The figures included in that section can serve as baseline reference for the graphs of working memory and task conditions presented here.

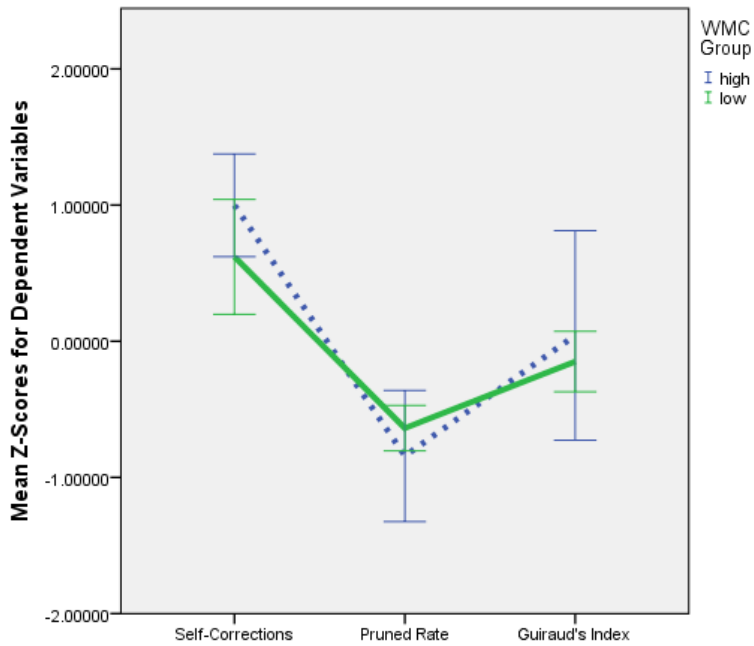


Figure 14, Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under guided conditions when planning time was offered first.

When guided planning time was offered second, there was far more variability in the relationship between the WMC groups. Learners in the low WMC group had fewer self-corrections, a higher pruned speech rate, and a lower score on Guiraud’s Index than learners in the high WMC group; the differences between WMC groups for attempted accuracy and pruned speech rate reached statistical significance (see Figure 15).

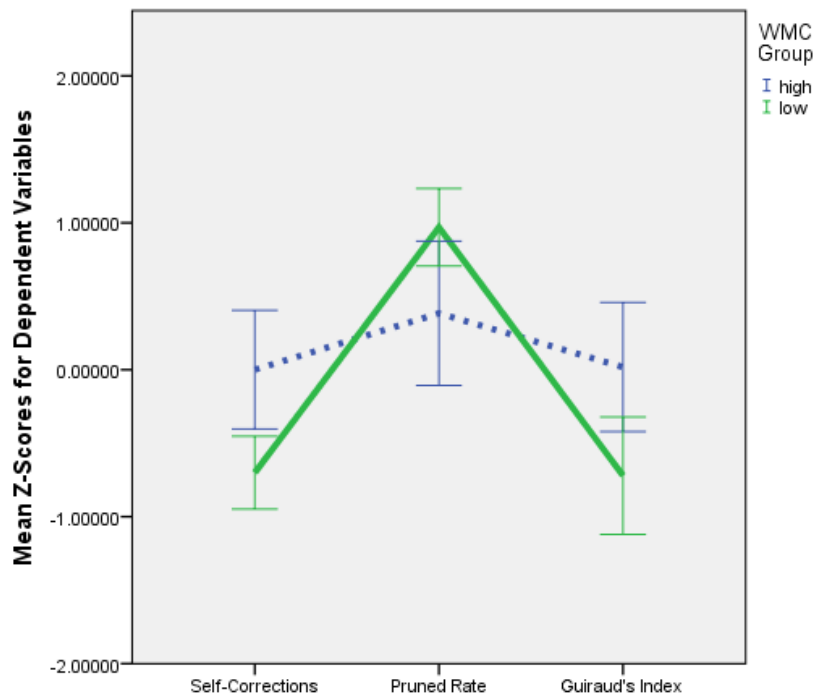


Figure 15. Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under guided conditions when planning time was offered second.

Learners in the high WMC group who had unguided planning time had more self-corrections a lower pruned speech rate than learners in the low WMC group, and levels of lexical complexity were very similar between groups. That is, for learners with low WMC, the unguided planning time benefited fluency whereas for learners in the high WMC group, the unguided time to plan benefited self-corrections (see Figure 16, below). These differences were statistically significant with respect to self-corrections and pruned speech rate.

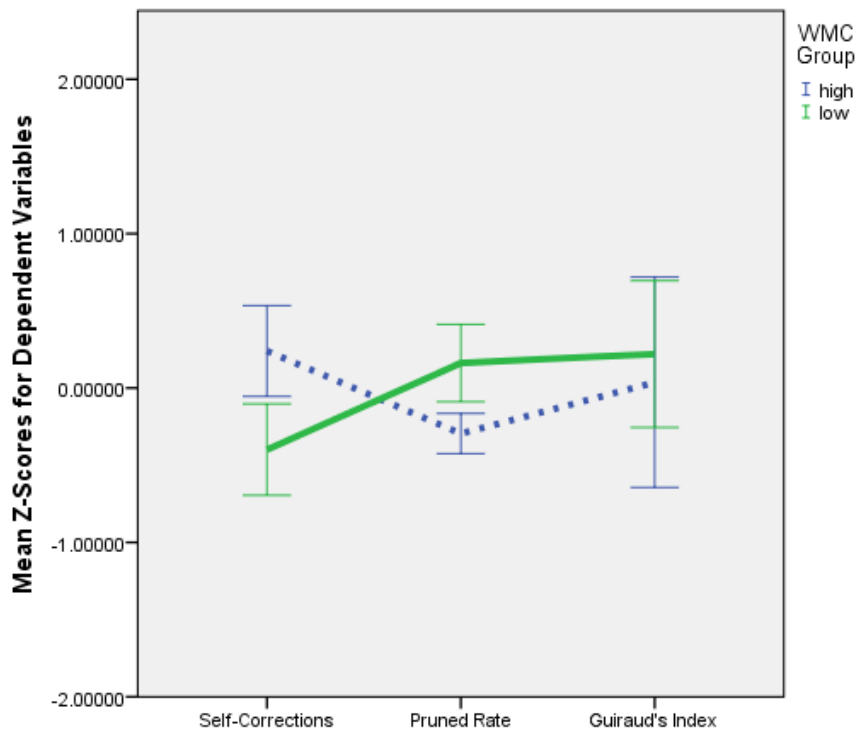


Figure 16. Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under unguided conditions when planning time was offered first.

Finally, when unguided planning time was offered after the no-planning tasks, there was even more variation between the two WMC groups. Learners with high WMC had high rates of self-correction, but were outperformed by the low WMC group with respect to every other variable (see Figure 17).

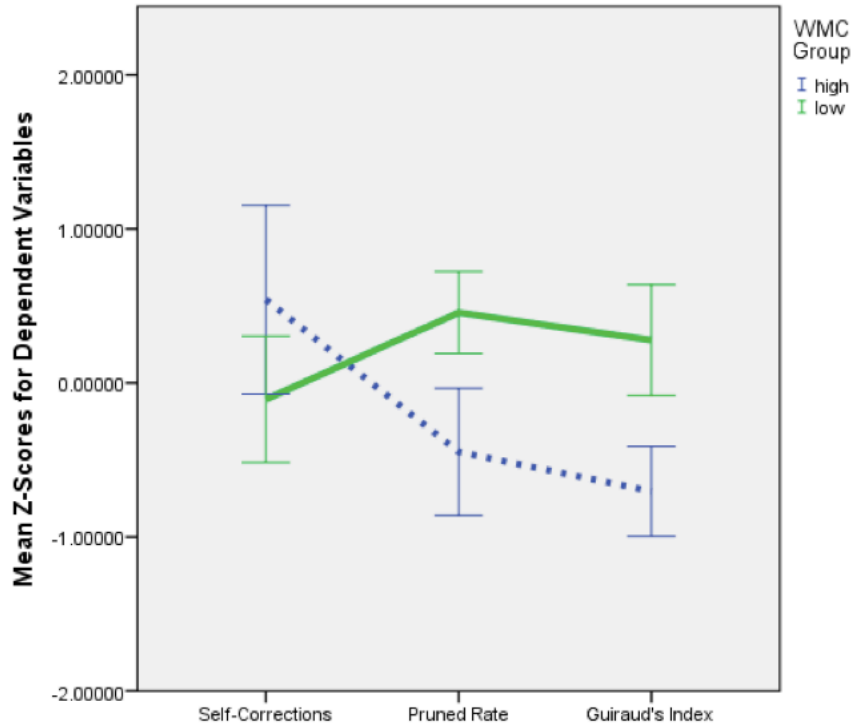


Figure 17. Mean z-scores for dependent accuracy, fluency, and complexity variables for high and low WMC groups under unguided conditions when planning time was offered second.

