The Effects of Cognitive Task Complexity on Writing

Mark Wain Frear

A thesis submitted to Auckland University of Technology in fulfillment of the requirements
for the degree of Doctor of Philosophy (PhD)

December, 2013

Department of Language and Culture

Supervisors: Professor John Bitchener and Dr. Dale Furbish

i

TABLE OF CONTENTS

v

LIST OF TABLES

LIST OF FIGURES

ATTESTATION OF AUTHORSHIP

_____

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Auckland, December 2013

_____

Mark Wain Frear

ACKNOWLEDGEMENTS

Firstly, my thanks go to Professor John Bitchener who provided not only the opportunity but also the guidance to get me through this PhD. John is the teachers' teacher, and I am grateful for the time he has invested in my work. My appreciation also goes to my other supervisor, Dr Dale Furbish, for his time and his much appreciated feedback.

I would also like to express my gratitude to Dr Eric Landhuis for his invaluable help on aspects of the statistical analyses.

Recognition should also go to AUT University where I was provided with work and financial assistance during the course of this thesis.

SPECIAL ACKNOWLEDGEMENTS

Jon Donne wrote: "no man is an island", and though writing a PhD often feels like an island, the truth is that a PhD is a communal endeavour that requires the forbearance of those around you.

The most thanks must go to my mother Jan Bakker and her husband Marten Bakker. Jan and Marten provided me with a place to sleep and eat for the many, many long months it took to finish this work. Without their help and support, finishing on time would have been impossible.

My brother, Dr David Frear, was an invaluable asset during the early stages of the thesis and has always been a stalwart companion. My thanks also go to his wife Amy, my beloved nephew and niece, Jack and Joy, and my sister Rachel who always fronted up with the cash when times got tight.

Recognition must also go to the people who supported me with food, coffee, movies, and friendship: Rudd Hughes, Karen O'Shea, Deidre Alderson, Meryn Crocker-London, Comic-shop Kelly Sheehan, Dolomite, and Bruce Kendall.

Finally, to my father Gary Frear, who passed away before the completion of this thesis, I dedicate my final word.

ABSTRACT

This thesis contributes to research on the effects of cognitive task complexity on the written complexity of second language learners, an area currently underrepresented in research. In addition, the variables pre-task planning time and post-task editing time, which bookend the writing process, are investigated, as are the potential effects of task motivation on syntactic output.

This research was conducted on 94 intermediate English learners from three Auckland language schools. Two variations of the positivist/normative approach were used: an experimental model investigating cause and effect between the main variables, and an associative model to explore the strength of association between task motivation and syntactic output. Participants were placed in three groups in which three letter-writing tasks of varying cognitive complexity and Likert-scale questionnaires were performed over two sittings. Each group had different conditions; these were no planning time, 10 minutes pre-task planning time, and 10 minutes post-task editing time. Additional factors investigated were the potential benefits of using a patently low complexity task, the effects of modality, the use of a non-standard measure of syntactic complexity, and the efficacy of Robinson's predictions for resource-directing and resource-dispersing variables.

The results suggested that increases in cognitive task complexity may adversely effect dependant clause production, but benefit lexical production; however, the inclusion of pre-

task planning time appeared to aid dependant clause production while slightly

decreasing lexical complexity. Proficiency issues rendered the post-task data unusable.

Potential support for Robinson prediction was posited for lexical complexity in the non-

planning group, and both syntactic and lexical complexity in the planning time group;

however, this support came with the caveat that the results were framed within a limited

attentional framework.

Overall, the patently low complexity task appeared useful, though the potential for the low

complexity task to be a different task type remains a potential issue. Modality only

appeared to benefit task complexity development when it was compounded with pre-task

planning time. The new measure of syntactic complexity revealed significant results not

noticeable using the older measure.

Regarding task motivation and subordination, negative perceptions of task relevance and

expectancy may have contributed to decreases in motivation and thus attention demanding

syntax. However, a further negative construal was posited to have elicited either positive or

neutral results. Negative task construal, and its theorised reduction of attention maintenance

across Dörnyei's (2002, 2003, 2009, 2010) motivational task processing system, may be

especially effective during the attention demanding formulation of dependent clauses under

increasing cognitive task complexity.

CHAPTER ONE

INTRODUCTION

1.1 Issues of interest

Writing is an important skill for second language learners who wish to study at university level in a non-native language. It could be considered an essential means of communication with which university students interact and exchange knowledge with teachers.

There are many second language students, studying at an intermediate level of language, who are looking towards improving their writing skills to a quality commensurate with those required for university education. As such, this study spotlights these students, with a particular focus on the development of complex (elaborate and varied language) writing skills.

In this thesis, four dimensions (complex task design, complex output, pre-task planning, and post-task editing) that could be considered important elements in the development of complex writing skills in intermediate second language learners are explored. Firstly there is *complexity*, which refers to both the modifications during task design that make a task difficult to complete, and language in the written output that can be considered varied and elaborate (Ellis, 2003).

Complexity as language output can be viewed as important because it is posited to push learners to create more elaboration and structure in their developing language ability, make language use more efficient, bring the language being learnt in line with target language

use, enable the efficient and precise expression of complex ideas, and can be a sign that acquisition is taking place (Cheng, 1985; Cook 1994; McLaughlin, 1990; Skehan, 1996; Swain, 1985).

The second and third dimensions related to writing development are *pre-task planning time* and *post-task self-editing time*, which are two activities that bookend the writing process that may have an effect on the production of complex language. Finally, *task motivation*, which focuses on a participant's attitudes towards the tasks they are performing, is addressed because the attitude a student has towards a task may have subsequent effects on performance (Dornyei & Kormos, 2000).

Complexity is a central term in this research, and as mentioned above, it can be understood in two ways. Firstly, complexity can appear in the written samples of students, which is *complexity as output*. In this thesis, complex output specifically refers to the development of syntactic (a measure of subordinate clauses) and lexical (a measure of lexical variety) items in a students' work. The syntactic elements in particular are viewed as an important developmental stage through which learners move as their written skills develop (see 2.2.5.8). Secondly, complexity can be viewed as an element of task design, which is complexity as input. Complexity as input is termed *cognitive task complexity* and it refers to the elements manipulated during task design that may cause increases in a participants' cognitive workload when trying to complete the task.

In this thesis, cognitive task complexity was modified in the task instructions by increasing *reasoning demands* (in which participants had to form opinions based on provided information), and by increasing the *number of elements* (which was the number of items

needing to be considered while forming an opinion). Reasoning demands and number of elements have been described as *resource-directing* task dimensions (Robinson, 2001a, 2005, 2007a, 2007b). Robinson claims that increasing the complexity of a task along resource-directing dimensions should result in increases in the occurrence of complexity as output (for example, the types of lexical and syntactic measures used in this thesis).

Central to this thesis is the causal relationship between both types of complexity and how the relationship between cognitive task complexity and complexity as output can be tested against Robinson's (2001a, 2001b, 2005, 2007a, 2007b) Triadic Componential Framework, a typology that has used cognitive task complexity as a component for sequencing and grading tasks.

A review of the literature on task-based language learning, in which complexity is manipulated in the input, has revealed that most of the work is in the oral modality (spoken language). Additionally, the smaller body of research that does focus on complexity and the written modality appears skewed towards analysing accuracy (whether the language used conforms with language norms) in the output as opposed to complexity (varied and elaborate language). This could be viewed as problematic if researchers and teachers are to develop a comprehensive knowledge of tasks and their effects on accuracy, complexity, and fluency (Skehan, 1998), thus addressing the need for the balanced development of each of these dimensions in a students' emerging language abilities.

Pre-task planning time is one of two types of planning time that have been the focus of language learning research. Broadly, there is the type of planning that happens while a task

is being performed, and there is the type of planning that happens before a task is performed (Ellis, 2005). This thesis focuses on the pre-performance planning aspect, specifically the type of pre-task planning referred to as *strategic planning*, which is where students are provided time to prepare for a task. This preparation might include thinking about task content (Ellis, 2005) and the language required to complete the task. In the small group of research that has used strategic pre-task planning time in conjunction with cognitive task complexity, there have been favourable results for complex output (Ellis & Yuan, 2004; & Ishikawa, 2006). As a result, pre-task planning has been added as an extra variable with the understanding that it may contribute to positive complex outputs (increases in elaborate and varied language).

Post-task self-editing time is an activity in which the participants are provided time to change any of the output written during the performance of the task. There is a large body of work concentrating on editing as it relates to the accuracy elements of writing; however, there appears to be no research focusing specifically on post-task editing and its effects on complex output.

Task motivation is the final element addressed in this thesis and it differs from the other parts of the research because it focuses in how task performance is affected by an individual's personal resources such as aptitude, confidence, motivation, or intelligence (Ellis 2003). This section is motivated by claims that the environment in a classroom may have a stronger impact on motivation than had previously been considered (Dörnyei, 2003b) and Robinsons' (2001b) suggestion that higher levels of motivation could lead to temporary increases in the level of resource pools currently used to meet the needs of a

pedagogical task.

1.2 Literature preview

This section previews the literature relevant to this thesis. Section 1.2.1 addresses cognitive task complexity and predictions by the Cognition Hypothesis (Robinson, 2001a, 2001b, 2005, 2007b) and the Limited Attentional Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001). Next, section 1.2.2 previews research utilising cognitive task complexity in both the oral and written modalities. The oral modality is included in this research because it is the bigger group of research and it informs much of the current understanding of the effects of cognitive task complexity on output. Additionally, comparisons between the oral and written modality may help to better understand the effects of mode in this thesis. Subsequently, section 1.2.3 addresses research dealing with planning time in both modalities. Finally, section 1.2.4 mentions post-task self-editing, and section 1.2.5 previews the relevant literature on task motivation.

1.2.1 Complexity

The effects of increased cognitive task complexity (elements modified in task design that make it harder to perform) on language development are often interpreted using two theories that hypothesize different outcomes. The Cognition Hypothesis (Robinson, 2001a, 2001b, 2005, 2007b) posits that increases in certain types of cognitive task complexity should lead to increases in complex language output; whereas, the Limited Attentional Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001) suggests that excessive increases in complex input will likely lead to degradations in complex output. Degradations in complex output can be viewed as inhibiting development as the use of increasingly

5

simple language constructions may be at odds with the notion that increasingly mature language users display increasingly complex elements in their production.

While there is a growing body of research that addresses cognitive task complexity and the predictions of both the Cognition Hypothesis and the Limited Attentional Model, most of the research has addressed the oral modality, not the written modality. This imbalance can be considered problematic considering that students also need to develop complex writing skills and that the results from research in the oral modality may not be directly applicable to writing.

When working on the development of complex writing skills, the difference in effect for modality may be an important issue. Some particular characteristics of writing have been suggested as reasons why the written modality is different from the oral mode. These characteristics include increased planning time and the recursive nature of writing's effect on processing, reduced mental stress between conceptualizing and encoding ideas, improved monitoring of output, and more selective control over cognitive elements of the writing process (Ellis & Yuan, 2004; Grabe & Kaplan, 1996; Grabowski, 2007; Kormos & Trebits, 2012).

As a result, the findings from the oral modality studies (Gilabert, 2007a, 2007b; Gilabert et al., 2009; Iwashita, McNamara, & Elder, 2001; Michel, Kuiken, & Vedder, 2007; Robinson, 1995b; Robinson, 2007b) cannot be assumed to be directly applicable to writing, especially when it comes to a dimension like complexity, which might be viewed as more suited to the conditions of writing. This is because the written modality is believed to have

different processing requirements (see above) that may be more suited to the resource heavy requirements of complex language production

Because of the issues mentioned above, the literature on the effects of complexity on both the oral and written modalities are reviewed. To determine if modality has had any mediating effects on cognitive task complexity and written complexity, it is advisable to view the research on the oral modality to provide some indication if the results differ.

1.2.2 Oral and written modalities

The current body of research exploring the effects of cognitive task complexity on complex and accurate oral output has mostly revealed positive outcomes for accuracy and some for lexical complexity, but no strong results for syntactic complex output (Gilabert, 2007a; Gilabert, 2007b; Gilabert et al., 2009; Iwashita et al., 2001; Michel et al., 2007; Robinson, 1995b; Robinson, 2007b). This has been viewed as partial support for the Cognition Hypothesis, which claims that increases in cognitive task complexity can have beneficial effects on the accuracy dimension of language development; however, there is little evidence to support the notion of dual increases in complexity and accuracy. Of interest to this thesis is the observation that the oral modality shows limited support for cognitive task complexity on complex output.

For the purposes of this thesis, studies in the written modality (meaning research where the analysed output was written) can be divided into two groups: one with variables different to this thesis and one with similar variables. The group with different variables is small. There are only 3 studies of which only one had positive effects for syntactic and lexical complexity (Ishikawa, 2006), and 2 for lexical complexity (Kormos, 2011; Ong & Zhang,

2010). The second group, which has similar variables (number of elements and reasoning demands) to this thesis, is also small, with contradictory findings for increases in lexical sophistication, and positive findings for type-token ratios that have not been corrected for text length (Kuiken & Vedder, 2007; Sercu et al., 2006). Other studies (Kuiken, Mos, & Vedder, 2005; Kuiken & Vedder, 2008) showed no positive results for lexical complexity. However, none of the studies that used similar variables to this thesis reported any effect for cognitive task complexity on syntactic complexity.

Further to the research on the oral and written modalities, there is a small body of work that focused specifically on comparing modalities (Grantfeldt, 2008; Kormos & Trebbits, 2012; Kuiken & Vedder, 2011); however, no clear evidence was found supporting modality as a noticeably effective element in tasks where cognitive task complexity was utilised in the task design.

As a result, in terms of syntactic complexity, the findings above provided no clear support for the Cognition Hypothesis (Robinson, 2001a, 2001b, 2005, 2007b); however, the lexical complexity results could be viewed as offering partial support. Additionally, the body of research addressing the effects of cognitive task complexity on complex and accurate output in both modalities shows little difference in the results that can be attributed to mode. However, it should be noted that an imbalance in the two bodies of work makes it difficult to make any strong judgments about the value of modality. To date, the body of work on the oral modality is larger than the body of work on the written modality.