5.7.2 WMC and No-Planning Conditions

Figures 14, 15, 16, and 17 indicate the relationship among the standardized dependent variables according to planning order and type of planning, but as this was a repeated-measures study, it is important to also consider the relationship among variables under the no-planning conditions. For the first, unplanned task, both Group 1 and Group 2 were considered together (as there was no possibility of a previous planning task influencing their performance). There was almost no difference between the high- and low-WMC groups when they completed the first task under the no-planning condition (see Figure 18).

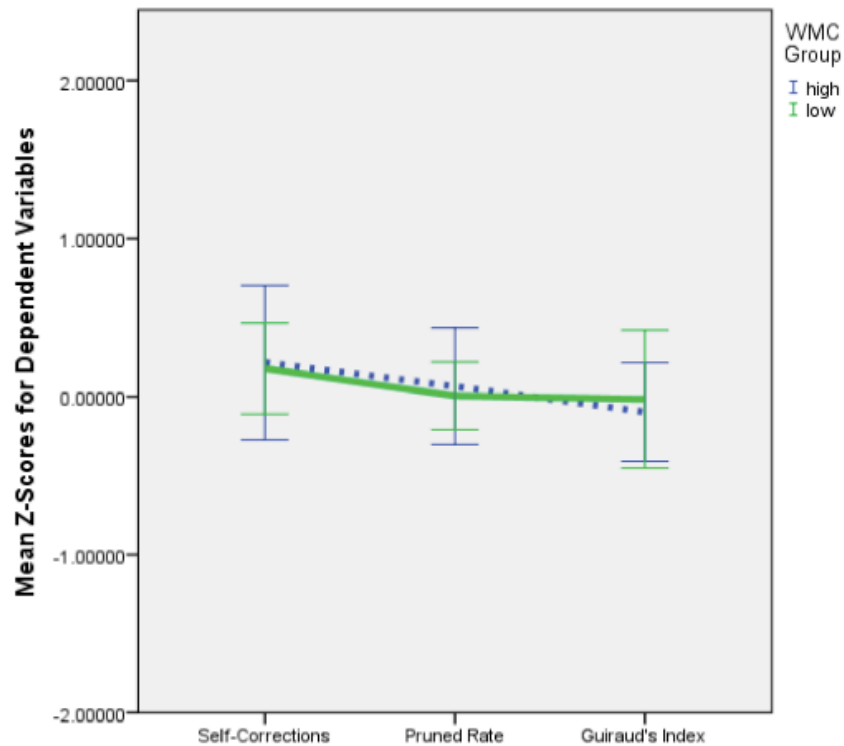


Figure 18. Mean z-scores for all dependent measures according to WMC group when learners had no time to plan for the first story-telling task.

However, when learners had either unguided or guided time to plan before completing the unplanned story-telling condition, there were differences between the WMC groups with respect to the dependent variables. Learners in the low WMC slightly outperformed learners in the high WMC group on all dependent measures when they had guided planning time for the first task. Learners in the low WMC performed very close to the mean on all measures, and there was more variation within the high WMC group, with learners appearing to prioritize self-corrections and complexity over fluency. Figure 19 offers a line graph of these relationships.

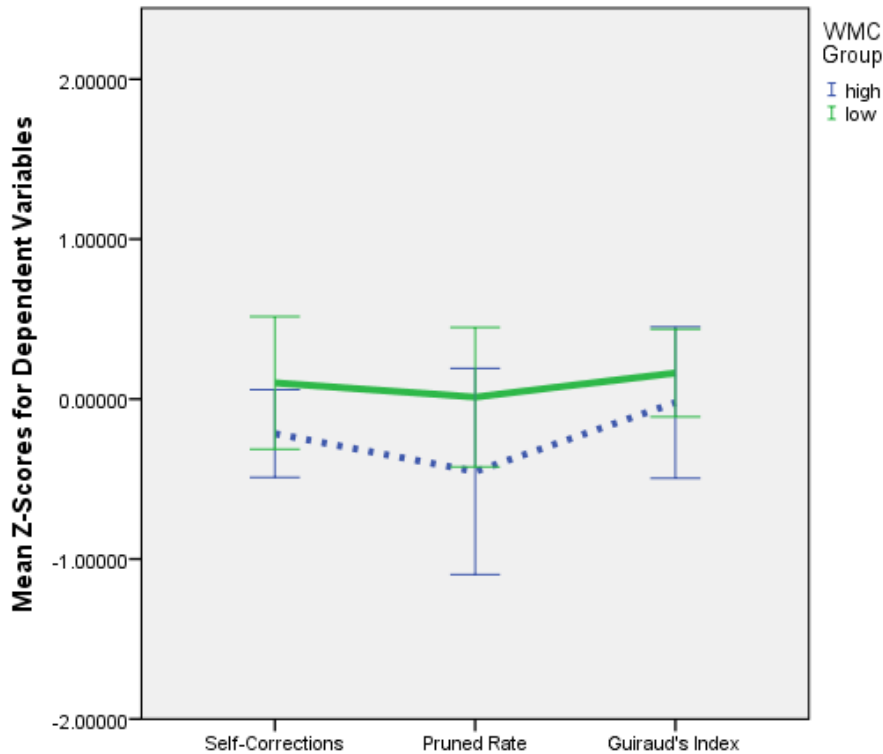


Figure 19. Mean z-scores for all dependent measures under no-planning conditions according to WMC group when learners had guided planning for the first story-telling task.

Finally, when participants began the story-telling portion of the experiment with unguided planning time, their subsequent unplanned performances had the opposite relationship between working memory groups. Learners with high WMC outperformed learners with low WMC on all dependent measures, and for the high WMC all outcome variables were measured at relatively similar levels. The pruned speech rate of low WMC learners was higher than their number of self-corrections per 100 words (see Figure 20, below).

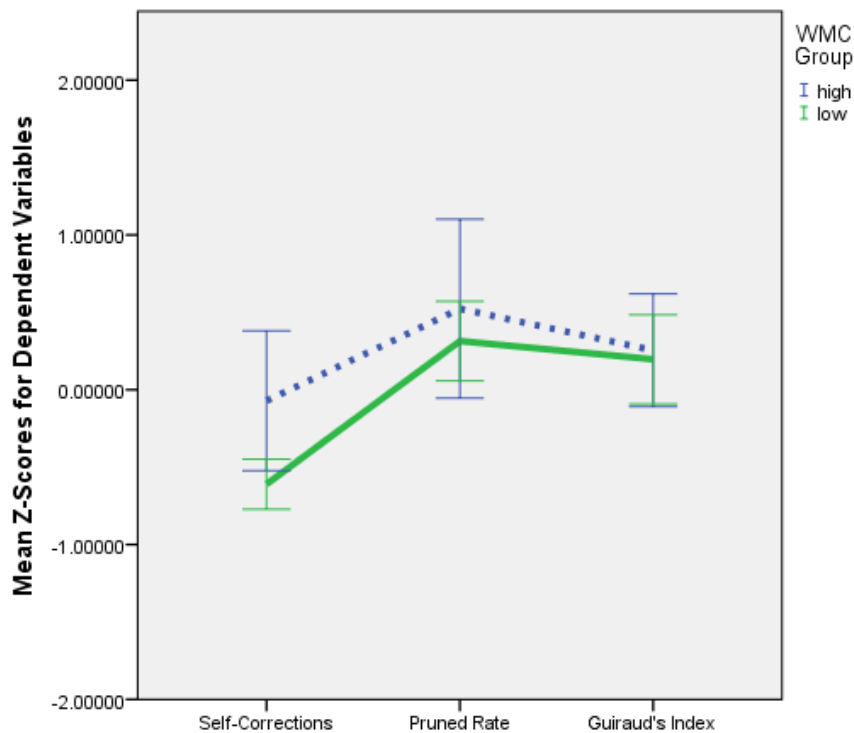


Figure 20. Mean z-scores for all dependent measures under no-planning conditions according to WMC group when learners had unguided planning for the first story-telling task.

5.8 Specific Accuracy Measures

The global accuracy measure of percentage of error-free clauses was used in the initial analyses, but it is necessary to consider the types of errors that learners made under the various task conditions. Participants were asked to specifically focus on subject-verb agreement and correct use of plural nouns, and to keep their stories in the past tense. Because participants used tense to add emphasis, and there were many tense shifts within the stories, it was impossible to tell when shifts in tense were due to error or due to learner-internal decisions about the narrative. For this reason, the number of tense shifts was not calculated as a specific accuracy measure. However, the number of lexical errors, the number of errors with plural nouns, and the number

of errors with subject-verb agreement were calculated. The repeated measures analysis of covariance was conducted with the specific accuracy measures.

The ANCOVA did not indicate a significant effect of planning conditions on any of the accuracy measures (all $p > .05$). Across all planning conditions, the largest portion of participant errors came from lexis, followed by subject-verb agreement, followed by plural nouns. While none is statistically significant, there were slight differences in the proportions of types of errors across planning conditions. For example, when learners were offered guided planning time, incorrectly used plural nouns accounted for the smallest percentage of their total errors. When learners had unguided planning time they made fewer errors with subject-verb agreement than in either the unguided condition or the no-planning condition, but they made more errors with plural nouns than in either of the other conditions. The unplanned condition resulted in the smallest percentage of lexical errors. See Figures 21 – 23, below, for pie charts representing the proportion of total errors from each of the categories.

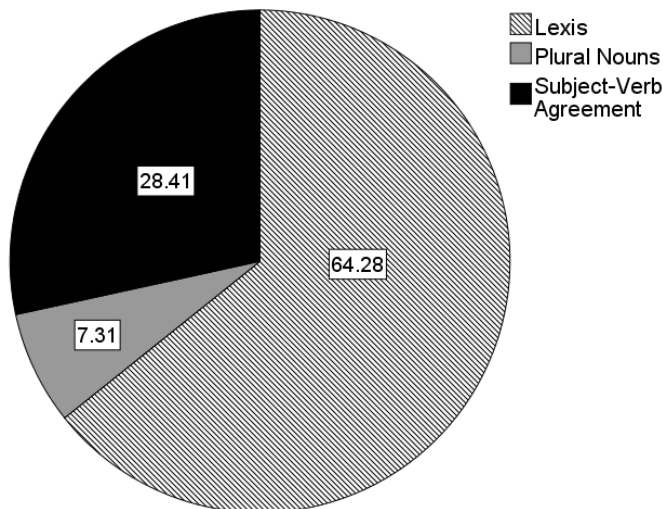


Figure 21. Percentage of errors due to lexis, plural nouns, and subject-verb agreement when participants were offered guided planning time

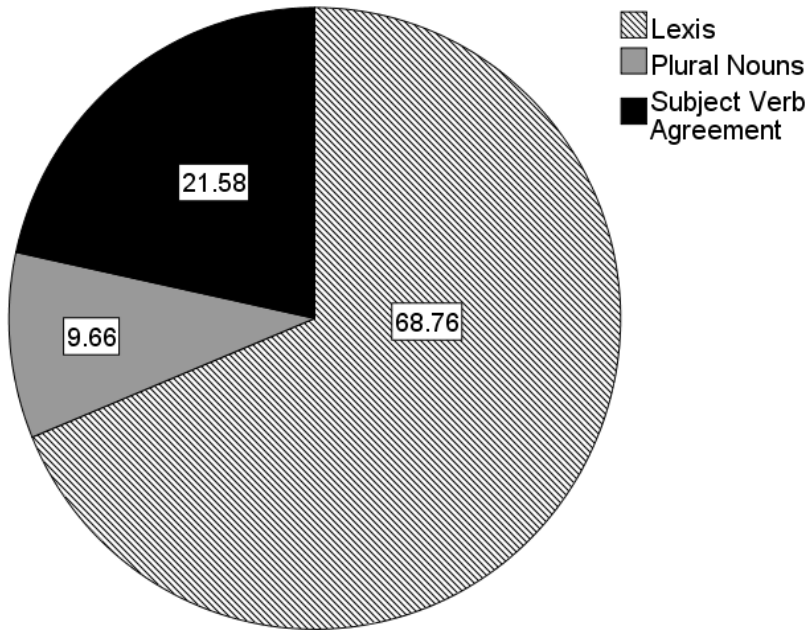


Figure 22. Percentage of errors with lexis, plural nouns, and subject-verb agreement when learners completed stories with unguided planning time.

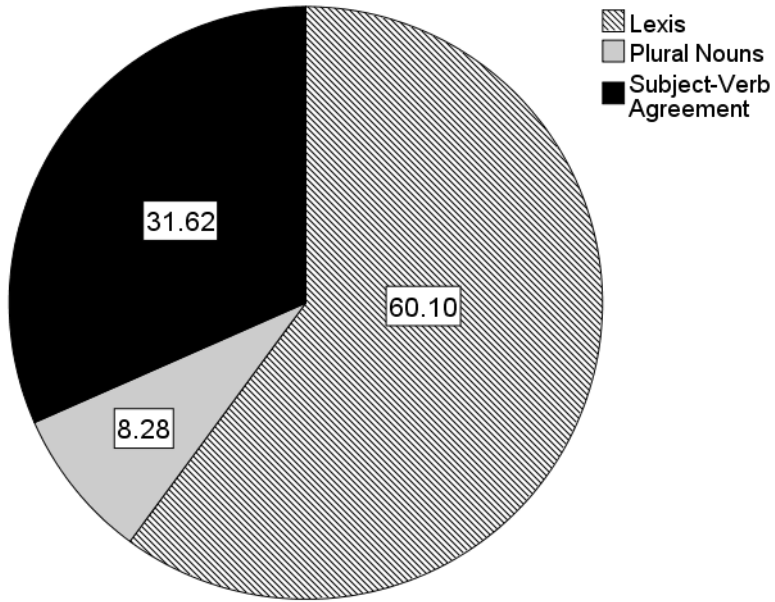


Figure 23. Percentage of errors with lexis, plural nouns, and subject-verb agreement when learners had no time to plan.

Chapter 6: Discussion and Conclusions

6.1 The Influence of Task Order on Learner Fluency

The first research hypothesis predicted that the provision of planning time for the initial story-telling task would benefit learner fluency on the subsequent, unplanned task under both planning conditions (guided and unguided). Based on the findings of Nielson (in press), who found that pre-task planning time improved the fluency of learner discourse on subsequent, unplanned tasks, this hypothesis predicted that when learners had either guided or unguided pre-task planning time as they told their first stories, their output from the second, unplanned story-telling task would be more fluent than the first, planned task. This hypothesis was only partially confirmed. There was a significant interaction between the order of planning time (whether it was offered first or second) and the type of planning time (either guided or unguided) with respect to learners' pruned speech rates, such that the predicted effect was obtained only for guided planning. When learners had guided planning time (Group 1) before their first story-telling task, their mean pruned speech rate was approximately 86 syllables per minute, and then their second, unplanned task was about 96 syllables per minute. However, when learners began with *unguided* planning time (Group 2), their mean pruned speech rate was about 128 syllables per minute, and then their speech rate declined during the subsequent, unplanned task to 108 syllables per minute. This drop in fluency indicates that giving participants unguided planning time for the first task did not result in a more fluent unplanned performance

during the second story-telling task. It is worth pointing out that when the situations were reversed, the speech rates were much closer together. Participants in Group 1 had a mean pruned speech rate of about 107 syllables per minute when they completed the first story under the no-planning condition, and then they improved their speech rate to approximately 110 syllables per minute when they completed the second task with guided planning time. Group 2 began with a mean pruned speech rate of about 102 syllables per minute during the first, unplanned task, and then completed the second story, with unguided planning time, at about 109 syllables per minute.

While the results do not fully support the notion of a carryover effect as described by Nielson (in press), they are interesting, nonetheless. Setting aside the issue of task repetition or any carryover effect, it is first important to note that the two different types of planning time (guided and unguided) appear to have very different effects on learner fluency. When examining just Task 2 (the first story-telling task), learners with no planning time had a mean pruned speech rate of about 100 syllables per minute, learners with unguided planning time had a mean pruned speech rate of about 126 syllables per minute, and learners with guided planning time had a mean pruned speech rate of about 93 syllables per minute (see Figure 2). The large difference in speech rate among the stories told under various planning conditions (unguided, guided, and no planning) was statistically significant, suggesting that different types of planning time influence learner fluency in different ways (see Figures 2 and 4 for plots of the different mean frequency scores under various conditions).

This is probably because when learners were given no planning time or were given unguided planning time (without instructions for how to use it), they prioritized fluency over other aspects of their speech, which is what has been suggested in previous research on pre-task planning (Ellis, 2009, Ortega, 1999, 2005; Skehan, 1998, 2009). However, when learners were given specific instructions for how to use their planning time, they attempted to comply with the very detailed guided planning instructions and focus on content and form, which had a detrimental effect on their fluency.

Given the effect that guided planning time had on participants' fluency, there was no way that the first research hypothesis could be supported. Further, the first hypothesis was based on the findings from Nielson (in press), which came from a study that allowed learners to use images to support their story-telling, making the task less cognitively demanding.¹⁴ In that study, Nielson hypothesized that the pre-task planning time (which was unguided) offered learners both a chance to practice telling a story and to notice areas where their planned performances differed from their actual performances so that they could correct them during the subsequent, unplanned stories.

Nevertheless, once we consider guided planning time as a treatment that slows fluency, we do see a carryover effect (in the opposite direction). Learners who had guided planning time in the first task improved the fluency of their subsequent, unplanned tasks by ten syllables per minute, so we might expect to

¹⁴ “Here-and-now” story-telling tasks (generally told in the present tense with visual support) are simpler and less cognitively demanding than the “there-and-then” tasks used in the present study. Previous research has demonstrated that “there-and-then” tasks promote accuracy and complexity over fluency (Gilbert, 2007; Robinson, 1995b).

see the fluency of learners in Group 1, who completed the first story-telling task without planning time, decrease during the second task, which was completed with guided planning time. However, learners who had no planning time for their first story-telling task improved their second, guided task by three syllables per minute. The act of repeating the task is what allowed them to improve their fluency under the most complex, guided condition when it came as the second task; the practice effect apparently outweighed the carry-over effect.