To conclude, the body of work on complexity and different modalities has shown that, contrary to suggestions that writing may be a modality better suited to obtaining positive outcomes for complex output, there is no strong evidence supporting the benefits of complexity and modality. Additionally, research on writing, with a focus on cognitive task complexity as the independent variable, is a small group in which a variety of dependent variables have been used. For the small group utilizing number of elements and reasoning demands as the independent variables, the results for syntactic complexity appear to be non-significant. Subsequently, the non-significant results for syntactic complexity do not appear to support the predictions made by either The Cognition Hypothesis or the Limited Attentional Capacity Model.

1.2.3 Planning time in the oral and written modalities

As with the research previewed above, even though the focus of this thesis is the written modality, including the literature on research in the oral modality helps to provide a clearer picture of whether or not the written modality is a contributing factor in written complexity research. As such, the literature review on planning time and writing includes work in both the written and oral modalities.

Pre-task planning time can be *guided* or *unguided.* Guided planning time requires that the participants receive direction regarding how or what to plan for an upcoming task, and unguided planning time is when no help in preparing for an upcoming task is provided (Foster & Skehan, 1996).

Research utilizing various amounts of unguided pre-task planning time has revealed positive effects on complex output; however, modality does not seem to have had any

noticeable effect, with the written modality showing less favourable effects (possibly due to a lesser number of research projects in the field). As a result, it is difficult to predict whether there will be any positive effects for cognitive task complexity and unguided pre-task planning time based on mode.

For the pre-task planning time studies in the oral modality, which utilise various amounts of planning time, most of the studies (Foster, 1996; Gilabert, 2007b; Levkina & Gilabert, 2012; Mochizuki & Ortega 2008; Ortega, 1999; Sangarun, 2005; Skehan & Foster, 1996; Skehan & Foster, 1997; Skehan & Foster, 2005; Tavokoli & Skehan, 2005; Wendel, 1997; Yuan & Ellis, 2003) seemed to show positive results for complex output, with some positive results for both lexical and syntactic complexity across a variety of different types of complexity measures.

Half the studies that utilised pre-task planning on writing showed positive effects for planning on complexity (Ellis & Yuan, 2004; Ishikawa, 2006) and half did not (Ong & Zhang, 2010; Rahimpour & Safarie, 2011). If the results are reduced even further to account for variables similar to this study (meaning those in the written modality that used resource–directing cognitive task complexity (as specified by Robinson) and unguided pre-task planning time), the number of results is even smaller.

Ellis and Yuan (2004), and Ishikawa (2006) both used unguided pre-task planning time in conjunction with different types of cognitive task complexity in the task input. Both of these studies have produced results with favourable outcomes for complex output. Noticeably, both sets of findings revealed favourable outcomes for syntactic complexity. This is particularly important given that syntactic complexity has not featured much in

other results, with few showing anything positive. However, the pool of positive findings for syntactic complexity is small, making generalization difficult.

In conclusion, pre-task planning time appears to be one factor promoting positive outcomes for complex output in the oral modality; however, there are contradictory findings in the written modality. This may partially be the result of a smaller body of research in the field of writing, which skews results in favour of the oral modality. Furthermore, few pre-task planning and writing studies specifically address increases in cognitive task complexity in conjunction with pre-task planning time, and fewer still vary the amount of cognitive task complexity in the task input. Subsequently, strong predictions based on past research and the specific variables used in this research are difficult.

If the small amount of studies and some of the differences in detail are discounted, a cautious optimism might be taken from these two studies in terms of positive outcomes for the effects of unguided pre-task planning time on written complexity.

1.2.4 Post-task self-editing time

Post-task self-editing has been included as an additional variable in this thesis because editing is an important component in the recursive process of writing and one where the effects of complexity do not appear to have been researched. The body of work on post-task self-editing time and cognitive task complexity is mostly non-existent; as a result, there has been no real context for this research. However, observations made about post-task activities (Ellis, 2003; Skehan & Foster, 1997) do not suggest any positive outcomes for complexity during the editing process.

1.2.5 Task Motivation

There are always a multiplicity of factors affecting the production and learning of a language, as such it is helpful to look beyond the causal relationships between cognitive task complexity, pre-task planning time, and post-task editing time towards other elements that may be affecting the production of complex output.

One such factor are the effects resulting from the relationship between learner attitudes and performance, which differs from other variables addressed in this thesis by being related to diffiulty. Difficulty, as described by Robinson (2005, 2007a, 2007b), can be viewed as an individual difference factor. Individual differences is the study of the unique characteristics a learner has which alter and personalize an individual's language learning process (Dörnyei, 2009).

Ellis (2008) claims that within the body of research on individual differences, two main phychological factors appear to have the strongest effect on language learning. These are aptitude and motivation. It is motivation, or more specifically a subset of motivation called *task motivation*, that is the focus of this part of the thesis.

Motivation, broadly described as relating to a learner's choice of an action and the persistence and effort invested in that action (Manolopoulou-Sergi, 2004), can be viewed as having two causal relationships to learning. In one sense it is viewed as being a strong support for learning, in another sense it is viewed as being the result of learning (Ellis, 2008). Causally, task motivation can broadly be viewed as motivation resulting from the learning experience, specifically it refers to motivational effects resulting from the students'

reactions to the characteristics of a task. This type of motivation has required a more detailed type of analytical approach than previous analyses of motivation.

Task motivation requires a situated micro-approach, which focuses specifically on the effects learning situations have on a learner's motivation and how that in turn affects performance. A situated micro-approach should provide a more detailed analysis than the more traditional macro-approach, which focused on motivation across learning communities (Dörnyei, 2003b, 2010).

Within this situated approach, three effective components have been identified: the teacher-specific motivational components, the group-specific motivational components, and course-specific motivational components Dörnyei (1994). This thesis focuses on the course-specific motivational components dimension as it encompasses students' attitudes towards tasks in the classroom.

Of the four factors (interest, relevance, expectancy, and satisfaction) claimed to be part of the make up course specific motivational components (Crookes & Schimdt, 1991; Dörnyei, 1994; Keller, 1983), relevance (a tasks' perceived relevance to students' goals needs and values) and expectancy (including perceptions of task difficulty and the effort required for completion) are dimensions that feature in the Likert scale questionnaires that are used for data collecting in this section.

Dörnyei (2002, 2005, 2009, 2010) has devised a model that seeks to explain the processes at work when student's motivation is effected by situational classroom elements such as expectancy and relevance. This model is called the motivational task processing system,

which seeks to clarify student processing during task motivation. Dörnyei (2002, 2009, 2010) claims there are three interacting processes (task execution, task appraisal, action control) that will effect motivation and thus the quality and quantity of learner output.

In addition to Dörnyeis' (2009) claim that task motivation may effect the quality and quantity of output, Dörnyei and Kormos (2000) stated that the negative effects of task motivation could affect output variables such as accuracy, complexity and fluency, something that the current thesis will explore by checking for any correlations between positive or negative attitudes of expectancy and relevance with positive or negative syntactic output.

Finally, when considering both Dörnyeis' hypothesis on the effects of task characteristics on motivation (and thus performance), and Robinson's (2001b) speculation that motivation affects attentional resources, it might be expected that student perceptions of task characteristics are especially effective when those tasks are intrinsically attention demanding. As a result, positive or negative task motivation on cognitively complex tasks might result in noticeable effects on syntactic output, which is also resource demanding to produce. To date, there does not appear to be any studies that have addressed the issues of task motivation (as expectancy and relevance) on the production of subordinate clauses.

1.3 Gaps in the research

A review of the current literature has produced the following issues regarding complexity and writing that will be addressed in this thesis.

*A. The body of research on cognitive task complexity on written complexity is modest and within the small group there are many variations in independent variables***.**

This thesis focuses on the cognitive task complexity dimensions number of elements and reasoning demands (see page 2), so as to expand on the small amount of research on these variables. In doing so, this research hopes to add to the small body of work on cognitive task complexity in the written modality.

B. The neutral effect for mode may be due to the small amount of research.

By adding to the smaller group of research on complexity and writing, the present study expects to contribute towards balancing the body of research that is skewed towards oral production and help to give a clearer picture of the effect of mode on complex production.

C. The results for cognitive task complexity may be obscured by the inability to accurately measure cognitive task complexity in task input.

Because there is no specific measure for accurately increasing or decreasing cognitive task complexity between complex tasks, none of the studies on complexity can rule out the possibility that this may obscure the effects of manipulating complexity. This thesis suggests that past non-significant results for cognitive task complexity on syntactic complexity, in studies using similar variables to this thesis (Kuiken & Vedder, 2008, 2012), may have resulted, in part, from the inability to accurately measure and apply cognitive task complexity between complex tasks. Too much complexity between complex tasks may overload students and appear to produce results similar to too little complexity. Thus, a baseline non-complex task might be required to more accurately track the effects of increases in cognitive task complexity.

D. Studies in the written modality using the same variables as this study (reasoning demands and number of items) have shown no significant increases or decrease in syntactic complexity when subordinate clauses are measured as one group.

This thesis proposes that non-significant results for syntactic complexity in past research may be partly the result of current measures of syntactic complexity, which are not sensitive enough because they do not account for the differential development of individual dependent clauses. Clustering dependent clauses into one group when measuring the mean length of dependent clauses, as a measure of complexity, may miss something about the developmental continuum on which dependent clauses are learned. As a result, dependent clauses are measured as both one group and as separate units in the present study to explore whether the two measures produce different findings.

E. Pre-task planning time could be a mitigating factor in the effect of cognitive task complexity and written complexity, especially syntactic complexity.

This thesis also focuses on adding to the small body of research utilizing unguided pre-task planning time and increases in cognitive task complexity in the written modality. By utilizing the same independent variables (number of elements and reasoning demands) in each aspect of this thesis, a small amount of consistency can be created by using the same types of cognitively complex independent variables while adding the additional variable of unguided pre-task planning time. Furthermore, the baseline non-complex task and more sensitive measures of syntactic complexity mentioned earlier are added to pre-task planning time in order to address the problems mentioned above.

F. The possible effects of post-task editing time on complex output have yet to be explored.

Self-editing time is an issue that does not seem to have been addressed in any studies that

focus on the effects of cognitive task complexity and complex output. As a result, this

thesis has added post-task self-editing time, with the aim of discovering any effect for

cognitive task complexity and self-editing time on syntactic and lexical complexity. The

same issues regarding baseline tasks, and measures of subordination are also accounted for

in this part of the thesis.

G. There are no studies on task motivation's effect on subordination.

Whereas the prior sections approached complexity in terms of a causal relationship between

task input and complex output, this section focuses on correlations between attitudes

towards the tasks and syntactic output. To date, there does not appear to have been any

studies that focus specifically on correlating student's attitude towards complex tasks, and

their complex written performance, specifically the production of dependent clauses. As

such, this study takes a situated micro approach (Dörnyei, 2003b), which as a more fine-

tuned method of studying motivation, concentrating specifically on how classroom

elements effect performance. In this case, those class elements are the students attitudes

towards the complex pedagogical tasks (task motivation) performed during data collection.

1.4 Methodological approach

This study is a within-subject experimental design that uses both an experimental and

associational quantitative approach. These approaches are considered the most suitable for

the aims of this thesis in answering the research questions.

This thesis takes the position that the learning or acquisition of a language is not a random process but one that is subject to rules, thus it is possible to investigate it using scientific methods designed to ascertain any types of causal relationships between variables. This type of orientation is called a positivist/normative approach and is typically one that utilizes quantitative methods of data collection and analyses (Ellis & Barkhuizen, 2005).

The current thesis takes a quantitative methodological approach; however, it is one that is divided along the lines suggested by Mackey and Gass (2005). Mackey and Gass propose that quantitative approaches can be viewed as experimental or associational. Experimental quantitative research involves the manipulation of independent variables to test for effects on dependent variables, thus determining if there is cause and effect. This is consistent with the aims of this research, which is to find the effects of cognitive task complexity, pre-task planning time, and post-task editing time on written complexity.

Associational quantitative research seeks to determine the existence of relationships between variables without making direct claims about causality. This is also consistent with the aims of this thesis, which seek to establish any strength of association between students' attitudes towards complex tasks and written complexity.

1. 5 Outline, content, and structure

This thesis consists of 6 chapters. Chapter 1 is the introduction, which provides an overview of the thesis. Chapter two contains the literature that reviews the relevant research informing the research questions. Complexity as an independent and dependent variable is examined along with explanations of the Cognition Hypothesis (Robinson, 2001a, 2001b, 2005, 2007b) and The Limited Attentional Model (Skehan, 1998, 2003; Skehan & Foster,

1999, 2001), two popular theories used to explain the effects of complexity. Literature pertaining to both the oral and written modalities is reviewed, as modality is a potential mitigating factor in complex output. Both oral and written modalities are also assessed in the literature on planning time for the same reason. Next, task motivation, another potential effective factor is reviewed, and finally, the gaps in the research that inspired the research questions are identified.

Chapter three illustrates the methodological approach underlying this research as well as the methods used to collect and analyse the data. A positivist/normative approach is considered appropriate, as this thesis focuses on determining whether causality exists between complex input and complex output, and determining the strength of association between participant attitudes and written complexity. The research questions are presented followed by explanations of why the data collecting instruments, letter-writing tasks, and Likert scale questionnaires are considered the appropriate instruments for this research. Additionally, explanations are given for the types of statistical analyses used on each group of data.

Chapter 4 presents the findings from the statistical analyses of the data from the letter writing tasks and the Likert scale questionnaires. The effects of cognitive task complexity on lexical and syntactic complexity are addressed first. Syntactic and lexical complexity are analysed both within groups and between groups. First is the within groups analyses, which includes the standard measure of mean length of T-units used by previous research and a non-standard measure of mean length of T-units in which dependent clauses are separated. Syntactic and lexical complexity are analysed separately under three conditions, no planning time, 10 minutes pre-task planning time, and 10 minutes post-task editing time.