While the task with guided planning time was predicted to be the most helpful to learners in terms of offering processing support, it turned out to be the most challenging. This is likely due to the combination of requiring learners to tell a story without any visual support while simultaneously attempting to adhere to all the requirements and suggestions on the guided planning sheet. While other researchers have found that learners improve accuracy and fluency with guided planning time (e.g., Mochizuki & Ortega, 2008), those planning conditions were more prescribed. Learners were given specific examples of what to say and what structures to use, which would make the task easier and more supported. In this experiment, under the guided planning conditions, learners were asked to think of coordinate and subordinate clauses, make their descriptions as accurate and detailed as possible, and keep all of their verbs in the past tense, among other things, but they weren't given any examples of how to do so related to the stories they were telling, so while having planning time provided processing support, having detailed task requirements took it away.

6.2 Guided and Unguided Planning as Predictors of Fluency and Complexity

The second research hypothesis predicted that participants would produce more fluent and more complex language under (a) the guided planning condition and (b) the unguided planning condition than under (c) the no-planning condition. As discussed in Section 6.1, given that the least fluent speech under all conditions was generated as a result of guided planning time, part (a) of this hypothesis is clearly not supported. However, part (b), at least with respect to fluency, accurately predicted outcomes, as learners produced more fluent speech with unguided planning time than without planning time, regardless of whether they had that unguided planning time for the first story they told or the second (see Figures 2 and 4). This finding is in line with previous research on pre-task planning time, which has generally demonstrated improved fluency under unguided planning conditions (Ellis, 2009).

We have not yet discussed the complexity of learner speech, which was operationalized with Guiraud's Index, as well as the number of clauses per T-unit. With respect to Guiraud's Index, the results of the initial repeated-measures analysis of covariance showed a significant main effect for planning time and lexical complexity, with planned performance having more lexically complex speech than unplanned performance; once the group was divided into type of planning time/order of planning time, there was no statistically significant relationship between the various combinations of conditions, the descriptive statistics for which indicate that, overall, guided planning time generated more lexically complex speech than unguided planning time or no planning time. Both types of planning time were generally superior to unplanned time, with one exception, which occurred when

learners had unguided planning time before the no-planning condition. After having had unguided planning time for the first story-telling task, learners in this group (Group 2) produced unplanned speech that was very slightly more lexically complex than under the unguided planning condition ($M = 5.33$ with unguided planning time and $M = 5.39$ under the subsequent, no-planning condition). Given the influence of task repetition on learner performance across the dependent variables, it is likely that it was the act of repeating the story-telling task that kept the lexical complexity relatively equal across task conditions, rather than the no-planning condition itself fostering more lexically complex speech.

Structural complexity did not appear to be influenced by either task repetition or planning conditions; participants generated nearly the same number of clauses per t-unit regardless of whether or not they had pre-task planning time or whether planning time came first or second (see Table 12). This concurs with previous planning research, which has also failed to find an effect for structural complexity (e.g., Gilabert, 2007a, in a study on “here-and-now” versus “there-and-then” tasks with and without planning time). This finding is also supported by previous research on learner speech and accuracy, fluency, and complexity measures, which has suggested that structural complexity has more to do with the inherent narrative structure of a task, and less to do with task conditions (e.g., Tavakoli & Foster, 2011). Given the effect of both guided and unguided planning time on learners’ fluency and complexity, the second hypothesis accurately predicted participants’ fluency under unguided planning conditions, as well as their lexical complexity under both planning conditions.

6.3 Accuracy of Participants' Performance

The third research hypothesis predicted that learners would produce more accurate speech under the guided planning condition than under the unguided condition. The between-groups analysis of variance conducted to identify differences between participants in Group 1 (guided planning time) and Group 2 (unguided planning time) on just Task 2 (the first story-telling task) did not detect a significant difference between the groups in terms of the percentage of error-free clauses; the descriptive statistics reveal that Group 1, which had guided planning time, had an average of 61% error-free clauses, while Group 2, which had unguided planning time, had an average of 64% error-free clauses. When the percentage of error-free clauses with either type of planning time was compared to the percentage of error-free clauses without planning time in the repeated-measures analyses of covariance, there was no difference between the planning conditions, either. In other words, the number of error-free clauses produced by the participants does not appear to be related to the provision of planning time, the type of planning time, or the order of planning time.

One problem with accuracy measures is that it is impossible to know whether or not learners stand a chance of being able to improve their performance on them, especially in experimental settings such as these, where the researcher has no knowledge of what learners have been taught or of what they have previously demonstrated mastery. If learners do not know how to supply correct subject-verb agreement, put verbs into the past tense, or use plural markings appropriately, then no amount of pre-task planning time will help them improve their use of these structures.

For this reason, a measure of attempted accuracy was selected for inclusion in these analyses: self-corrections per 100 words. Following Gilabert (2007a, 2007b) as well as Shiau & Adams (2011), this measure was intended to capture learners' *focus* on accuracy. It was calculated by counting the number of times learners self-corrected their speech, dividing by the total number of words in the speech sample, and then multiplying by 100.

The initial analysis of covariance indicated a significant effect for number of self-corrections per 100 words and planning group (guided, unguided, or no-planning). In the between-groups ANCOVA for Task 2 only, learners in the guided planning group had the most self-corrections per 100 words ($M = 4.49$, $SD = 2.65$), learners in the no-planning group had a similar, but lower, number of self-corrections ($M = 3.92$, $SD = 2.83$), and learners in the unguided planning group had far fewer self-corrections per 100 words ($M = 2.33$, $SD = 1.78$); see Figure 1. As with the other dependent variables, task repetition appeared to influence this accuracy measure during the repeated measures analyses. For example, in three of the four conditions (guided planning first, no planning second; no-planning first, guided planning second; no-planning first, unguided planning second) learners produced an average of at least one self-correction more per 100 words during the first task they completed, regardless of whether or not planning time was offered. The one exception occurred when learners had unguided planning for the first task and then no planning time for the second task, in which case they produced nearly the same number of self-corrections per 100 words under both unguided and no-planning conditions.

To summarize, when offered as part of the first story-telling task, guided planning time fosters a greater focus on linguistic accuracy than unguided planning time. When planning time is offered before the second story-telling task, task repetition interacts with the planning condition to influence learners' focus on accuracy. There is no evidence that learners produce more accurate speech under any of the planning conditions; however, there is evidence that learners who are given guided planning time focus their attention on accuracy; whereas learners who are given unguided planning time do not, offering some support for the third research hypothesis, which predicted that learner output would be more accurate under the more complex guided planning condition than under the unguided planning condition, based on findings from other guided planning studies (e.g., Sangarun, 2005) and other studies of learners' self-corrections (e.g., Gilabert, 2007b; Shiau & Adams, 2011). Because the guided planning task was the most complex, participants' focus on accuracy and improved lexical complexity lends support to Robinson's Cognition Hypothesis, which predicts that increasing the cognitive demands of a task can shift focus toward accuracy and complexity at the expense of fluency (2001, 2005c, 2011).

In addition to the global accuracy measures of number of error-free clauses and self-corrections per 100 words, specific accuracy measures were also calculated to determine the extent to which learners were able to focus their attention on specific types of structures during the planning treatments. While there was no statistically significant effect for a difference between the types of errors made by each group, there were some small between-groups differences. The unplanned discourse had the fewest errors with subject-verb agreement and the most errors with plural nouns and

lexis. The discourse produced with guided planning time, during which learners were specifically asked to focus on avoiding errors with subject-verb agreement, lexis, and plural nouns, had the fewest errors with plural nouns, but fell in between the unguided condition and the no-planning condition with respect to the proportion of lexical and subject-verb agreement errors (see Figures 21, 22, and 23). Given the small differences among planning conditions, the lack of statistical significance, and the lack of an effect for the percentage of error-free clauses, it is not possible to say anything conclusive about how planning time affects the production of learner errors.

6.4 WMC and learner production

6.4.1 WMC and accuracy across planning conditions

The fourth research hypothesis predicted that WMC would play a role in participants' ability to attend to accuracy, regardless of the provision of pre-task planning time. Although planning time was predicted to improve accuracy (and guided planning time did generate more attempts at accuracy through self-corrections), given the difficult "There-and-Then" nature of the story-telling task along with the requirements to focus simultaneously on both content and form, learners with high WMC were predicted to perform more accurately than learners with low WMC. There were no between-groups or within-subjects differences in terms of the percentage of error-free clauses, and there was no interaction between planning time and the WMC covariate. However, because the initial repeated-measures ANCOVA identified an aptitude-treatment interaction between the provision of planning time and learners' WMC in terms of the number of self-

corrections per 100 words, the data were split so that learner performance within WMC groups could be considered along with their attempted accuracy.