Secondly, the between groups analysis examines syntactic and lexical complexity between the three separate conditions mentioned above. Finally, the analyses of the Likert scale questionnaires are presented, focusing on correlations that show any significant strength of association between task motivation and complex syntactic output.

Chapter 5 is the discussion of results section in which interpretations of the statistical analysis are made in relation to the research questions and previous relevant literature. The implications for these findings are presented, as are the limitations of this research. Finally, suggestions for future research are suggested.

Chapter 6 is the conclusion. In this chapter, the findings contribution to theory, research, and pedagogy, is discussed. Additionally, the limitations of this thesis are mentioned, as are suggestions for areas of future research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview

The main focus of this research is the effect of cognitive task complexity on the lexical and syntactic complexity of written output. Centrally, this research looks to test whether the predictions for cognitive task complexity, made by Robinson (2001a, 2001b, 2005, 2007a), in the Cognition Hypothesis are confirmed in the written modality, which is a mode that has received less attention than the oral modality in the field of cognitive task complexity. Section 2.2.2 addresses cognitive task complexity, starting with a specification of the types of cognitive task complexity used in this research. In the following section (2.2.3), cognitive task complexity's role in traditional syllabus design is discussed, starting with an explanation of traditional syllabus design problems and the way in which Robinson suggests that cognitive task complexity might be used to circumvent problems, such as the need to consider individual learners needs, and the problem of sequencing and grading syllabi based only on chunks of language.

The next section (2.2.4) provides a description of the Cognition Hypothesis in which the theory supporting Robinsons' predictions is explained along with the Limited Attentional Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001), a model that predicts some different outcomes for the effects of cognitive task complexity. Finally, Robinsons' distinctions between *resource-directing* and *resource-dispersing* elements are clarified from section 2.2.4.3 to 2.2.4.5. This is important as cognitive task complexity, which can be viewed as a resource-directing element, is predicted to generate positive output as a

result of increases in complexity; whereas, pre-task planning time (another variable used in this thesis), which can be viewed as a resource-dispersing element, is predicted to generate negative effects on output as a result of increases in complexity.

In sections 2.2.5 to 2.2.5.5, past research in the field of cognitive task complexity in both the oral and written modalities is reviewed. It has been suggested that the nature of the written modality incorporates more planning time (Kormos & Trebits, 2012) and selective control over different parts of the cognitive processes (Grabowski, 2007) involved with production than is usually the case with the oral modality. As a result, it is one of the proposals of this research that modality may be an important factor when manipulating cognitively complex variables in search of favourable complex outputs. As most of the research on cognitive task complexity has been performed in the oral modality, these results are reviewed first as they have, for the most part, informed the current understanding of the effects of cognitive task complexity on output. The research pertaining to cognitive task complexity and the oral modality is reviewed with a focus on whether or not the results favour Robinson's predictions. Following this, the same is done with studies on the written modality. Subsequently, the written modality studies, which utilise the same dependent variables as this study, are reviewed separately to ascertain if there are any noticeable differences that might be due to variable type.

Sections 2.2.5.6 to 2.2.5.8 clarify what is meant by complex output, specifically as it relates to syntactic complexity and the type of measures used in this and other studies. It is suggested that current measures of dependent clauses used in research on cognitive task complexity and writing may not be providing a clear picture of the effects of cognitive task

complexity on the subordination skills of intermediate writers. A more detailed type of analysis is suggested.

Sections 2.3 to 2.3.6 review the additional variable of *planning time*, which is added to increases in cognitive task complexity. Once again, most of the work is done in the oral modality; as a result, the oral mode is reviewed and considered in light of Robinson's predictions. Following that, the same is done for the written modality and planning time, after which a comparison is made regarding planning time and the different modes. Finally, studies that have focused specifically on cognitive task complexity and pre-task planning time are reviewed.

Section 2.4 focuses on the final additional variable, which has been added to the investigating of cognitive task complexity on written complexity. The final variable is *post-task editing time*, which does not appear to have been previously researched in conjunction with cognitive task complexity and complex output. Given that the combination of cognitive task complexity and self-editing on written complexity has not been addressed, and that self-editing is a regular part of the writing process for learners wishing to produce quality work, this appears to be a variable that requires some consideration.

In section 2.5, the literature focuses on attitudes towards tasks and task motivation. Task motivation is the study of the effects of classroom elements (such as pedagogic tasks, on a participant's motivation) and how they subsequently affects a participant's output. The current thesis investigates how the motivational dimensions of relevance (a tasks' perceived relevance to students' goals, needs, and values) and expectancy (including perceptions of task difficulty and the effort required for completion) are correlated with the production of

dependent clauses. A review of Dörnyei's (2002, 2005, 2009, 2010) motivational task processing system is provided as a potential explanation for the processes involved in task motivation, and an explanation of how task motivation may relate to Robinsons' predictions for the effects of motivation.

2.2 Complexity

Before reviewing the literature on complexity, it should be noted that complexity is often researched in conjunction with accuracy; however, the decision was made to focus solely on complexity in this research. This decision was made for two main reasons. Firstly, the detailed investigation of complexity, in conjunction with the various supplementary issues addressed in this research, was considered a large enough body of work in itself without needing to add accuracy. Secondly, it was the position of this research that not every investigation into one of the three variables (accuracy, complexity, and fluency) automatically requires the addition of the other variables for the research to be informative and of value to the body of literature.

The term *complexity* informs much of this research, referring to the central independent variable (cognitive task complexity) common to this thesis' research questions; and also referring to the types of measures analysed in the written output (syntactic and lexical complexity). In the following sections, the different uses of the term complexity are reviewed, starting with cognitive task complexity as an independent variable. This is followed by cognitive task complexity as a central element in the design of pedagogical tasks as proposed by Robinsons' (2001a, 2005, 2007a, 2007b) Cognition Hypothesis. Following this, cognitive task complexity's' effect on both the written and oral modalities is assessed. Finally, complexity, as a measurement of written complex output, is discussed

with a specific focus on syntactic complexity addressing why subordinate output is considered complex and why there is a need to measure it.

2.2.1 Overview

The first dimension of complexity reviewed is that of cognitive task complexity, the central independent variable of this research. Cognitive task complexity can be simply viewed as one dimension of task complexity, which has been described as the degree to which any task is innately difficult or easy (Ellis, 2003).

In a summary of various accounts of criteria suggested for task grading, Ellis (2003) listed the following four elements believed to affect task complexity: (a) Input, information that is supplied as part of the task and can be manipulated by the task designer. (b) Conditions, the interactional requirements of a task. (c) Process, the type of cognitive processes needed to perform a task, such as deduction, inference, or calculation (Prabhu, 1987). (d) Outcomes, whether or not the outcome of a task is simple and general, or complex and precise.

Of the four dimensions of task complexity (input, task conditions, task process, and task outcomes) shown to affect the innate difficulty of a task, it is the *input dimension* (information that is supplied by the task and manipulated by the designer) that is central to this study. Of specific interest is the input types that are posited to affect the *process dimension*, which are the cognitive demands required for processing (e.g., deduction, inference, or calculation). This type of input has been termed cognitive task complexity

Importantly, cognitive task complexity has been suggested as a central component for the design and sequencing of tasks in a task-based syllabus (Robinson, 2001a, 2005, 2007a,

2007b). Robinson has suggested that variations of certain types of cognitive task complexity (in the design of pedagogic tasks) will have particular effects on attentional resources (the amount of attention that can be applied to cognitive processes like deduction, inference, or calculation). Subsequently, Robinson believes that directing these attentional resources may have positive effects on learners' language production and the ability to learn the target language.

In conjunction with the additional variables pre-task planning time and self-editing time, this study mainly investigates cognitive task complexity as it affects lexical and syntactic written complexity, and the way these effects relate to predictions by the Cognition Hypothesis (Robinson, 2001a, 2005, 2007a, 2007b), which predicts positive outcomes for variations in certain types of cognitive task complexity on complex output.

Additionally, this thesis considers whether modality is a moderating factor between cognitively complex input and the production of complex output. A review of research in the oral mode, in conjunction with research on task complexity and the written modality, may provide some insight and context as to whether mode is a contributing factor for the results of this study.

Finally, this section concludes with a review of what is meant by complexity when this term is applied to complex syntactic output of the type utilised in this research. This research posits that one of the reasons that previous studies similar to this one (Kuiken & Vedder, 2008, 2012) have shown no significant results for cognitive task complexity on syntax may be the insensitive nature of measuring all subordinate clauses together.

Separating dependent clauses during analyses may provide a clearer view of the effects of cognitive task complexity on syntactic complexity.

2.2.2 Cognitive task complexity

Complexity in SLA (second language acquisition) can be viewed as relating to two dimensions: the elements of a task, which is called cognitive task complexity; and the elements of performance and proficiency, which is L2 complexity (Housen & Kuiken, 2009). This study spotlights the causal relationship between both dimensions of complexity, specifically the effects of manipulating task complexity (in this case, cognitive task complexity as described in Table 1) on particular elements of linguistic complexity (defined as measures of written performance that are believed to represent complex language output: see section 3.5).

Broadly, cognitive task complexity (as an elements of task design) can result from the interplay of two elements that can be manipulated in the design of pedagogical tasks (Ellis, 2003). These elements are *types of information* and *amounts of information* (Brown, Anderson, Shilcock, & Yule, 1984), which are posited to place varying demands on learners' cognitive resources (see Table 1).

Table 1: Types of cognitive task complexity

Cognitive task complexity	Examples
1. Types of information	Examples of types of information
Static information does not change during a task. This is considered the easiest element.	Describe the items in a picture.
Dynamic information changes during the task. This is considered to be more demanding than a static task.	Describe what is happening in a series of pictures that show a sequence of events.
Abstract information requires the participant to have an opinion or defend their position. This is considered the most difficult	Choose an item (maybe one restaurant) and explain why this choice is a good one.
2. Amounts of information	Examples of amounts of information
The number of elements that are used in a task	Describing the actions of one person would be easier than describing the actions of five people.
The relationship between elements refers to the relationship between the varying *number of elements* and the *static*, *dynamic* or *abstract information* types	An example of *number of elements* and *abstract information* would be the following: Choose an item (maybe one restaurant from two) and explain why this restaurant is better than the one not chosen. Subsequently, Choose an item (maybe one restaurant from four) and explain why it is better than the three not chosen.

An important dimension of this research is testing the predictions of the Cognition Hypothesis, which posits a central place for cognitive task complexity in the written modality. Robinson incorporated the dimensions of cognitive task complexity mentioned above into the Cognition Hypothesis, which is the theoretical underpinning of a detailed taxonomy called the Triadic Componential Framework. In the Triadic Componential

Framework the terms complexity and difficulty (which have often been conflated) are clarified (see below). After separating and clarifying these dimensions, Robinson goes on to describe a clear and detailed foundation on which task sequencing decisions can be made.

In the Triadic Componential Framework, Robinson (2001a, 2001b, 2005, 2007a) proposed three dimensions: (a) Task complexity, which is the process where inherent task characteristics have a cognitive effect on second language learners. The cognitive effects are the demands put on information processing, which include elements such as attention, memory, and reasoning (Robinson 2001a). (b) Task conditions, which are the interactive demands of a task, e.g., whether information goes only one way from one person to another, or whether information is exchanged between people. Robinson described this as the participation requirements of any given task. (c) Task difficulty, which refers to factors that relate to individual learners. These factors include elements such as aptitude, confidence, motivation, and intelligence. These are the resources that an individual brings to a task (Ellis 2003).

In this thesis, cognitive task complexity is a combination of *abstract information* (henceforth referred to as *reasoning demands*) and *numbers of elements,* which are situated in Robinsons' taxonomy as dimensions of task complexity. It is a crucial part of this research to discover whether cognitive task complexity induces output consistent with theories posited in the Cognition Hypothesis (Robinson, 2001a, 2007a), thus adding evidence supporting or negating the use of cognitive task complexity as an effective language learning tool, and as an effective element of task sequencing as proposed in Robinsons' Triadic Componential Framework (2001a, 2001b, 2005, 2007a).

2.2.3 Cognitive task complexity and syllabus design

Cognitive task complexity has been suggested by Robinson as a means by which certain problems associated with traditional syllabus may be overcome. Early theories informing changes in the field of language teaching, and the resulting changes to syllabus design, can be linked to at least two central factors. These are the idea that language structures needed to be taught in conjunction with a focus on developing meaning expression (Widdowson, 1978); and the notion that the theories of acquisition underpinning traditional linguistic syllabi were not supported by research on language learning (Long & Crookes 1992).

Traditional synthetic syllabi were identified as being incompatible with the constantly evolving, research-based, approaches to language teaching that have been informing changes to classroom pedagogy since the 1970s. Importantly, Corder (1967) noted that learners appeared to be resistant to the types of external sequencing methods used in traditional syllabi.

Traditional syllabi were found to have a number of problems. For example, given that the language samples being taught were selected to conform to the syllabus designer's linguistic specifications, these samples did not represent how people spoke (Long & Crookes, 1992). Subsequent research showed that language items were part of "complex mappings of form function relationships" (Long & Crookes 1992, p 31); a view inconsistent with the traditional practice of teaching decontextualized language samples (language samples that were isolated and taught without any connection to the context in which they would normally exist as part of a functioning and usable language system). Secondly, traditional syllabi viewed groups of learners collectively, thus failing to account for individual differences in developmental rates (Ellis, 2003). This point touches on a

central problem with traditional syllabi, which is the notion that individual development was not a consideration given that syllabi were based on external considerations.

Broadly, task-based approaches attempt to avoid the types of problems mentioned above, and are a way in which the individual developmental needs of students might be addressed. This might be achieved by providing real-world tasks that mirror real world language requirements while simultaneously allowing participants to use their own resources (e.g., current language levels and aptitude) to complete pedagogical language learning tasks.