As discussed in Section 6.3, learners in Group 1 (the one that received guided planning) tended to produce more self-corrections on the first story they told, regardless of whether they received planning time for the first task or the second task. However, once the participant pool was split into two groups by WMC (high and low), some differences emerged (see Figures 9 - 12). First, under the guided planning condition, the group of participants with high WMC did not follow the same pattern as the learners in the low-WMC group. Learners with high WMC produced significantly more self-corrections per 100 words under guided planning conditions regardless of whether planning time came first or second, while learners in the low-WMC group produced more self-corrections during whichever task came first, regardless of whether or not guided planning time had been offered. This dramatic difference in learner behavior across conditions partially supports the fourth hypothesis, which predicted that learners with high WMC would produce more accurate output with either guided or unguided planning time. While they did not demonstrate an improvement in their performances in terms of accuracy measures, they did demonstrate a better ability to focus on accuracy (per the planning instructions) than learners with lower WMC.

Learners in Group 2 (who completed the planned task under unguided conditions) behaved similarly, regardless of WMC (see Figures 11 and 12). They produced slightly more self-corrections per 100 words when they told stories without pre-task planning time, regardless of whether the planning time came first or second.

What is noteworthy about this group is that although the patterns were the same, learners with high WMC produced more self-corrections under both conditions than the learners with low WMC, who generated fewer self-corrections per 100 words under both unguided and no-planning conditions. These findings suggest that learners with high WMC were able simultaneously to monitor their performance and correct their speech in real time more frequently than learners with lower WMC (and, therefore, more limited processing resources), which is in line with theories of WMC explained by attentional control (e.g., Kane et al., 2007).

The differences among WMC groups with respect to the number of self-corrections per 100 words is logical given that this dependent variable demonstrates the ability to focus on one task in the process of interference. As Engle (2010) points out, individual differences in WMC boil down to the ability to maintain attention to the task at hand while simultaneously attending to interference from competing resources (523). When learners are monitoring and correcting their speech, they are paying attention to the story being told, the target discourse, and how their actual discourse differs from the target discourse. The “There-and-Then” story-telling task is ideal for highlighting these differences because learners are required to remember a narrative without visual support while also verbalizing it. As Robinson (2011) suggests, WMC might be especially significant in “There-and-Then” tasks (p. 23).

6.4.2 WMC and fluency

The fifth research hypothesis predicted that learners’ WMC would have more influence over their unplanned performance in terms of fluency than it would over their performance under either guided or unguided planning conditions. This

hypothesis was predicting that the facilitative effect of planning time would help equalize task conditions for learners with low WMC. The initial analyses of covariance revealed a significant interaction between planning time and WMC in terms of learner fluency when learners had pre-task planning time for the second story-telling task. When the data were split and the analyses re-run with WMC as a grouping variable, there was a significant interaction between +/- planning time and WMC on fluency; learners had the same mean fluency scores regardless of WMC without planning time, but learners with high WMC were less fluent than learners with low WMC under either guided or unguided planning conditions (see Figure 13). There was no significant interaction between type of planning (guided or unguided) or planning order and WMC group; however, given the small n-sizes when the groups were split and then divided into groups by WMC, it is possible the lack of significance was due to the low power associated with small n-sizes. There was a clear difference between learners' speech rates according to WMC among the four conditions (guided planning first, unguided planning first, guided planning second, unguided planning second).

Before discussing these differences, it is important to reiterate that both types of planning time were originally predicted to help compensate for individual differences in WMC (and, therefore, improve fluency), so this hypothesis predicted that participants with low WMC would have more fluent speech under guided and unguided planning conditions. However, as discussed in Section 6.1, learners who had guided planning time produced the least fluent speech, learners who had unguided planning time had the most fluent speech, and those who told stories

without the benefit of any pre-task planning time fell somewhere in between the two planning conditions. This is likely because the guided planners were asked to attend to both content and form, and attempted to do so, which slowed down their pruned speech rate while the unguided planners, free to direct their attention to linguistic features of their own choice, favored fluency.

Regardless of WMC, learners who were given guided planning time for the first story-telling task had a lower pruned speech rate after the guided planning time than when they told the second story without any planning time (see Figures 14 and 19). That is, when participants in Group 1 (the guided planning group) completed the second story-telling task without the provision of any planning time, they improved their fluency. Learners in the low WMC group were more fluent in both conditions than learners in the high WMC group. When the task order was switched, though, the pattern among WMC groups changed. Learners in the low WMC group improved their pruned speech rate by approximately 15 syllables per minute with guided planning time when it was offered for the second task (essentially improving the fluency of their performance for the second task they completed, regardless of whether the task was completed with guided planning time or no planning time). However, learners in the high WMC group slowed their speech for the second story-telling task when it was completed with guided planning time, mirroring the pattern of the high WMC group who had pre-task planning time for the first story-telling task.

The results of learner fluency with guided planning time suggest that learners in the high WMC group are affected by the guided planning conditions and are

attempting to focus their discourse on form and content simultaneously, which slows their pruned speech rate. On the other hand, learners who have lower WMC appear to benefit from task repetition, improving the fluency of their second stories regardless of whether planning time is offered for the first task or the second task. In other words, with respect to learners' fluency, task repetition appears to be the treatment that maximizes fluency for learners with low WMC, whereas learners with high WMC, who attend to the task requirements regardless of the order of the provision of planning time, benefit from unguided planning time when the goal is to improve fluency. The difference among WMC groups, task conditions, and the order of the provision of planning time suggests a complex relationship between task conditions, individual differences in learners' WMC, and learner output.

Manipulating task conditions by offering unguided planning time allowed learners more time to prepare for the story-telling tasks without specifically directing them to focus on form or content; the results for all participants support previous planning research and suggest that learners prioritize fluency when given unstructured planning time (Ortega, 1999; Robinson, 2005a). However, the results presented here also suggest that, in terms of fluency, learners with low-WMC benefit more from unguided planning time than learners in the high-WMC group. When learners in the low-WMC group had unguided planning time for the first story-telling task, their speech rates were higher than the high-WMC group (see Figure 16) and fluctuated significantly more when planning time was taken away for the second story-telling task (see Figure 20). Learners in the high WMC group in Group 2 (unguided planning time) were much more consistent across planning conditions, with a mean

pruned speech rate of 120 syllables per minute with unguided planning time and then a mean pruned speech rate of 117 syllables per minute in the subsequent, unplanned condition. With the other half of the participants in Group 2 (unguided planning time), learners in the low WMC group improved the fluency of their planned performances during the second, unplanned condition, whereas learners with high WMC were slightly less fluent on the second, unplanned task. These results suggest that learners with low WMC shift their focus to improving their speech rate whereas learners with high WMC were less focused on speech rate (presumably because they were attending to content or form).

The fifth research hypothesis predicted that WMC would play a larger role in the fluency of learner discourse under the unplanned conditions, and therefore learners with low WMC would be able to be more fluent with planning time than they would without planning time. The results partially support this. Learners with low WMC were far more fluent with unguided planning time than they were under the no-planning conditions, and they were more fluent than learners with high WMC under both conditions. Learners with high WMC had relatively similar pruned speech rates regardless of whether they completed tasks with unguided planning time or no planning time, again suggesting that improving speech rate was not the focus of their attention when they were offered planning time.

As discussed in section Section 6.2, overall, within subjects, the guided planning condition generally resulted in less fluent discourse than story telling without planning time. However, learners with low WMC did not follow this pattern; they improved their fluency for the second task they completed, regardless of whether

or not planning time was offered. In other words, the pruned speech rate of these learners did not appear to be affected by guided planning time, which does not support the fifth research hypothesis. This is likely because offering guided planning time made the task even more complex. While learners were given time to plan, they were also told what to do, and their attention was directed to content and form. It is possible that learners with low WMC were unable to focus their attention on both content and form simultaneously, so they were unable to capitalize on the benefits of the resource-directing nature of guided planning time.

6.4.3 WMC and speech complexity

The final research hypothesis was that WMC would have a greater influence on speech complexity under the unplanned condition than either of the planning conditions. The between-groups differences with respect to complexity were the least straightforward, probably because there was no significant interaction between planning and WMC on either lexical or structural complexity. Because the differences between and within groups were not statistically significant, these data must be interpreted cautiously; however, they reveal some trends with WMC that align with observations presented in Sections 6.4.1 and 6.4.2.