To facilitate the development of individual needs, researchers proposed the use of task-based approaches that utilised *psycholinguistic-based task elements*. Psycholinguistic-based task elements are properties of a task that are expected to induce participants to employ the types of language use and cognitive processing that have an effect on how language learners process/acquire language (Ellis, 2000; Long & Crooked, 1987). It would be essential to be able to sequence any tasks that utilise these elements in a way that facilitates the best situation for learning. Additionally, Candlin (1987) claimed that tasks need to be selected and sequenced in a principled way with Skehan (1996) adding that any principled criterion for task sequencing would be a corollary for the allocation of attentional demands (an issue that Robinson claims to address in the Cognition Hypothesis).

Any principled criterion for task sequencing must account for form/meaning issues. Form/meaning issues refer to the problems associated with trying to simultaneously teach language forms (words and grammar), language forms relationships to other language forms, and the meanings and usage of those language forms in practical real world conditions. Long (1985, 1991) proposed a distinction between focus-on-formS and a focus-

on-form. Focus-on-formS is a focus on the type of isolated linguistic forms found in traditional syllabi; and focus-on-form is described as the process where instruction engages a learners' attention to language structure while the primary focus of the learner is on the meaning of the content (Ellis 2003).

Focus-on-form is an important element of task-based approaches because learners tend to prioritize meaning (VanPatten, 1989, 1990) when performing purely meaning driven tasks, something that is unlikely to promote the necessary development of all three dimensions (accuracy, complexity, and fluency) of interlanguage.

Interlanguage (Selinker, 1972) is the developing language knowledge of a second language learner, and it is believed to have three dimensions (Skehan, 1998): fluency, being able to communicate in real time; accuracy, the ability to produce language that is consistent with intended language norms; and complexity, the using of elaborate and cutting edge structures (Ellis 2003). A task-based syllabus would have to sequence and grade tasks in a way that incorporated a focus-on-form, with the intention of developing all three dimensions of interlanguage.

Cognitive task complexity has been posited as a principled criterion with which to sequence and grade tasks while accounting for the developmental issues mentioned above. Robinson (2007a) argued that pedagogic tasks should be developed and sequenced for language students based on the increasing of cognitive task complexity, and that these increases should be done gradually until the complexity levels mirror those of real world demands. It is Robinson's contention that sequencing tasks based on cognitive task complexity could

address the issue of form/ meaning development and any attendant attentional allocation issues.

2.2.4 Competing theoretical approaches

When reviewing the results of research on task complexity and Robinsons' taxonomy of task elements, it is also important to account for the Limited Attentional Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001) also referred to as the Trade-off Hypothesis, which predicts potentially different outcomes when increasing cognitive task complexity.

In the literature, these two approaches are often mentioned together when comparing research results that speculate on the apportioning of attentional resources and their effect on output. This is an issue that arises during discussion on the effects of cognitive task complexity on the output factors accuracy, complexity, and fluency. This is especially the case with accuracy and complexity, which are viewed as more resource demanding than fluency.

The output aspects accuracy, complexity and fluency (Skehan, 1996, 1998) are theorized to utilise different language systems; for example, fluency employs readily available, formulaic chunks of language. Alternatively, complexity and accuracy require more processing resources when drawing on a rule based system, which is a system of rules upon which a person can generate correctly-formed or creative language (Ellis, 2008).

Importantly, though accuracy and complexity are associated with the use of the rule based system, accuracy and complexity can be viewed as separate and competing dimensions, with accuracy associated with control over existing forms and complexity associated with

restructuring and risk taking (Ellis, 2005). Moreover, Robinson and Skehan predict different outcomes for increases in cognitive task complexity on accuracy and complexity.

2.2.4.1 The Limited Attentional Model

A central issue dividing the Cognition Hypothesis and the Limited Attentional Capacity Model is the proposal by Robinson that the concurrent development of accuracy and complexity is possible when cognitive task complexity is increased in task input, a point disputed by Skehan.

Very broadly, the Limited Attentional Capacity Model (also referred to as The Trade-off Hypothesis) advocates that cognitively demanding tasks require trade-offs from limited attentional resources (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001). Building on the work of VanPatten (1990), the Limited Attentional Capacity Model proposes that when resource limits are reached, learners will focus on meaning, something that Skehan (1996) claimed they are naturally and unavoidably predisposed towards. Additionally, given that complexity and accuracy are viewed as separate and competing dimensions, both dimensions will be competing for limited attentional resources, with one dimension likely receiving less attention than the other.

Additionally, the Limited Attentional Capacity Model (Skehan & Foster, 2012) notes that trade-offs in performance will not always be apparent because other factors may influence the outcome; however, the need for trade-offs in performance can be viewed as the default position. Furthermore, Skehan suggests that increases in complexity and accuracy are unlikely to happen concurrently. Any dual increases in complex or accurate output as found in Foster and Skehan (1999), Skehan and Foster (1997), and Tavakoli and Skehan (2005)

are probably not the result of increases in complexity (as suggested by the Cognition Hypothesis) but instead could be the result of other mediating factors e.g., a situation where certain types of structures that lead to increases in accuracy might be working in conjunction with types of information manipulation that effect subordination (Skehan, 2009; Skehan & Foster, 2012).

This thesis only focuses on complex output, and therefore will not provide any results or discussion regarding the issue of the dual development of accuracy and complexity. However, Skehan's Limited Attentional Capacity Model is important to acknowledge, as it is closer to the more orthodox theories regarding limited attentional resources and their theorized effects on output under cognitive stress.

2.2.4.2 The Cognition Hypothesis

As stated earlier, Robinson's Cognition Hypothesis (2005, 2007a) posits that cognitive task complexity is a means by which tasks in a task-based syllabus can be sequenced, with increases in cognitive task complexity hypothesized as having positive outcomes for learner output. Contrary to prediction made by the Limited Attentional Capacity Model, Robinson also claims that increases in cognitive task complexity could lead to dual increases in accurate and complex output. In the following section Robinsons' (2001a, 2001b, 2005, 2007a) theoretical underpinnings for the Cognition Hypothesis are briefly reviewed.

Robinsons' assertion that increases in cognitive task complexity lead to increases in language production is grounded in theories about first language development and the notion that language has two modes. These are a simpler mode for expressing simple meaning and a more complex mode for expressing complex meaning.

Robinson suggests that increases in cognitive task complexity, leading to increases in accurate and complex language, mirrors aspects of first language development, specifically that conceptual development creates the conditions for first language development (Cromer, 1991) and that there are developmental parallels between adults learning L2 and children's development of L1 (Slobin, 1993*).* As a result, Robinson (2001a, 2005) has claimed that pedagogic tasks used in task-based syllabi should be sequenced using increases in cognitive task complexity and that this sequencing would be similar to the way that children meet increasingly complex demands when learning a first language. Thus, increases in conceptual development should lead to subsequent increases in the language needed to express those concepts by L2 learners.

Robinson hypothesises that there may be a pragmatic and a syntactic mode to language output (Givon, 1985, 2009), the uses of which are related to cognitive demands. Increasingly complex cognition may be coded and produced by the complex linguistic forms needed to express the increases in complexity. Robinson states that less cognitively demanding tasks require the less complex pragmatic mode (more focused on meaning expression), and more complex tasks would utilise the syntactic mode (which utilises more lexically and syntactically complex language forms).

One of the points upon which the Limited Attentional Resource Model and The Cognition Hypothesis differ is the nature of attention as it facilitates the learning process. Whereas Skehan claims that the default position is one where limited attentional resources lead to trade-offs during the production of accuracy and complexity, Robinson has claimed that the default position may be more flexible.

The notion that limited attentional resources might have a negative effect on the positive predictions for cognitive task complexity is accounted for by Robinsons' approach to the theory of multiple resource pools. The Cognition Hypothesis (Robinson, 2005, 2007a) is influenced by Wickens (1980, 1989) who wrote that there might be multiple pools for resource allocation that could allow for multiple attention-demanding activities to be carried out if they draw on different modalities. For example, Robertson and Gilabert (2007) suggested that driving a car and listening to a radio are possible as they utilise visual and auditory modes, which are different modalities. The notion of multiple resource pools has allowed Robinson to claim that the dual development of accuracy and complexity may be possible without the need for trade-offs.

In terms of complex output, which is the production dimension focused upon in this thesis, the Cognition Hypothesis predicts that increases in task complexity should result in increases in complex language as a way of expressing the increasingly complex demands of the tasks. However, Robinson (2005, 2007a) and Robinson and Gillabert (2007) state that it is important that increases in complexity are made along resource-directing dimensions, as opposed to resource-dispersing dimensions, both of which are discussed in the following sections.

2.2.4.3 Resource-directing and resource-dispersing dimensions

In Robinson's Triadic Componential Framework (2001a, 2001b, 2005, 2007a), the list of task design characteristics that are proposed to have an effect on task performance are further divided into two categories, which are believed to have different effects on learners when increased in complexity. These dimensions are *resource-directing* dimensions and *resource-dispersing* dimensions.

2.2.4.4 Resource-directing dimensions

Robinson categorizes cognitive task complexity as a resource-directing dimension. A resource-directing dimension is believed to make cognitive/conceptual demands on a learner as complexity is increased. It can be manipulated in task design so that more or less demanding language is needed to express the variations in complexity.

In addition, as a result of increasing the complexity of a resource-directing dimension, Robinson claims that learner attention is channelled in a way that facilitates noticing. Noticing is the conscious attention to language forms, which is considered a prerequisite of learning (Schmidt, 2001). Subsequently, the noticing, facilitated by increases in resource-directing complexity, is posited to lead to more accurate language output, as well as more complex language output. The increases in complex language output is viewed as resulting from a move from the less complex pragmatic mode (which is meaning focused) to the syntactic mode (which utilises lexically and syntactically more complex language).

2.2.4.5 Resource-dispersing dimensions

Robinson also identifies resource-dispersing elements. Contrary to the resource-directing dimensions, the resource-dispersing elements have an effect more in line with the negative effects for increased complexity predicted by Skehan and Foster (2001). When increased in complexity, resource-dispersing elements do not direct learners' attention to the specific language needed to meet the demands of a more complex task; but instead, attention is dispersed and unfocused, making the completion of any task more difficult. For example, planning time is considered a resource-dispersing element.

When a task is made more complex by removing planning time, the smaller amount of processing time might affect attentional focus, thus making it more difficult to access and utilise a learner's language resources (something that may already be difficult for learners of language who are not as fluent as native speakers). This will not promote more accurate or complex language, but instead encourage the learner to use a more pragmatic (meaning focused) language mode, focusing more on fluent language production. The Cognition Hypothesis can be viewed as requiring a balance of resource-directing and resource-dispersing in different measures to meet the developmental ends of achieving fluency, complexity, and accuracy.

In this thesis, the independent variables relate to the Triadic Componential Framework in the following way. The central independent variables used in 3 of the 4 research questions are the cognitive complexity elements reasoning demands and number of elements, which are viewed by the Triadic Componential Framework as resource-directing elements. Planning-time, which is an additional independent variable featuring alongside cognitive task complexity in one of the questions, is a resource-dispersing element (see above). When used together, Robinson claims that cognitive task complexity should be increased, in conjunction with the addition of planning time, to achieve positive results for complex output.

2.2.5 Cognitive task complexity in oral and written modalities

As noted in the previous sections, cognitive task complexity has been suggested as having a central place in the sequencing and grading of task based syllabi. The previous sections have reviewed Robinsons' prediction that cognitive task complexity might have positive outcomes for complex output. However, it is important to remember that cognitive task

complexity cannot be considered to be operating in isolation. One variable that may play a part in the relationship between cognitive task complexity and output (not withstanding other factors like task conditions and learners' individual differences, which can affect task completion) is the effect of mode of production (Kuiken & Vedder, 2011).

There are reasons to believe that mode might be a contributing factor if cognitive complexity were to have a positive effect on measures of complex output. One of the essential elements separating oral and written production is planning time (both pre-task and online) and the effect they have on processing. Grabe and Kaplan (1996) state that fewer time constraints are a central issue for writing processing, in conjunction with the recursive process of planning, encoding (transforming ideas into words and syntactic structures), and revising.

In a study about on-line planning, Yuan and Ellis (2003) state that there may be a reduced load on writers (compared to speakers) when it comes to apportioning attention between the conceptualizing and the linguist encoding of a message. Kormos and Trebits (2012) also make the point that pre-task planning time may be integrated into the writing process; as a result, writing students naturally have more time to plan what they will produce prior to writing. Additionally, Kormos and Trebits state that writers may also have more available attentional resources to monitor what they are producing during the encoding process (where thoughts are transformed into words and syntactic structures). In this case, writers would have more attentional resources to apply to this process than speakers would.

Finally, Grabowski (2007) notes that writing production is more self-determined than oral production, with the writer having more selective control over different parts of the

cognitive process (an attribute that can be seen as conducive to both the task-based theory that participants need to use their own resources to improve, and the increased demands made by cognitive task complexity). It is possible, given the previous assertions, that the proposed existence of planning time inherent in writing may be viewed as a type of complexity facilitating feature.

Considering that most of the work to date on cognitive task complexity has been on the oral modality, and that the mode of production may be an important factor affecting production, a review and comparison of the literature on both oral and written studies may shed light on the importance of modality and complexity. It is possible that given the differences in modalities mentioned above, results that support the Cognition Hypothesis, especially in terms of complex output, may be evident in the written modality.

In the following section, the research in the field of cognitive task complexity and the oral modality is reviewed and considered in relation to the claims that increasing cognitive task complexity is predicted to have positive effects on complex production. Following this, the written modality is reviewed, starting with work that utilises various types of cognitive task complexity as independent variables, then finishing with a review of the research that utilises independent variables similar to this study. Finally, the results from both the oral and written modality are considered in terms of how they appear to relate to predictions regarding issues of cognitive task complexity and the proposed effects of modality, with a specific focus on complex output.

2.2.5.1 Oral modality

As mentioned above, most work on manipulating cognitive task complexity has been in the area of oral production (both monologic and dialogic). Research has used a variety of independent and dependent variables while manipulating task complexity along the dimensions that Robinson refers to as resource-directing and resource-dispersing.