With respect to lexical complexity, learners in the low WMC group followed the same pattern regardless of whether or not they received either type of planning time for the first task or the second: their lexical complexity improved for the second story-telling task. This pattern of learners with low WMC improving their performance simply as a result of repeating the story-telling task is consistent with what was observed with guided planning and fluency and self-corrections. However,

for both fluency and self-corrections, learners with low WMC improved their performance when given unguided planning time; with lexical complexity, we do not see any influence of unguided planning time. Learners with high WMC demonstrated less predictable patterns in their lexical complexity. When either type of planning time was offered first, they improved their lexical complexity slightly on subsequent, unplanned tasks. However, when the unplanned condition was presented first, learners with high WMC improved their lexical complexity when presented with guided planning time for the story-telling task, but they decreased their lexical complexity when they were presented with unguided planning time. Given the lack of an effect for the WMC covariate or the between-groups WMC grouping variable, these data merely suggest trends in group behavior and should not be interpreted as indicative of what can be expected from groups based on WMC and lexical complexity.

The structural complexity variable of clauses per t-unit remained nearly uniformly consistent across planning conditions (see Table 12). When the data were split into WMC groups; however, there were larger within-groups differences between two of the planning conditions. Learners with high WMC improved their structural complexity under the guided planning condition, and learners with low WMC improved their structural complexity under the unguided planning condition. There were no differences between the guided planning condition and the no-planning condition, neither for high-WMC nor for low-WMC learners.

Because the differences within and between groups on both the structural and lexical complexity measures were not statistically significant, we must exercise

caution in basing generalizations on them. Despite this, because they support patterns evident through the other between-groups WMC analyses, they can help us understand how learners are allocating their attention under the various planning conditions. As far as lending support to the sixth research hypothesis, there is no evidence that WMC interacts with task conditions to affect lexical complexity, but there is some evidence to suggest that for learners with low WMC, unguided planning time can improve structural complexity, and for learners with high WMC, guided planning time can offer similar results.

6.7 Synthesis of Results by Planning Condition and WMC

The relationship among planning conditions (guided, unguided, or none), WMC, and task repetition is complex, and the results of the present study suggest that all of these factors play an interrelated role as predictors of learner accuracy, fluency, and complexity. As each of the dependent variables was considered separately throughout the discussion thus far, this section will summarize the effect of each condition on participants' output as a whole.

6.7.1 Guided Planning Time

Guided planning time promoted a focus on attempted accuracy and lexical complexity, as evidenced by the statistically significant increase in self-corrections and Guiraud's Index. However, there was a trade-off with fluency, and this condition resulted in lower pruned speech rates. With respect to working memory, learners with high WMC generated the most self-corrections under the guided planning

condition and had a slightly lower speech rate than participants with low WMC, who generated far fewer self-corrections.

6.7.2 Unguided Planning Time

Unguided planning time promoted a focus on fluency for all learners, and the fastest pruned speech rates were recorded whenever learners were offered time to plan without parameters or suggestions. Learners with high WMC had slower pruned speech rates during unguided planning time than learners with low WMC, but they had far more episodes of self-correction. Learners with low WMC had very few self-corrections when they completed the stories with unguided planning time.

6.7.3 No Planning

In general, stories told without planning time included more self-corrections than stories told with unguided planning time, but fewer than those told with guided planning time. They were also less fluent than those told with unguided planning time, but more fluent than those told with guided planning time (see, e.g., Figure 2). WMC played a significant role in the number of self-corrections generated without planning time; learners with high WMC had far fewer self-corrections under no-planning conditions than learners with low WMC. When they were not offered any planning time, learners in both WMC groups generated discourse with similar pruned speech rates.

6.7.4 Task Repetition

Learner output on the second story-telling task under all conditions (guided, unguided, and no-planning) was affected by learners' having immediately performed

the first story-telling task, and there was a significant interaction between +/- planning time and the order of the provision of planning time. In Group 2 (guided planning), learners self-corrected fewer errors for their second stories, regardless of the task conditions, and their pruned speech rates were closer to those of their first stories. When learners began without planning time, their accuracy, fluency, and complexity ratings were very similar to one another. However, when they completed a guided task first, they tended to prioritize self-corrections and complexity under the second unguided task, whereas learners who began with an unguided task prioritized fluency and lexical complexity under the subsequent, no planning conditions. The effect of task repetition is more pronounced once WMC is taken into consideration.

6.7.5 Task Repetition and WMC

Learners with high WMC did not follow the same pattern of fewer self-corrections under the second story-telling task; when they were offered guided planning time for the second task, they made far more self-corrections than learners in the low WMC group (see Figure 10). The same pattern is evident with pruned speech rate; learners in the low WMC group improved their fluency for the second task, regardless of conditions, while learners in the high WMC group slowed their speech rate to cope with the requirements of the planning time. Finally, with respect to lexical complexity, learners in the low WMC group improved their score on Guiraud's Index for the second task they completed, regardless of the task conditions, whereas learners with high WMC had levels of lexical complexity that fluctuated with task conditions, regardless of planning order.

6.7.6 WMC

When learners with high WMC were offered the guided or no-planning conditions, they produced far more self-corrections with guided planning time, regardless of task order (suggesting that they prioritized self-corrections based on the planning instructions, as appropriate). They were also more consistent between groups with respect to their performance on guided and unguided tasks; in other words, the order of the provision of planning time had less of an impact on the performance of the learners with high WMC. Learners with low WMC were more affected by the order of the story-telling tasks than by the task conditions. This is especially evident with self-corrections (see Figure 10). While learners with high WMC attempted to focus on both content and form with guided planning time, with trade-offs in fluency, learners with low WMC did not. In other words, the conditions that promote a focus on accuracy and fluency are not the same for learners with high and low WMC; with low WMC, it makes more sense to give learners time to repeat the same task multiple times without specific instructions, whereas with high WMC, learners are able to take advantage of the benefit of guided planning instructions.

6.8 Conclusions and Directions for Future Research

This study was undertaken with three main goals: to confirm previous findings that the order of the provision of planning time influences learner production; to determine whether or not learners' WMC interacts with planning conditions to affect performance on a complex, "there-and-then" narrative; and to investigate the relative benefits of guided and unguided pre-task planning on the

accuracy, fluency, and complexity of learners' discourse. The results demonstrate that task order had a clear effect on learners' production. When learners began the series of story-telling tasks with either type of planning time, their output on the subsequent, unplanned task varied according to whether they had first received guided or unguided planning time. Figures 7 and 8 illustrate this difference particularly well, demonstrating that learners who had completed a previous story with guided planning time prioritized attempted accuracy and complexity on the second task, whereas learners who had begun with unguided planning time prioritized fluency the second time around.

With respect to the role of WMC, this study clearly demonstrates that task conditions can affect learners with high and low WMC in different ways. Learners with high WMC are more likely to monitor and correct their speech when necessary for task requirements (e.g., when complying with guided planning instructions to prioritize content and form), and at the same time, they produce fewer self-corrections under unguided and no-planning conditions. Learners with low WMC, on the other hand, produce self-corrections as a result of task order, lowering the number of self-corrections and improving their fluency when they tell the second story in a sequence, regardless of task conditions. These differences in performance between groups of learners with high and low WMC is consistent with the notion that WMC is at least in part an issue of attentional control (Kane, Bleckley, Conway, & Engle, 2001; Kanet et al., 2007). That is, learners with high WMC are better able to direct their attention to complex demands (demonstrated by more attention to attempted accuracy to comply with task requirements) whereas learners with low WMC are unable to direct their

attention in this way and therefore respond to the provision of planning time with output that requires less attentional control: improved fluency.

Guided planning time and unguided planning time also have very different effects on learners' production. Under the most complex, guided condition, learners focused on accuracy and lexical complexity at the expense of fluency, in line with Robinson's Cognition Hypothesis (2001, 2005c, 2011), which predicts that increasing task complexity along resource-directing lines will improve accuracy and complexity. When learners were able to direct their own planning time, they prioritized fluency over accuracy, and demonstrated better lexical complexity than under the unplanned condition, but lower scores on Guiraud's Index than under guided planning conditions. This is in keeping with previous research on planning time, which often finds an effect for fluency under unguided conditions, probably due to learners' prioritization of fluency over accuracy (Skehan 1998, 2009).

While this study demonstrated that learners can be pushed to *focus* on accuracy through manipulations in task conditions, it did not demonstrate any improvement in the accuracy of learners' performance or in the types of structures they attempted to improve. This is in line with many other planning studies, which fail to show improvements in accuracy measures as a result of planning time (see Mochizuki & Ortega, 2008 for a review). It is likely that while learners attempted to produce more accurate language, they needed more practice and more feedback on error in order to actually be able to do so.

This dissertation research has clear implications for instructed SLA. First, it is possible to manipulate learner performance in classroom settings by changing task

conditions, allowing instructors to devise sets of tasks that shift learners' focus to accuracy, fluency, and complexity so that they can have balanced amounts of production practice. Further, learners' WMC interacts with task conditions; learners with low WMC do not demonstrate the same simultaneous attention to content and form (as evidenced by self-corrections) on the most complex, guided planning task. Therefore, in heterogeneous classrooms with learners of varying WMC—like the ones chosen for this study—it may be necessary to offer more processing support during guided planning, perhaps by allowing learners to keep their notes with them during the story-telling task, in order to achieve the same results across WMC groups. However, in instructed situations where it is possible to divide learners into groups according to their WMC (as in the groups established for the study presented here), treatments can be tailored to learners' individual differences, so students with high WMC can be offered complicated instructions for task completion, promoting a focus on accuracy and complexity, whereas learners with lower WMC can be given fewer instructions but more time to repeat the story-telling tasks.