The combination of this range of variables has mostly yielded results in terms of accuracy, which could be viewed as partial support for Robinsons' approach; however, it does not appear to demonstrate much support for the notion that increases in cognitive task complexity have a simultaneously beneficial effect on accuracy and complexity, or that there is much beneficial effect for complex output alone, with syntactic complexity showing little in the way of positive results.

There are a number of studies that have utilised the cognitive task complexity dimension *here-and-now* in oral complexity research. Here-and-now refers to a less complex task condition where participants refer to situations happening now as opposed to a more complex condition *there-and-then*, where participants refer to action happening in the past. Here-and -now is considered to be a resource–directing dimensions by Robinson.

Robinson (1995) manipulated task complexity along here-and-now and contextual support dimensions. Support dimension refers to whether or not the participant was able to view the picture material used in the task during task completion. In the here-and-now task, the participants looked at the picture prompts while performing; in the there-and-then, the picture prompt was unavailable while the participant was speaking. The findings from his study showed that more complex narrative tasks (there-and-then) produced a non-

significant but strong trend for accurate language when accuracy was measured as target-like use of articles, but showed no results for increases in task complexity on complex syntactic output.

Task complexity was also manipulated along the here-and-now dimension by having students narrate stories from pictures in a study by Iwashita, McNamara, and Elder (2001). The findings from this study supported Robinson's (1995) study, with the more complex task eliciting more accurate language when measured as error free clauses in the total number of clauses. However, there were no significant results for complexity. This could be viewed as a partial validation of Robinson's Cognition Hypothesis because increases along the here-and-now resource-directing dimension resulted in increases in accurate output; however, it is referred to as only *partial* support because there are no concurrent positive results for complex output as predicted by the Cognition Hypothesis.

In further studies utilising here-and-now, Gilabert (2007a, 2007b) and Gilabert, Baron, and Llanes (2009) performed a number of studies analysing cognitive task complexity on oral output; however, complex output was only included in one of these studies. Gilabert (2007a) and Gilabert et al. (2009) both utilised cognitive task complexity along the here-and-now dimension and different task types, but neither of these studies addressed the effects of cognitive task complexity on complex output. In terms of how these studies related to Robinson's predictions, Gilabert (2007a) found no results to support the Cognition Hypothesis, though Gilabert et al. (2009) produced some support for the Cognition Hypothesis' assertion regarding cognitive complexity and interaction.

Of more relevance to this thesis is Gilabert (2007b), who did measure complex output in a study that manipulated here-and-now (resource-directing) and planning time (resource-dispersing) dimensions. The findings revealed no result for increased complexity on syntactic complexity. Additionally, less lexical complexity was discovered in the more complex conditions. The only partial support for the Cognition Hypothesis was in the significant findings for accuracy.

Michel et al., (2007) provided some support for the Cognition Hypothesis in a study that manipulated cognitive task complexity, but used a different dimension than the studies by Gilabert. Michel et al. (2007) utilised task conditions (both monologic and dialogic) and increases in cognitive task complexity along the number of elements dimension. Accuracy, complexity and fluency were all measured, with some positive effects found for increases in accuracy in the complex monologic tasks. In terms of complex output, the complex task had more diverse speech, meaning a significant effect on the percentage of lexical words; however, there was no effect for increases in number of elements on syntactic complexity.

Finally, Robinson (2007b) used neither the here-and-now or the number of elements dimension used in the studies mentioned above. Instead, the resource-directing dimension reasoning demands was used, and the output was measured using psychological state terms (the use of words associated with psychological states that are supposed to signify the development of syntactic complexity in children). Robinson reported that increases in complexity along the reasoning-demands dimension resulted in increases in complex speech.

The literature on task complexity and oral production has mostly shown findings with positive outcomes for accuracy, which can be viewed as partial support for the Cognition Hypothesis' claims that increases in task complexity can have beneficial form effects on the accuracy dimension of language development; however, there is little to support the notion of dual increases in complexity and accuracy.

The results for cognitive task complexity on oral complex output do not provide strong support for the Cognition Hypothesis, with one result favouring lexical complexity, but no strong results for increases in cognitive task complexity on syntactic complex output. Though Robinson's (2007b) study provides positive results for increases in cognitive task complexity on complex output, the complexity measures (psychological state terms) used by Robinson are, to date, not commonly used in studies of cognitive task complexity and complex output. As a result, the findings are difficult to compare with those other studies in the field.

Problematically, there are relatively few studies addressing the effects of cognitive task complexity on complex output. In addition, those few studies have utilised various types of cognitive complexity and various task conditions, making generalisations about the results difficult.

It is possible that one of the factors contributing to the lack of positive results for cognitive task complexity on complex output is that of modality. The lack of positive results for increases in syntactic complexity could be viewed as consistent with the notion that the processing requirements of oral production in conjunction with increases in cognitive task

complexity are not the optimal circumstances for the type of syntactic processing required for positive outcomes in syntactic complexity.

If stronger evidence supporting Robinson's (2001a, 2001b, 2005, 2007a, 2007b) theories about positive outcomes for cognitive task complexity on complex output were to be found, it seems that the processing conditions of the written modality may be more conducive to the claims of the Cognition Hypothesis.

2.2.5.2 Written modality

As with the studies in the oral modality, there are relatively few studies on cognitive task complexity and writing, and these studies utilise various types of cognitive task complexity as the independent variables in conjunction with various different measurements in the analysis of the results.

This section is separated into two parts. Section 2.2.5.3 reviews studies of cognitive task complexity and writing, which have utilised a variety of independent variables and a variety of language measures that are different to the ones used in this thesis. Following this, section 2.2.5.5 focuses specifically on studies that have utilised the same types of cognitive task complexity and similar output measures as those used in this thesis.

2.2.5.3 Written modality and various types of complexity

In a study on narrative tasks, Ishikawa (2006) manipulated task complexity along here-and-now dimensions. The findings indicated that increases in task complexity led to increases in accuracy, syntactic complexity and fluency. Cognitive task complexity was increased along the here-and-now dimension and measured as both syntactic and lexical complexity.

Ishikawa used S-nodes, which are the equivalent of finite or non-finite verb phrases, in the measurements (e.g., want to be, stand, walk, has been, did) Syntactic complexity was analysed as the following: S-nodes/T, which are the S-nodes per t-unit; C/T, which is the number of clauses per T-unit; S-nodes/C, which is the number of S-nodes per clause; and DC/C, which is the dependent clauses per clause.

Ishikawa's (2006) study could provide the clearest results in favour of the beneficial results of increases in cognitive task complexity on written output. The results from Ishikawas' study can be viewed as supporting Robinsons' assertion about the positive effects of increases in cognitive task complexity, as well as the idea that the processing requirements of writing may be more suited to cognitive task complexity and complex output. However, it should be noted that Ishikawas' study also includes pre-task planning (though it is not mentioned as one of the independent variables in the study) because five minutes preparation time was allowed prior to the commencement of the task. Skehan (2009) states that it is not known if Ishikawas' results would be applicable in a non-pre-task planning situation. As a result, Ishikawas' study, and its findings, might be considered more suitable to section 2.3, which reviews studies that include pre-task planning time.

Ong and Zhang (2010) analysed different levels of pre-task planning time, ideas and macro structures, and draft availability. Complexity was increased through the manipulation of planning time, the availability of writing assistance, and the availability of a first draft. As a result, the most complex conditions were the ones where there was no planning time, no writing assistance, and no access to the first draft. The results revealed significant effects for the removal of planning time on increases in fluency and lexical complexity, meaning increased complexity, actualized by the removing of planning, time led to marginally

higher lexical complexity. The removing of ideas and macro structures lead to significant increases in lexical complexity. Finally, no results were provided for the effects of planning time, ideas and macro structures, and draft availability on syntactic complexity.

An interesting point, which should be noted about Ong and Zhangs' research, is the influence of planning time. Counter to findings on the benefits of planning time on some measures of complexity, this study appears to show that the addition of planning time had a negative effect. This result would be counter to Robinsons' assertion that, as a resource-dispersing variable, the exclusion of planning time might lead to negative effects on complexity. Additionally, an inspection of this study reveals that there may have been a conflating of variables as the task instruction requests the participants to give an opinion and to provide specific reasons to support the answer. As a result, the cognitive task complexity dimension reasoning demands was an unaccounted for variable in this research that may have had an effect on the results.

Finally, Kormos (2011) manipulated complexity through the use of picture prompts, one of which gave narrative context while the other did not. Additionally, L1 users (native language users) and FL users (foreign language users) results were compared. Kormos utilised a large range of measures from which a significant effect was found for lexical sophistication (concreteness) and a significant difference in lexical variety between L1 and FL writers. However, no significant results were found for increases in complexity on syntactic complexity for second language learners.

Regarding the preceding studies, of the results for cognitive task complexity (using variables different than those used in this thesis) and writing, only Ishikawa (2006) had

positive effects for syntactic complexity. This result, in conjunction with the unmentioned inclusion of pre-task planning time, would be consistent with Robinsons' claim that the addition of planning time helps with directing resources to complex language production. However, the opposite seemed to be true in Ong and Zhangs' (2010) study where it was found that increasing resource-dispersing complexity, by removing planning time, had a marginally positive effect on lexical complexity.

Considering the small number of studies and the contradictory results, it is difficult to make any clear prediction based on these findings. These studies have used a variety of different independent variables; as a result, the differences in the findings for written complexity could have been caused by the use of different cognitive task complexity variables because various complexity variables may not elicit uniform effects on output. Furthermore, if the various effects of different complexity variables are combined with variations in other factors (e.g., individual differences and the inexact nature of applying levels of cognitive task complexity in the input), then the contradictions in the findings are not surprising.

What can be understood from these studies is that they have, for the most part, demonstrated little effect for cognitive task complexity on syntactic complexity. Additionally, mode does not seem to have had much effect given that both the oral and written groups of findings found a small effect for lexical complexity and (with the exception of Ishikawa) no effect for syntactic complexity.

2.2.5.4 Task complexity studies comparing modalities

In the sections above, no clear difference between the results of the written and oral modalities was found. In relation to the Cognition Hypothesis' predictions for increase in cognitive task complexity, the results have not been strong, with nearly no findings for syntactic complexity in either modality, and sporadic results for lexical complexity. To complete the review of cognitive task complexity, this section reviews those studies that have focused specifically on the issue of modality.

In a study on the effects of mode, Grantfeldts' (2008) findings revealed significant increases in lexical complexity in the written mode; conversely, there was more accuracy in the oral mode. However, this was not the case for grammatical complexity, which was analysed with a sub-clause ratio and the occurrence of advanced syntactic structures, as there was less grammatical complexity found in the writing than there was in the oral modality.

Kuiken and Vedder (2011) investigated cognitive task complexity and the effect of mode. Using both oral and written tasks that increased in reasoning demands and number of elements, the output was analysed for accuracy and complexity. The syntactic complexity measures were number of clauses per T-unit (a syntactic unit used to measure writing production) and AS unit (a syntactic unit used to measure oral production), and a dependent clause ratio. Lexical complexity was measured using a type-token ratio that corrects for text length. Accuracy was measured as total errors per T-unit (writing) and AS unit (oral), accounting for three degrees of errors.

Kuiken and Vedders' results showed that in both the written and oral modes, complexity influenced accuracy, with the more complex tasks having more accuracy. However, cognitive task complexity had no effect on lexical complexity in either the written or oral modality. Though there was no effect for cognitive task complexity on written syntactic complexity, increases in cognitive task complexity lead to less syntactic complexity (fewer dependent clauses) in the oral mode.

Interestingly, Kuiken and Vedders' (2011) results for complexity on dependent clauses in the oral modality may provide some evidence about the difference between modality and the level of cognitive task complexity applied to learners during a task. It is possible that the level of cognitive task complexity used in the tasks was enough to tax the participants during the oral production of dependent clauses, which can be viewed as being more taxing to process (see section 2.2.5.8). Conversely, the same amount of cognitive task complexity might not have been enough to negatively affect written output, given the different processing requirements of the written modality.

In a study of complexity and mode, Kormos and Trebits (2012) examined complexity across both oral and written modalities while investigating the relationship between complexity and elements of aptitude (phonological sensitivity, inductive ability, grammatical awareness, and rote learning ability) on narrative tasks. Cognitive task complexity was manipulated by increasing the need to conceptualise the plot of narrative tasks, which is viewed as resource-directing. Outputs were analysed using general measures as follows: fluency (speech rate) complexity (lexical variety was D-formula; syntactic complexity was clause length and ratio of subordinate clauses) and accuracy (ratio

of error free clauses) Task specific complexity (ratio of relative clauses) and accuracy (ratio of error free relative clauses, ratio of error free verbs, ratio of error free past tense verbs).

Kormos and Trebits' results showed that the participants displayed increases in lexical complexity in the writing mode, but lesser lexical complexity in the oral mode. However, results for syntactic complexity were not significantly different between modes. Separate from considerations of mode, increases in cognitive task complexity elicited more written complexity, measured as clause length and the ratio of relative clauses, which Kormos and Trebits consider to be support for the Cognition Hypothesis. As a result, there is some support for the Cognition Hypothesis, but only partial support for the benefits of mode, with the written mode benefiting lexical complexity, but not syntactic complexity.

Kormos and Trebbits (2012) suggest that the increases in lexical performance in the written modality may be due to the availability of online planning (inherent in writing) or due to teacher encouragement of increased lexical usage during writing. Regarding the absence of any significant effect for mode on syntax, Kormos and Trebbits suggest that the resource-directing nature of the tasks were the same for each modality; and as a result, the same syntactic features were elicited by the tasks.

Generalising the results of the studies, which specifically compare complexity and modalities, should be done cautiously considering that the body of work is small, and that within the small body of work, there are variations in the types of variables being used. Overall, the studies that focus on comparing modalities and cognitive task complexity show results that are not much different from previous work in the different modalities. There

appears to be no firm support for modality as a noticeably effective element influencing the output of tasks where cognitive task complexity is utilised in the input.

In terms of complexity (separate from comparisons of mode) there appears to be partial support for Robinson (2001a, 2001b, 2005, 2007a), with some effect for cognitive task complexity on lexical complexity, but little on syntax; this is similar to previous findings in the both oral and written modalities.

Despite the lack of evidence about the influence of mode, it is worth noting that in Kuiken and Vedders' study, mode may have influenced the different levels of syntactic complexity between the oral and written work. Though there was no effect for cognitive task complexity on written complexity, the oral modality did have less syntactic complexity as a result of cognitive task complexity. Problematically, no generalisation can be inferred from this one study.

2.2.5.5 Written modality, reasoning demands, and number of elements

Of the small group of studies on cognitive task complexity and writing, it is the work of Kuiken and Vedder (2007, 2008, 2011, 2012) and Sercu, De Wachter, Peters, Kuiken, & Vedder (2006) that has used independent and dependent variables similar to those used in this thesis. Kuiken and Vedder increased task complexity using number of elements and reasoning demands when researching the effects of cognitive task complexity on lexical and syntactic complexity in written output.

In terms of lexical complexity, Kuiken and Vedder had positive findings for increases in lexical sophistication (Kuiken & Vedder, 2007) and type-token ratio (Sercu et al., 2006) as

a result of increases in cognitive task complexity; however, this was only when using a type-token measure that did not account for text length, and the increases in lexical sophistication were inconsistent between groups. Kuiken et al., (2005) and Kuiken and Vedder (2008) found no increases in lexical variation resulting from increases in cognitive task complexity. Kuiken and Vedder (2012) found that the findings for lexical sophistication were inconsistent, with some one group of students showing significant increases in sophistication, and the other not.

For measures of syntactic complexity, Kuiken and Vedder (2008, 2012) and Sercu et al. (2006) analysed the number of clauses per T-unit and the number of dependent clauses per clause (with dependent clauses analysed as one group). The results of these studies found no effect for cognitive task complexity on syntactic complexity; therefore, in terms of syntactic complexity, these findings provided no clear support for the Cognition Hypothesis (Robinson, 2001a, 2001b, 2005, 2007a, 2007b); however, the lexical complexity results could be viewed as partial support.

No strong support for Robinsons' theory can be inferred from these results, with some effect for cognitive task complexity of lexical complexity only noticeable in type-token ratio measures that do not account for text length, and inconsistent results between groups when measured as lexical sophistication. As with the previous studies, these results seem consistent with work in the oral modality, with no effect found for syntactic complexity, and limited success for lexical complexity. Given the similarity in mode, independent variable, and dependent variable, it is these results that might be expected to be most relevant to this thesis.

As far as comparing reasoning demands, number of elements, and effect for mode, the results do not appear to show significant differences to previous work on cognitive task complexity in the oral modality, and cognitive task complexity (various measures) in the written modality. As a result, from the small amount of studies using similar independent and dependent variables as this thesis, there seems to be no significant effect for modality and the combination of these particular variables.

In terms of syntactic complexity (separate from considerations of mode), the non-significant results bear some consideration. The lack of any significant effect could be viewed as related to a number of issues. Firstly the inexact nature of measuring cognitive task complexity in the task input makes it difficult to know what type of effects are being generated between complex tasks. For example, too little variation in cognitive task complexity may not effect change; conversely, too much complexity between tasks may lead to participants being overloaded thus encouraging students to revert to more meaning focused (simplistic) output. In such cases, this thesis suggests that the addition of a baseline task, which is patently lower in cognitive task complexity, may provide clearer insight into whether cognitive task complexity is having any effect.

Another factor that may contribute to the non-significant results (regarding increases in number of elements and reasoning demands on written syntactic complexity) is the nature of the measurement. Measures of subordination, as utilised in studies by Kuiken and Vedder, are more complex that they appear (Bulte & Housen, 2012). It is possible that measures of subordinate clause length that collapse all dependent clauses into one inclusive group are imprecise and they miss certain aspects of performance that a more detailed subordinate measure might reveal.

2.2.5.6 Task complexity, written output, and syntactic complexity

In the studies above that use cognitive task complexity and measures of complex output, there have been various types of measures used to capture complex performance in written output. In this thesis, there are two general measures: lexical complexity and syntactic complexity, of which the syntactic measure of complexity measures the ratios of dependent clauses per T-unit.

In this thesis, the dependent clauses are measured as both one group, and also separately (adjectival, nominal, and adverbial). This has been done because the two different approaches may provide different results, with the detailed analysis provided by separating the dependent clauses revealing more about the effects of reasoning demands and number of elements on syntactic complexity than previous studies (Kuiken & Vedder, 2008, 2012; Sercu et al, 2006), which had non-significant findings for syntax.

2.2.5.7 T-units as measures of complexity

A justification for the use of T-units is required as they are an important aspect of this thesis and not without controversy. T-units are widely considered to be a good measure for writing (Ellis & Barkhuizen, 2005) because they provide a quantifiable unit for measuring segments of written language, and they are also viewed as a useful indicator of developmental progress in writing ability (Hunt, 1965; Hudson, 2009). The term developmental progress, as it relates to complexity, is the understanding that writer maturity correlates with an increase in the average length of T-units (Cooper, 1976).

The term maturity needs to be clarified, however. Hudson (2009) writes that linguistic maturity can be viewed as the types of linguistic patterns that are often found in examples

of mature writing and that often age and ability increase with maturity in the same way. Hudson goes on to state that two of the objectives of writing research are to determine those patterns that represent maturity, and research the influences that determine how these patterns develop. In this case, increases in the length of T-units, which often increase with age and ability, are the mature pattern of interest and are part of complex language development.

However, Crowhurst (1983) has stated that though longer T-units are a noted occurrence in the writing of superior writers, long T-units can be also be linked with flawed writing (Hake & Williams, 1979). Additionally, there can be a relationship between the length of a T-unit, its quality, and genre, with narratives and arguments having different relationships between length and quality in a young students' writing (Crowhurst, 1980). As a result, T-unit length, in and of itself, cannot be considered an acceptable measure without some qualifications.

This thesis does not take a default position that longer T-units are inherently better than shorter T-units. More complex doesn't necessarily mean better (Ortega, 2003; Pallotti, 2009) and short T-units may, in the hands of a good writer, be used to express meaning in a concise way that in some cases is more appropriate than long T-units. However, being able to access a variety of types of sentences is a good skill (Weaver, 1996), and the ability to use extended T-units, as a pattern recurring in mature writing, is something that should be learned as part of the arsenal of techniques used by those looking to become mature writers (Hudson, 2009).

2.2.5.8 Dependent clauses as measure of complexity

Wolfe-Quintero, Inagaki, and Kim (1998) state that measuring writing, of the type done in writing research in a second language, is not just about measuring the ability of students to write well, but can also focus on written development as it manifests and advances through the developmental process.

This thesis is focused on measuring the effect of cognitive task complexity on complex syntactic output, which represents an intermediate stage of written complexity. Norris and Ortega (2009, p.561) suggest how subordination may be situated as an indicator of language at an intermediate level.

> This theory of language posits that development proceeds from: (i) the expression of ideas first by means of mostly parataxis (i.e. coordination) or the sequencing of self-standing words, sentences, and clauses; through (ii) an expansion by which hypotaxis (i.e. subordination) is added as a resource to express the logical connection of ideas via grammatically intricate texts; to finally (iii) the emergence of and reliance on grammatical metaphor (achieved through nominalization, among other processes), which leads to advanced language that actually exhibits lower levels of subordination or grammatical intricacy but much higher levels of lexical density and more complex phrases (as opposed to more clauses).

This suggests that, as a measure of complexity, the ratio of dependent clauses per T-unit used in this study may be better suited to students who are generally considered to be intermediate learners because subordination has been viewed as a strong indicator of complexity for intermediate level language learners (Norris & Ortega, 2009).

However, viewing subordinate complexity as complex only in as much as it reflects stages along a developmental continuum may miss other dimensions contributing to the complexity of dependent clause production. Bulte and Housen (2012, p.36) refer to measures of subordination as hybrid measures, which "not only capture (syntactic) diversity, depth and compositionality but also difficulty". This study acknowledges that when measuring the ratio of dependent clauses per T-unit, other evidences of complexity may exist in the T-units; and that by ignoring these other elements of complexity, evidence of the effects of cognitive task complexity on other important aspect of writing may be missed.

Having acknowledged that subordination is likely a more complex measure than it appears to be, this thesis, while remaining focused on subordination, seeks to expand the purview of the syntactic measure by following Wolf-Quintero et al. (1998), who suggest that subordinate clauses could be analysed separately instead of collapsed into one group, as is normally the case in subordinate clause measures. Studying the effects of cognitive task complexity on each individual dependent clause does not appear to have been done in conjunction with studies on cognitive task complexity and writing.

In the previous section, the notion of maturity in T-units was explored; however, the more specific notion of subordinate clauses as complexity needs to be explored further in order to be clearer about what is being measured outside of the quantifiable length of a T-unit as no more than just a comparison of syntactic items that are visibly short or long.

There is some agreement that subordination complexity may be an issue of processing requirements. For example, Bygate (1999) states that T-units that have higher subordination require more complicated computation than T-units that evidence lesser subordination. In

addition, Lord (2002) researched the way in which students processed input both with and without subordination. It was noted that students, who scored lower in the activities containing subordinate clauses in the input, could have done so as a result of processing difficulty associated with processing the semantic meaning of the input. Alternatively, this may not have been the case with the input that was constructed using simple sentences without subordination.

Subordinate clauses may be a more difficult set of linguistic structures to process as they add a cognitive burden, which could interfere with the forming of correct meaning related outputs (Lord, 2002). This may be related to working memory capacity (Cheung & Kemper, 1992). Additionally, Cheung and Kemper (1992) have suggested that the act of embedding clauses can place extra demands on working memory as a result of the need to create and manipulate various syntactic elements. Mendhalson (1983) writes that subordination represents a more difficult and mature type of expression than simple coordination because subordination is a means to express relationships, complicated propositions, and form coherent organization between statements that are related.

If the assertions above are true, and if it can be assumed that different dependent clauses are learnt-acquired at different rates, then it is possible that the measure of syntactic complexity used in this study (ratio of dependent clauses to T-unit with each dependent clause analysed separately) could be measuring the effects of cognitive task complexity on processing heavy syntactic items at different stages along on the learning continuum. As a result, it is the assertion of this thesis that the separating of individual dependent clauses may be a more sensitive measure than ones where dependent clauses are collapsed into one group.

Thus, the separating of dependent clauses may give a clearer picture of the effects of cognitive task complexity on the syntax of intermediate learners.

It should be noted that previous studies using the same independent variables (number of elements and reasoning demands) and similar dependent variables have shown results for increases in cognitive task complexity that appear essentially neutral. It is the contention of this thesis that a more detailed analysis of the complex output might provide greater insight into the effects of number of elements and reasoning demands on measures of subordination in the written modality.

2.2.6 Conclusion

Cognitive task complexity has been suggested as a means to sequence and grade pedagogical tasks in a task-based syllabus. For this suggestion to work, research is required that shows that cognitive task complexity has the positive effects on output posited by Robinson. Moreover, it is important to know which elements of output are being effected as there seems to be a multitude of different measures within the dimension of complexity alone. Additionally, it seems important that the effect of mode and cognitive task complexity be clarified.

Though there are arguments that suggest that the written mode may be more conducive to the application of cognitive task complexity on complex output, a viewing of the results for oral and written studies in cognitive task complexity have shown little in the way of obvious differences; however, it should be noted that research in the area of writing and cognitive task complexity is small, thus more research is needed if the effects of cognitive task complexity on written complex output is to be clearly understood.

Of the small amount of studies in the field on cognitive task complexity and written output, the variation of task conditions, dependent variables, and independent variables across a small number of studies make it harder to generalize the results. As a result, it is the aim of this thesis to add to the research on specific independent and dependent variables that have not had a large body of research performed on them. With this in mind, this thesis has focused on two cognitive task variables: number of elements and reasoning demands, in the written modality. The previous studies in this area (Kuiken & Vedder, 2007, 2008, 2012: Sercu et al., 2006) have provided results for cognitive task complexity on complex output that are mostly consistent with previous studies in complexity, showing no noticeable effects for mode, limited effects for lexical complexity, and no significant effects for syntactic complexity. Importantly, the studies by Kuiken and Vedder (2007, 2008, 2012) and Kuiken and Vedder in Sercu et al., (2006) have been described as partial support for Robinsons' Cognition Hypothesis, for the most part because of the positive effects found for cognitive task complexity on accuracy, not complexity.

Another recurring issue in the research on cognitive task complexity is the lack of a quantifiable measure in applying amounts of cognitive task complexity and the effect this might have on the results. Given the inexact nature of applying amounts of cognitive task complexity, this thesis advocates that a baseline task that has patently lower cognitive task complexity needs to be added to research that increases complexity between tasks. The reasoning being that two tasks of differing complexity may produce the same effect on output if the amount of complexity in each task overloads the participant's attentional resources. The addition of a patently lower cognitive task may provide a clearer indication that variations in cognitive task complexity are having an effect.

Measures of subordination can be viewed as measures of resource-depleting syntactic items, which may be susceptible to the resource draining qualities of cognitive task complexity given that syntactic production is naturally disposed towards drains on attentional resources. In addition, individual subordinate clauses are likely to be learnt/acquired at different intervals; as a result, the effects of cognitive task complexity may be more noticeable on the differently acquired/learnt subordinate clauses, thus separating the subordinate clauses may provide a more sensitive instrument for syntactic measurement.

2.3 Pre-task planning time

2.3.1 Overview

In this research, the main independent variable is the manipulation of cognitive task complexity as reasoning demands and number of elements. However, there are further independent variables added to different questions. Pre-task planning time is an additional independent variable that explores the combined effects of increases in cognitive task complexity and pre-task planning time on complex written output.