These pedagogical suggestions have clear implications for future research because there is a need for more classroom-based investigation of the longitudinal effects of manipulating task conditions. As Ellis (2012) points out, moving away from laboratory-based experiments, where participants work one-on-one with researchers, is an important step in examining the appropriateness of modifying task conditions in instructed settings, and the research discussed in this dissertation does just that. However, because all of these manipulations are intended to offer learners the chance to practice using the target language in a balanced way, their effects must

be considered as they accumulate over time. A longitudinal, classroom-based study that considers learners' WMC; their performance on tasks under guided, unguided, and no-planning conditions over time; and their language proficiency before and after semester- or year-long treatments is critical to understanding how these task manipulations affect acquisition, and not just performance. Further, if planning research along the lines outlined above were to take place over a semester or a year, then instructors could take specific accuracy measures into consideration, as well. They could use changing patterns in learners' accuracy to gauge how well the various treatments (task order, guided vs. unguided planning) promote an actual shift in accuracy. Finally, longitudinal research of this type would allow instructors to contrast even more types of task conditions (comparing, for example, a "Here-and-Now" task with guided planning time to a "There-and-Then" task with unguided planning time), offering the chance to collect more evidence to test the Cognition Hypothesis.

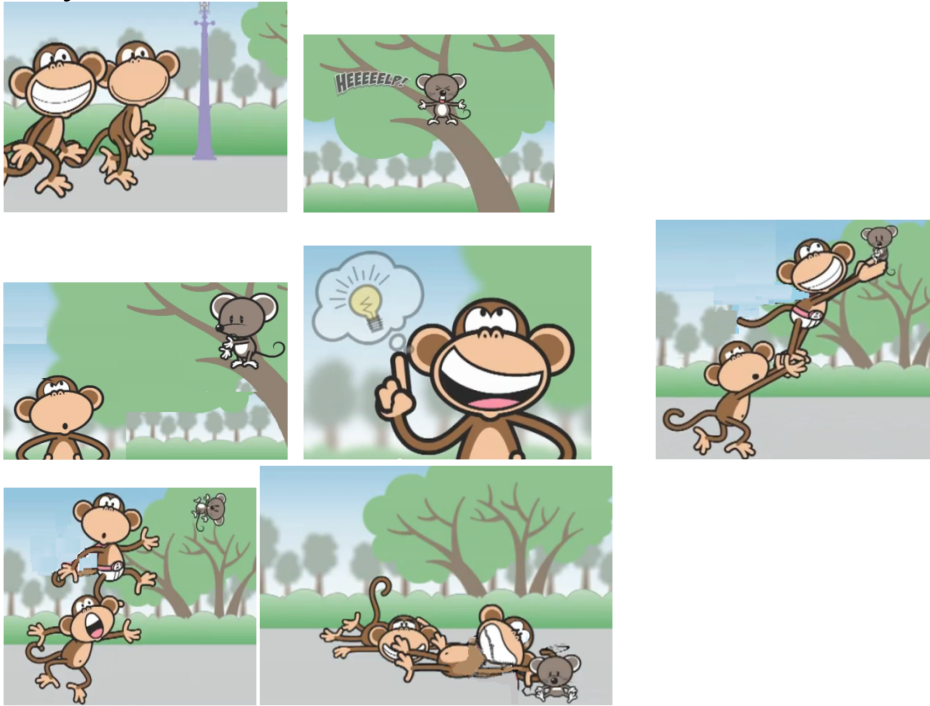
Appendices

Appendix A: Language background questionnaire

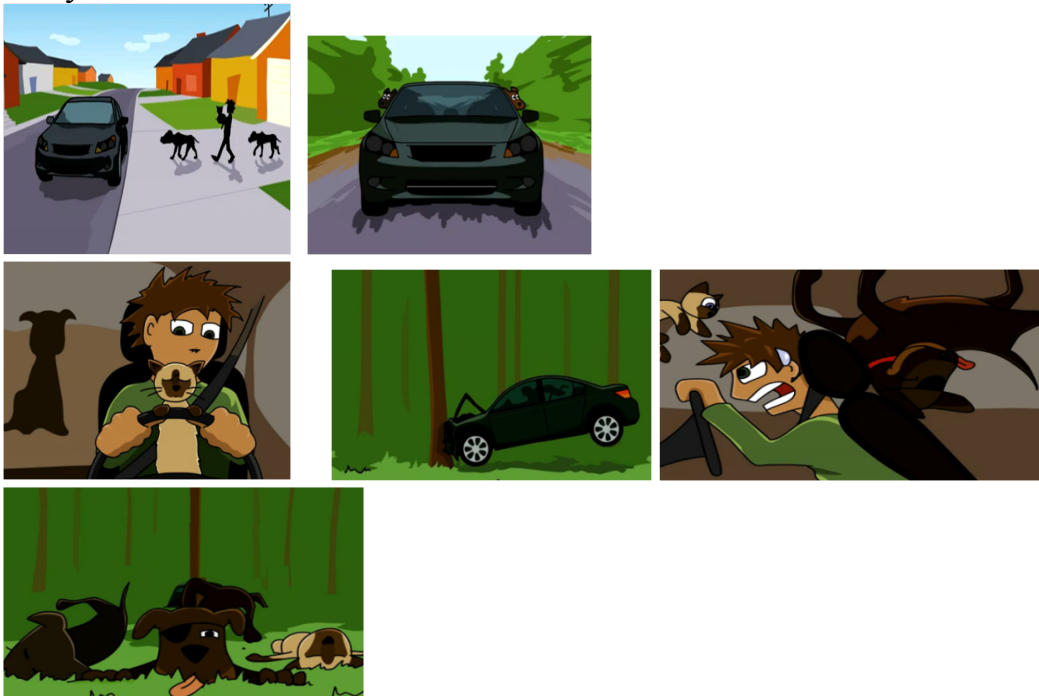
- 1) What is your native language?
- 2) How long did you study English before coming to the United States (in months and years)?
- 3) How long have you been in the United States (in months and years)?
- 4) What was the first ESL class you took at this community college?

Appendix B: Images for story-telling tasks A and B

Story A:



Story B:



Appendix C: Instructions for guided planning time

Instructions for Guided Planning Time

You have 10 minutes to plan your story according to the following directions. Please use this paper to take notes if you wish; you will not be able to keep the paper or the notes with you when you record yourself telling the story in 10 minutes.

- 1) Think about all of the details you want to include in your story. What happens first, second, third?

- 2) Review the images and think about how you will describe each one; try to make your descriptions specific. For example, instead of saying “the ball is on the table,” you could say “the small red ball that is on the table is about to fall off.”

- 3) Think of the transition words you will use to tell your story; remember to use coordinate and subordinate clauses during your story.

- 4) Think about how to keep your story in the past tense. What verbs will you need?

- 5) When you practice telling your story, make it as detailed and accurate as possible. Make sure that you keep the story in the past tense, that your subjects and verbs agree, that you differentiate between singular and plural nouns, and that you choose appropriate words to describe the images.

Appendix D: Instructions for unguided planning time

Instructions for Unguided Planning Time

You have 10 minutes to plan your story. Please use this paper to take notes if you wish; you will not be able to keep the paper or the notes with you when you record yourself telling the story in 10 minutes.

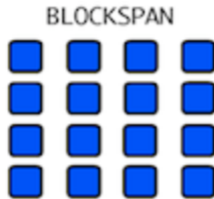
Appendix E: Instructional handout given to participants

Thank you for participating in the story-telling project! First, you will complete two games to test your memory; then you will tell two stories.