Different types of planning are believed to be a component of all modalities of language production regardless of how effortless or difficult that production may appear to be. Planning has been broadly characterized as an activity that requires the selection of linguistic elements so as to create the desired effect upon the recipient of the output (Ellis, 2005). Two broad types of planning have been studied in the field of language research. Pre-task planning, which can be subdivided into rehearsal and strategic planning, takes

place before any given performance; and with-in task planning, which takes place during performance (Ellis, 2005).

The planning variable included in this thesis is pre-task planning. As mentioned above, pre-task planning can be divided into two general subcategories: *rehearsal* and *strategic planning*. Rehearsal involves giving the participants of any given task the chance to prepare for the main task by doing what amounts to a practice task first. Strategic planning time is the type of pre-task planning time that allows the participants of a language-learning task to prepare by thinking about the content of the task prior to performance (Ellis, 2005). In this thesis, only strategic pre-task planning time is reviewed (and utilised in the data collecting); as a result, from section 2.3.2 onwards, strategic pre-task planning is just referred to as pre-task planning time.

Strategic planning time can be further divided into two general sub-sets, referred to as *guided* and *unguided*. Guided planning involves some type of intervention in the pre-task planning process where the participants receive guidance regarding how or what to plan for the upcoming task. Unguided planning involves no help in preparing for the upcoming task (Foster & Skehan, 1996). This thesis is primarily concerned with the effects of *10 minutes, unguided, pre-task planning* in conjunction with increases in cognitive task complexity on complex language. Ten minutes has been selected as it is viewed as providing the best opportunity for producing positive effects in the output (Mehnert, 1998).

2.3.2 Pre-task planning time and oral modality

One aspect of this thesis is the consideration of modality as a possible factor contributing to the mélange of intersecting dimensions that come to bear when cognitive task complexity is

applied to the production of complex language. When adding pre-task planning as an additional variable, it may also be worthwhile to consider whether pre-task planning time is subject to issues of modality because most of the work in pre-task planning has been in the field of oral communication. As a result, it is from the oral modality that most of the current perceptions on the effects of pre-task planning come; and as such, it appears to be the best starting point for understanding its effects.

In general, the results for planning time on oral production have produced the following generalisations about the effects of pre-task planning time. While there is little evidence to support the notion that pre-task planning time has positive effects on accuracy, there is evidence that there are positive effects for pre-task planning on fluency (Gilabert, 2007b; Kawauchi, 2005; Ortega, 1999; Skehan & Foster, 2005; Wiggelsworth, 1997; Yuan & Ellis, 2003). In regards to complexity, pre-task planning time has been linked with positive effects for both grammatical and lexical complexity.

Because the focus of this thesis is complex output, and pre-task planning time is one of the extra variables utilised in this study; the following review will mainly spotlight the effects of pre-task planning time on complex output, as these results are considered most relevant to this thesis. However, even within this purview, it should be noted that the research into the effects of pre-task planning and complexity have used a variety of different complexity measures, variations in independent variables, variations in the amount of pre-task planning time, and the application of guided or unguided planning time. As a result, it is difficult to make strong generalisations about the results.

In the research on pre-task planning time that utilises both guided and unguided planning,

there have been a number of studies in the oral modality in which complexity was one of the elements measured in the participants' output. Additionally, there were variations in the amount of pre-task planning time provided.

The longest pre-task planning time was provided by Sangarun (2005), who utilised 15 minutes pre-task planning time while analysing the performance of students on two types of tasks using different types of guided planning. Complexity was measured as sentence nodes per T-unit, and clauses per T-unit. The results showed significant effects for guided planning on both measures of complexity.

Ten minutes pre-task planning time, which has been considered an effective amount of planning time, was used by Foster (1996) and Foster and Skehan (1996). Foster (1996) utilised both guided and unguided pre-task planning in a study that found that the use of pre-task planning might be linked to increases in complexity, measured as variety of past tense forms and clauses per c-unit (c-units are similar to T-units, but are more suited to spoken discourse as they account for the inclusion of certain types of incomplete clauses in the analysis). The effect was more noticeable in tasks where guided planning was used in conjunction with the more complex decision-making task. While utilizing similar task types, Foster and Skehan (1996) tested combinations of both unguided and guided pre-task planning and task type. Complexity was measured using a variety of verb forms and clauses per c-unit. Increases in complexity, measured as increases in the variety of past tense forms, was found on both the personal information and narrative tasks (the two less complex tasks). Additionally, the degree of detail given during the guided pre-task planning was associated with the degree of complex output.

Further studies utilised smaller amounts of pre-task planning time, with Rutherford (2001), Skehan and Foster (2005), and Mochizuki and Ortega (2008) all utilizing only 5 minutes. Rutherford (2001) investigated 5 minutes guided pre-task planning time on accuracy and complexity. Complexity was measured as length of c-unit and number of clauses per c-unit; however, no significant effect was found for either measure of complexity. Though Rutherford found no positive effect for complexity, positive findings for complexity were discovered by Skehan and Foster (2005), and Mochizuki and Ortega (2008).

As well as using only 5 minutes planning time, both Skehan and Foster (2005), and Mochizuki and Ortega (2008) investigated guided and unguided pre-task planning time, with Skehan and Foster also analysing online planning. Skehan and Foster (2005) used decision-making tasks, which are considered cognitively complex. Under some conditions, extra information was provided after 5 minutes to account for online planning. Complexity was measured as subordinate clauses per AS-unit (described as an utterance made up of an independent clause, or sub-clausal element, with any subordinate clause or clauses related to either), and in the findings more complexity (more subordinate clauses) was found in the condition where guided planning was provided in the first five minutes before the task started. Mochizuki and Ortega (2008) investigated a picture story re-telling task with audio narrative stimulus. Complexity was measured as mean length of T- unit; mean number of clauses per T-unit; and number of relative clauses per T-unit. The results showed an effect for guided planning on the increased number of relative clauses produced in the output. The studies below only investigated unguided pre-task planning time; however, like the studies above that investigated a combination of both guided and unguided planning time, a variety of dependent and independent variables were used in the studies. Additionally, there are variations in the amount of pre-task planning time provided to the participants;

however, most of the studies below utilised 10 minutes unguided pre-task planning time.

Three of the studies (Skehan & Foster, 1997; Wendel, 1997; Yuan & Ellis, 2003) that utilised 10 minutes pre-task strategic planning time had positive results for complex syntactic output. Ten minutes unguided pre-task planning time in conjunction with participants' ability to recall and retell information from two films was investigated by Wendel (1997). The results for complexity, which were measured using T-units and lexical variety, indicated that pre-task planning was linked to increased T-unit scores. Skehan and Foster (1997) also found positive effects for unguided pre-task planning on complexity, with complexity measured as clauses per c-unit. Pre-task planning was linked to more clauses per c-unit in personal information and decision making tasks, which are the least complex and the most complex task respectively, with no result for the medium complex task narrative task. Finally, Yuan and Ellis (2003), investigated unguided pre-task planning on the accuracy, complexity and fluency of narrative tasks, and significant effects were found for complexity. Complexity was measured as clauses per T-unit, the number of different verb forms, and a mean segmental type-token ratio. The significant result for pre-task planning and complexity was for more clauses per T-unit.

Other studies that focused on 10 minutes unguided planning time had results that only showed positive findings for lexical complexity. Ortega (1999) focused on the effects of planning on narrative tasks utilizing a set of pictures as prompts. Complexity was measured as words per utterance and type-token ratio. Unguided pre-task planning was found to have had a significant effect on words per utterance but not type-token ratio. It should be noted that Ortega used no measures of syntactic complexity. Furthermore, Gilabert (2007b) investigated unguided pre-task planning time on complexity with tasks manipulated along

the here-and-now dimension. Complexity was measured with Guiraud's index of lexical richness, and S-Nodes per T-unit. Ten minutes planning time appeared to contribute to increases in lexical richness for both complex and less complex tasks.

Only Kawachi's (2005) research yielded no positive effects for complexity from the group of studies that used 10 minutes unguided pre-task planning time. Kawachi studied the effects of three types of planning (10 minutes was allocated to writing condition; rehearsal condition; and reading condition) on picture based narrative tasks; planning time was unguided in the writing pre-task planning condition. Complexity was measured as number of clauses per T-unit, number of words per T unit, number of subordinate clauses, and number of word types. No effect was found for any of the planning conditions on the complexity measures

As well as the studies that utilised 10 minutes unguided pre-task planning time, there are also studies that have used shorter amounts of planning time, and could be considered less effective given the general consensus that 10 minutes tends to be effective for encouraging positive results in pre-task planning time studies. However, there were both positive results (research yielded expected production outcomes) and negative results (research did not yield expected production outcomes) for research that used less than 10 minutes pre-task planning time.

The shortest pre-task planning time was 3 minutes, utilised by Elder and Iwashita (2005) who tested unguided pre-task planning time on narrative tasks in testing situations. No result was found for complexity, which was measured as number of clauses per c-unit. As well as having the shortest pre-task planning time, Elder and Iwashita's study was the only

one in the pre-task planning time group using less than 10 minutes that showed no positive results for complex output.

The studies by Tavokoli and Skehan (2005), and Levkina and Gilabert (2012) both used 5 minutes unguided pre-task planning time, and both studies had findings with positive outcomes for complex output. Tavokoli and Skehan (2005) investigated unguided pre-task planning on narrative tasks of different complexity. Complexity was measured as clauses per AS-unit, with significant results found for unguided planning on increases in complexity and proficiency. Finally, Levkina and Gilabert (2012) investigated unguided pre-task planning time in conjunction with variations in the cognitively complex variable number of elements. Complexity was measured as lexical diversity, measured as Guiraud's (1954) index of lexical richness; and syntactic complexity, measured as mean number of clauses per AS-unit. The results showed some positive effects for lexical complexity, but there were no effects for syntactic complexity.

There are limits to generalising the above findings for unguided and guided pre-task planning time on the oral modality (Levkina & Gilabert, 2012). The difficulty with generalising the results is due to the variety of planning time conditions, the variations in planning time, and the variations in types of guided planning time across the body of research. Additionally, there were also variations in the types of measures used to analyse complexity.

A lack of uniformity, or any extended analyses of one set of dependent and independent variables, over a number of research projects suggest that the results are spread too thinly, making it difficult to make strong predictions at this time. However, looking at these results

as they stand, and taking into account the issues mentioned above, the following claims

could be made about pre-task planning time and complexity in the oral modality. For the

most part, guided pre-task planning time was present in studies that had positive effects for

complex oral output, measured across a number of different lexical and syntactic

complexity measures.

Regarding studies that utilised unguided pre-task planning time on the oral modality, most

of the findings seemed to show positive results for complex output with some positive

results for both lexical and syntactic complexity across a variety of different types of

complexity measures. Given these results, and the general consensus that the written

modality has greater planning time built into the modality, the same or better results might

be expected in the next section.

2.3.3 Pre-task planning time and written modality

This section addresses the literature on pre-task planning in the written modality. With the

exception of Ishikawa (2006), all of the studies that used guided or unguided pre-task

planning time used 10 minutes.

Ojima (2006) and Johnson, Mercado, and Acevedo (2012) were the two studies that used

guided pre-task planning time, though only Ojima (2006) had positive effects for lexical

and syntactic complexity. Ojima investigated the effects of guided pre-task planning,

initiated as concept planning on the written accuracy, complexity, and fluency of three

Japanese ESL learners. Complexity was measured as number of words per T-unit and

number of clauses per T-unit. Ten minutes was given during the planning phase, while four

essays were written, two with the pre-task planning component and two without. The topics

71

of these essays are not clearly explained therefore it is difficult to tell how complex these assignments were in terms of cognitive task complexity. The results of this study found effects for guided pre-task planning on fluency and complexity. Strategic planning as concept mapping was found to have a positive effect on both measures of written complexity; however, it should be noted that the sample group of three students is very small.

Johnson et al., (2012) also investigated 10 minutes guided pre-task planning time, utilizing a number of different types of guided pre-task planning on the written output. Unlike Ojima, the essays had a resource-directing element in the input, as participants were asked to give opinions and reasons; however, there were no increases in cognitive task complexity between tasks. A wide variety of complexity measures were utilised, though no effects were found for any of the different types of pre-task planning on grammatical or lexical complexity.

Both studies above report conflicting results for the effects of guided pre-task planning. Once again, it should be noted that generalisations are difficult with such a small group, and even between only two studies, there are variations in variables and conditions, which make comparisons difficult.

The writing research that used unguided pre-task planning is also small, comprising only four studies. Ong and Zhang (2010) utilised both 10 and 20 minutes pre-task planning time, while Ellis and Yuan (2004), and Rahimpour and Safarie (2011) both used 10 minutes. Ishikawa (2006) only used 5 minutes, which is not considered an optimal time for pre-task planning; however, Ishikawa was one of only two studies that produced positive results for

unguided pre-task planning on complexity, and was the only study that increased complexity along *resource-directing* dimensions.

Though not purporting to be about pre-task planning, Ishikawa (2006) is mentioned here because 5 minutes unguided pre-task planning time was provided to the participants of a study where cognitive task complexity was used in writing tasks. Increasing tasks along the here-and-now dimension, positive effects were found for all measures of complexity, both syntactic and lexical.

Ellis and Yuan (2004) can also be viewed as utilizing a resource-directing dimension as participants were required to interpret information in the tasks; however, there was no increasing of complexity along resource-directing dimensions. The effects of narrative writing tasks and three planning conditions (pre-task, online, and no planning) on the complexity, accuracy and fluency of written output were investigated, with positive effects found for complexity and fluency. Of most interest to this study is the effect of the pre-task planning (which had no detailed guidance) on written complexity, which had three measures: syntactic variety, the total number of different grammatical verb forms; syntactic complexity, the ratio of clauses to T-units; and a mean segmental type-token ratio. The results showed an effect for 10 minutes strategic planning on the complexity measure, syntactic variety.