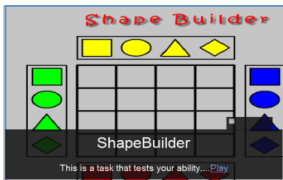
- 1) Go to this website: <http://tiny.cc/nielson1>
- 2) Click on “WM Task: HERE”

Your login is: **1003KNSLA2_**
Your password is: **my1fun**

- 3) Complete this game first:

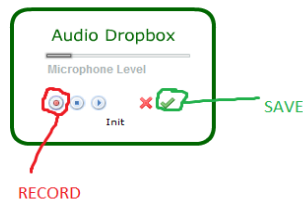
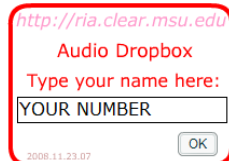


- 4) Complete this game second:



- 5) After you complete the two games, go back to this website: <http://tiny.cc/nielson1>
Follow the directions on the main webpage to complete the story-telling tasks.
When you are ready to record your stories:

Click here to record your story:



References

- Atkins, S. M., Harbison, J. I., Bunting, M. F., Teubner-Rhodes, S., & Dougherty, M. R. (2009). Measuring working memory with automated block span and automated letter-number sequencing. Poster presented at the 50th Annual Meeting of the Psychonomic Society, Boston, MA.
- Badcock, J., Michie, P., & Rock, D. (2005). Spatial working memory and planning ability: Contrasts between schizophrenia and bipolar disorder. *Cortex*, 41(6), 753 - 763.
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189 - 208.
- Baddeley, A. & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Vol. 8*, (pp. 47 - 90). New York: Academic Press.
- Bygate, M. (2001). Effects of task repetition on the structure and control of language. In M. Bygate, P. Skehan & M. Swain (Eds.), *Task-based learning: language teaching, learning and assessment* (pp. 23 - 48). London: Longman.
- Bygate, M. & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.) *Planning and task performance in a second language* (pp. 37 - 76). Amsterdam/New York: Benjamins.
- Christoffels, I. (2006). Listening while talking: The retention of prose under articulatory suppression in relation to simultaneous interpreting. *The European Journal of Cognitive Psychology*, 18, 206 - 220.
- Conway, A.R.A., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O., & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769 - 786.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62 - 101). Cambridge: Cambridge University Press.
- Cronbach, L.J. & Snow, R.E. (1977) *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Crookes, G. (1989) Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 367 - 383.
- Daneman, M. & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450 - 466.
- Daneman, M. & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25(1), 1 - 18.

- DeKeyser, R. M. (2007). Introduction: Situating the concept of practice. In R. DeKeyser (Ed.), *Practice in a second language. Perspectives from applied linguistics and cognitive psychology* (pp. 1 - 18). Cambridge: Cambridge University Press.
- DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 119 - 138). Massachusetts: Blackwell.
- Dörnyei, Z. (2005). *The psychology of the language learner*. New Jersey: Lawrence Erlbaum.
- Doughty, C. & Williams, J. (1998) *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Doughty, C., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, 7(3), 50 - 80.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics* 30(4), 474 - 509.
- Ellis, R. (2012). *Language teaching research and language pedagogy*. Chichester: John Wiley & Sons.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in WMC and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102 - 134). Cambridge: Cambridge University Press.
- Engle, R. W. (2010). Role of Working Memory Capacity in Cognitive Control. *Current Anthropology*, 51(s1), 17 - 26.
- Gass, S. (1997) *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.
- Gass, S., Mackey, A., Alvarez-Torres, M. J. & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49(4), 549 - 581.
- Gilabert, R. (2007a). The simultaneous manipulation of task complexity along planning time and (+/- Here and Now): Effects on L2 oral production. In M. Garcia-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44 - 68). Clevedon: Multilingual Matters.
- Gilabert, R. (2007b). Effects of manipulating task complexity on self-repairs during L2 oral production. *IRAL*, 45, 215 - 240.
- Gilabert, R. & Muñoz, C. (2010). Differences in attainment and performance of a second language: The role of WMC. *International Journal of English Studies*, 10(1), 19 - 42.
- Gilhooly, K., Wynn, V., Philips, L., Llogie, R., & Della Salla, S. (2002). Visuo-spatial and verbal working memory in the five-disc Tower of London task: An individual differences approach. *Thinking & Reasoning*, 8(3), 165 - 178.

- Guará-Tavares, M. G. (2009). The relationship among pre-task planning, working memory capacity, and L2 speech performance: a pilot study. *Linguagem & Ensino*, 12(1), 165 - 194.
- Harrington, M., & Sawyer, M. (1992). L2 WMC and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25 - 38.
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, 24, 541 - 577.
- Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24(2), 168 - 196.
- Kane, M.J., Bleckley, M. K., Conway, A. R. A., & Engle, R.W. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General* 130, 169 – 183.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in WMC as variation in executive attention & control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. N. Towse (Eds.), *Variation in working memory* (pp. 21 - 48). Oxford: Oxford University Press.
- Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate proficiency. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 143 - 164). Philadelphia: Benjamins.
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11(2), 261 - 271.
- Kormos, J. & Trebits, A. (2011). Working memory capacity and narrative task performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 267 - 286). Amsterdam : Benjamins.
- Long, M. H. (1981). Input, interaction and second language acquisition. In Winitz, H. (Ed.), *Native language and foreign language acquisition. Annals of the New York Academy of Sciences*, 379, 259 - 278.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In Bot, K. de. Ginsberg R. B & Kramsch C. (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39 - 52). Amsterdam: John Benjamins.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413 - 468). San Diego: Academic Press.

- Long, M. & Crookes, G. (1993). Units of analysis in syllabus design: The case for task. In G. Crookes & S. Gass (Eds.), *Tasks in a pedagogical context: Integrating theory and practice* (pp. 9 - 54). Clevedon: Multilingual Matters.
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.) *Individual differences and instructed second language learning* (pp. 181 - 209). Amsterdam: John Benjamins.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 52 - 83.
- Mochizuki, N. & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research* 12(1), 11 - 37.
- Morrison, A. & Chein, J. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, 18, 46 - 60.
- Newman, S., Carpenter, P., Varma, S., & Just, M. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computation modeling of planning and high level perception. *Neuropsychologia*, 41(12), 1668 - 1682.
- Nielson, K. (in press). Can planning time compensate for individual differences in working memory capacity? *Language Teaching Research*.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29, 557 - 582.
- O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27(3), 377 - 402.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109 - 148.
- Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77 - 109). Philadelphia: Benjamins.
- Owen, A., Downes, J., Sahakian, B., Polkey, C. & Robbins, T. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, 28(10), 1021 - 1034.
- Pallotti, G. (2009). CAF: Defining, refining, and differentiating constructs. *Applied Linguistics*, 30(4), 590 - 601.

- Payne, J. & Ross, B. (2005). Synchronous CMC, working memory, and L2 proficiency development. *Language Learning and Technology*, 19(3), 35 - 54.
- Payne, J. & Whitney, P. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, 20, 7 - 32.
- Robinson, P. (1995a). Attention, memory, and the "noticing" hypothesis. *Language Learning*, 45(2), 281 - 331.
- Robinson, P. (1995b). Task complexity and second language narrative discourse. *Language Learning*, 45(1), 99 - 140.
- Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287 - 318). New York: Cambridge University Press.
- Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA. In P. Robinson. (Ed.), *Individual differences and instructed language learning* (pp. 211 - 266). Philadelphia, PA: John Benjamins.
- Robinson, P. (2005a). Aptitude and second language acquisition. *Annual Review of AppliedLinguistics*, 25, 46 - 73.
- Robinson, P. (2005b). Attention and Memory during SLA. In C. Doughty and M. Long (Eds.), *The handbook of second language acquisition* (pp. 631 - 678). Massachusetts: Blackwell.
- Robinson, P. (2005c). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43, 1 - 32.
- Robinson, P. (2007). Aptitudes, abilities, contexts, and practice. In R. DeKeyser (Ed.), *Practice in a second language. Perspectives from applied linguistics and cognitive psychology* (pp. 256 - 286). Cambridge: Cambridge University Press.
- Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3 - 38). Amsterdam : Benjamins.
- Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111 – 141). Amsterdam: John Benjamins.
- Sawyer, M. & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319 - 353). New York: Cambridge University Press.

- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 206 - 226.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, Proceedings of CLaSIC 2010, Singapore, December 2 - 4 (pp. 721 - 737). Singapore: National University of Singapore, Centre for Language Studies.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- Shiau, Y. & Adams, R. (2011). The effects of increasing reasoning demands on accuracy and complexity in L2 oral production. *University of Sydney Papers in TESOL*, 6, 121-146.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510 - 532.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed second language learning* (pp. 69 - 93). Amsterdam: Benjamins.
- Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93 - 120.
- Skehan, P. & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183 - 205). New York: Cambridge University Press.
- Sprenger, A. M., Atkins, S. M., Colflesh, G. J. H., Briner, T. L., Buchanan, J. B., Chavis, S. E., Chen, S., Iannuzzi, G. L., Kashtelyan, V., Dowling, E., Bolger, D. J., Bunting, M. F., & Dougherty, M. R. (in preparation). A four-dimensional video game for measuring cognitive ability.
- Sunderman, G. & Kroll, J. (2009). When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, 30, 70 - 99.
- Swain, M. (1995) Three functions of output in second language learning. In G. Cook & B. Siedlhofer (Eds.), *Principle and practice in applied linguistics*. Oxford, UK: Oxford University Press, 125 - 144.
- Tavakoli, P. & Foster, P. (2011). Task Design and Second Language Performance: The Effect of Narrative Type on Learner Output. *Language Learning*, 61(s1), 31 - 72.

- Unsworth, N. & Engle, R. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114, 104 - 132.
- Van Patten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287 - 301.
- Van Patten, B. (1996). Input processing and grammar instruction. New York: Ablex.
- Williams, J. (2011). Working memory and SLA. In S. Gass & A. Mackey (Eds.), *The Routledge Handbook of second language acquisition* (pp. 427 - 441). New York: Routledge.
- Yuan, F. & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24, 1 - 27.