The final two studies did not appear to utilise recourse-directing variables and revealed no positive results for 10 minutes unguided pre-task planning on complexity. Rahimpour and Safarie (2011) found no effect for 10 minutes unguided pre-task planning or online planning time on a descriptive writing task when complexity was measured by calculating

73

the percentage of dependent clauses to total number of clauses. Problematically, there was no complexity variable manipulated in the input and thus the results are difficult to compare to this thesis. Ong and Zhang (2010) increased the complexity of writing tasks through a number of variables, one of which was the removal of pre-task planning time. Both 10 minutes and 20 minutes planning time were used during the research. The planning time was unguided and any writing assistance was provided in the prompts. The results of this study indicated that the removal of pre-task planning time was linked to increases in fluency and lexical complexity; however, no results were presented for syntactic complexity. This result is unusual, as other findings in this area claim beneficial increases for pre-task planning on complexity

An overview of the results for pre-task planning on writing appears to show contradictory results. Two early unguided pre-task planning studies (Ellis & Yuan, 2004; Ishikawa, 2006) and one guided pre-task planning (Ojima, 2006) yielded positive outcomes for pre-task planning on lexical and syntactic complexity. The later studies, one using unguided pre-task planning (Rahimpour & Safarie, 2011) and one using guided pre-task planning (Johnson, Mercado, & Acevedo, 2012) showed no effect for pre-task planning on complex output. Additionally, Ong & Zhangs' (2010) study, which was unguided, appeared to show that pre-task planning time actually had a deleterious effect on complex output.

There are a number of issues that may be responsible for the contradictory results. Firstly, the small amount of research on pre-task planning and written complexity make it difficult to generalize of the results. Clearly more research in this field is required. Additionally, the small amount of research in this field uses different types of pre-task planning, a variety of types of tasks, and a variety of different measures of complexity. The combinations of these

different variables also make generalising difficult as the different combinations could be affecting the findings

It is the contention of this thesis that more studies need to be performed on the phenomena of pre-task planning and written complexity, with a view of taking a more exhaustive approach on a smaller number of variables before widening the scope of the research. Presently, it is difficult to make strong generalisations based on the present group of studies.

2.3.4 Pre-task planning time comparing modalities

The suggestion that the written modality might have a favourable effect with pre-task planning time on complex output is not clearly bourn out by comparing the results of studies that look at oral and written modalities. The results for the written modality seem to be evenly split, with results for guided and unguided pre-task planning on written complexity showing both positive, negative and neutral results for the combinations of different variables, conditions, and types of complex measures.

The studies in the oral modality, against expectation, seem to have a clearer result for unguided and guided pre-task planning time on lexical and syntactic complexity across a variety of independent variables, conditions, and measures. However, it should be noted that the oral modality has had a larger number of studies performed than those in the written modality, so any account for the two modalities should consider this imbalance.

This thesis seeks to add to the small number of studies in the area of unguided pre-task planning time on written complexity, a group that is even smaller when the additional variable of increases in cognitive task complexity is added.

2.3.5 Pre-task planning time, complexity, and the written modality

Once again, the paucity of research that focuses specifically on the area of cognitive task complexity, written complexity, and pre-task planning makes the area fertile ground for further research; however, the lack of a large body of past studies means there is less work in which to contextualize new findings.

In the preceding work on pre-task planning time and written complexity, there are few actual studies that combine cognitive task complexity with the pre-task planning time variable. Additionally, those studies that manipulate amounts of cognitive task complexity between tasks (which is to say, increases tasks along what Robinson would call resource-directing dimensions), in conjunction with manipulating complexity by adding or subtracting pre-task planning (resource-dispersing), are few.

Viewing how previous studies, which utilised pre-task planning time, cognitive task complexity, the written modality, and complex output; are different from this thesis may show more clearly where the gap lies for further research.

Discounting the complexity manipulated by the addition or removal of planning time, which is considered the manipulation of a resource-dispersing element by Robinson, the writing task used by Ellis and Yuan (2004) could be viewed as complex along a dimension not specified in their research. The tasks used by Ellis and Yuan require participants to

interpret information, which could be a resource-directing dimension according to the Cognition Hypothesis. However, there were no variations in cognitive task complexity between tasks along resource-directing dimensions as is the case in this thesis. This was also the case with Johnson et al., (2012), who provided essay writing tasks that contained the resource-directing dimension reasoning demands; however, there was no increasing or decreasing of complexity along the resource-directing dimension between tasks.

Ishikawa (2006) manipulated here-and-now (a resource-directing dimension) and also used 5 minutes unguided pre-task planning time. This thesis differs from Ishikawa by using a different resource-directing variable. Additionally, Ishikawa's pre-task planning time is only half the length of the time provided in this thesis.

Finally, Ong and Zhang (2010) manipulated complexity (excluding planning time) through the availability of writing assistance and the availability of a first draft. These methods of manipulating complexity are different from the resource-directing variables used in this thesis.

None of the above studies utilise a combination of pre-task planning time, cognitive task complexity as reasoning demands and number of elements manipulated between tasks, and the written modality. This is one of the gaps addressed by this thesis.

2.3.6 Conclusion

The preceding sections on pre-task planning have shown that pre-task planning, both guided and unguided, have had positive effects on complex output; however, there does not appear to be any noticeable effect for modality, with the written modality showing less

favourable effects (possibly due to a lesser number of research projects in the field). As a result, it is difficult to predict whether there will be any positive effects for cognitive task complexity and unguided pre-task planning time based on mode.

Of the studies that use pre-task planning on writing, half showed positive effects for planning on complexity and half did not. If the results are reduced even further to account for variables similar to this study, meaning those that are in the written modality and utilise cognitive task complexity and unguided pre-task planning time, the number of results is even smaller.

Ellis and Yuan (2004), and Ishikawa (2006) both utilise unguided pre-task planning time along with some type of cognitive task complexity in the task input. Both of these studies have produced results with favourable outcomes for complex output. Noticeably, both studies show favourable outcomes for syntactic complexity. This is particularly important given that syntactic complexity has not featured much in the results, with few showing anything positive. However, as noted earlier, the number of different independent variables used in these studies and the small amount of research in this field makes generalisations difficult.

Discounting the small number of studies and some of the differences in detail, a cautious optimism might be taken from these two studies in terms of positive outcomes for the effects of unguided pre-task planning time on written complexity.

2.4 Post-task editing time

The final additional variable is *self-editing time*, which happens during the post-task phase. Ellis (2003) describes the post-task phase as having three pedagogic goals. The first goal is to provide the participants with a chance to repeat the task. The second purpose of the post-task phase is to promote reflection about the task performance. The final goal is to promote attention to form; particularly those that the participants may have had trouble with during task performance. Ellis (2003) states that during the post-task phase, students can focus-on-formS, referring to Long's (1991) distinction where attention is focused on form rather than meaning.

Of the three goals mentioned above, editing time, as used in this thesis, would fit best with goal three, a chance to focus-on-formS. During the post-task phase, the participants would be engaging in undirected editing, which would require the participants to use their own resources to isolate and modify the written work. In the case of this thesis, the results would be focused on any possible post-task modifications to complex language, not accuracy.

The goal of unguided focus-on-formS during the post-task phase could be viewed as examining whether participants, freed from the cognitive pressures of performance during task production, are inclined to increase lexical and syntactic complexity during the post-task phase using their own resources. To date, there does not appear to have been any work focusing on written tasks that combine increases in cognitive task complexity and self-editing time focusing on complex output.

As a result, there is little research that can be used to contextualize the results for self-editing time for this thesis. Past studies on self-editing have concentrated on the widely

researched field of accuracy, whereas research on self-editing and complexity appears to be non-existent. Given that accuracy is viewed as correcting forms that are already controlled by the user, and complexity involves the learner pushing to use new forms and taking chances, the work on self-editing and accuracy would appear to be covering different areas of language production and is likely not transferable to self-editing and complexity.

Skehan and Foster (1997) and Ellis (2003) have described post-task requirements as weak, and Ferris, Liu, Sinha, and Senna (2012), in a study about written corrective feedback, noted that formal knowledge seemed to play a limited and possibly counter productive part in the self-editing process. It is possible from these few comments that the expectation for self-editing time on complexity may be low, and if the observations made by Ferris et al. (2012) hold true for this research, it does not seem likely that participants will engage in syntactic processing as a result of the provision of editing time. It is difficult to forecast the results given that there is no previous conclusive evidence upon which to make a prediction.

2.5 Attitudes towards tasks and task motivation

2.5.1 Introduction

The previous sections of this thesis (investigating evidence of causal relationships between cognitive task complexity, pre-task planning time, and post-task editing time on the production of complex output) might be viewed as somewhat narrow. What is meant by narrow is that the previous sections have a specific focus on the cause and effect relationships between a small group of indepednent and dependent variables; and therefore, do not take into consideration the fact that these realtionships do not exist in a vacumm. There is always a multiplicity of factors affecting the production and language learning.

To date, there does not appear to be any research focusing specifically on the effects of student attitudes on complex output (syntactic complexity). The final section of this thesis explores the potential additional effects on syntactic complexity resulting from learners' attitudes towards the relvance (a tasks' realation to students' goals, needs and values) and expectancy (the perception of difficulty and effort needed to complete a task) of the writing tasks performed during data collection.

The variables addressed in this section of the thesis differ from the previous variables because they are associated with the *difficulty* of task performance, as opposed to the *complexity* of task performance. As mentioned earlier (section 2.2.2), the concept of difficulty can be viewed as relating to *learner factors* (Robertson & Gilabert, 2007), which accounts for variations in performance based on an individuals' traits (e.g. anxiety, working memory capacity, motivation, aptitude).

The difficulty/ learner factors dimension is part to the wider issue of *individual differences*, which has been described as the study of the unique characteristics of an individual that alter and personalize the course of their language learning process (Dörnyei, 2009). Within the larger framework of individual differences, there are the two main phychological factors believed to have a strong effect on language learning. These are apptitude and motivation (Ellis, 2008).

Motivation is the focus of this section of the current thesis, specifically the area of motivation addressing the relationship between a learners' attitude towards a task and their performance. This area of research is termed *task motivation*, which examines the motivational effects resulting from a learners' reactions to the characteristics of tasks.

2.5.2 Motivation

Motivation is a much-researched phenomenon, one so widely researched that covering all the researched aspects of motivation would be too large to fit comfortably into this literature review. As a result, this thesis provides a very broad explanation of motivation before moving to the more specific area of task motivation.

Dörnyei (1999) has described the concept of motivation as hard to pin down in areas of educational psychology and applied linguistics, as it is a concept with multiple definitions. However, there is some consensus amongst researchers that motivation can be described as being "..related to a persons' choice of a particular action, persistence with it, and effort expended on it" (Manolopoulou-Sergi, 2004, p. 248). Additionally, motivation can be viewed as having two causal relationships to learning. In one sense it is viewed as being a strong support for learning, in another sense it is viewed as being the result of learning (Ellis, 2008).

What can be understood from these two statements about motivation is that a learners' action or persistence in completing a task in a task-based syllabus may be affected by the motivational traits a student brings to a writing task, and also the motivation created by the learning experience. It is mostly the motivation created by the learning experience that is central to this part of the thesis.

2.5.3 Motivation and learning situations

Dörnyei (2003) describes how previous macro-approaches to motivation (meaning research that analyzed motivation across whole learning communities) may not have provided an adequately detailed analysis of the motivational elements associated with the SLA

classroom. Changes in the approach to studying motivation were influenced by the notion that the environment in a classroom may have had a stronger impact on motivation than had previously been considered (Dörnyei, 2002; Julkunen, 2001; Dörnyei & Csizér, 1998; Dörnyei, 2010). As a result, a situated micro-approach to motivation was considered an appropriate method for the current component of this thesis, which focuses on how motivation is affected in the learning situation by the students' attitudes towards task characteristics.

Dörnyei (1994) suggested that leaning situations could have three effective components. Firstly the *teacher-specific motivational components* can be described as relating to how motivation is affected by a teachers' personality type, the style of teaching and feedback, and the types of relationships developed with the students. The second dimension is the *group-specific motivational components,* which relates to the dynamics that occur within a learning group, with Dörnyei (1994) referring to the social groups in learning environments as having the potential for strong impacts on students and their cognition. Thirdly there are the *course-specific motivational components*, which accounts for the motivational effects of syllabi, teaching materials and methods, and the types of tasks used in a classroom.

This thesis is concerned with the motivational effects of the course specific motivational components dimension, which would include students' attitudes towards tasks. Within the course specific motivational dimension, there are hypothesized to be four motivational elements that could affect a students' attitude towards a task. These four elements are *interest*, *relevance*, *expectancy*, and *satisfaction* (Crookes & Schimdt, 1991; Dörnyei, 1994; & Keller, 1983).

1. Interest is connected to intrinsic motivation (which is motivation based on an individual's internal factors as opposed to external factors) and relates to a student's inate curiosity about themselves and their environment.

2. Relevance is the degree to which class instruction is perceived to be related to a student's goals, needs and values.

3. Expectancy can be viewed as the student's perception of success in completing a task and includes perception of task difficulty and perceived amounts of effort required for completion.

4. Satisfaction, which concernes extrinsic motivational issues (motivation related to external factors affecting an individual) such as praise or curse marks; and intrinsic issues such as pleasure deriving from task participation and pride.

Of the four factors mentioned above, relevance and expectancy are the dimensions that feature in the Likert scale questionnaires used during the data collection component of the current thesis. To better understand how elements like relevance and expectancy might affect output, Dörnyei (2002, 2009, 2010) has developed a model of task motivational processing that speculates on the motivational processes involved with task motivation. This model is called *the motivational task processing system*. In the following section, the elements affecting task motivation are reviewed, including an explanation of Dörnyei's motivational task processing system.

2.5.4 Task motivation

Dörnyei (2002, 2005, 2009) states that task motivation has been described as a combination of motives that are both generalised and specific (Julkunen, 1989) and that both these

motives contribute to the amount of commitment a student might make to any given task. The generalized and specific dimensions of task motivation are described as being analogous to the more conventional terms trait motivation and state motivation.

Trait motivation (generalized) can be understood as coming from dispositions that are constant and enduring; whereas state motivation (situation specific) can be understood as temporary, coming from reactions to temporary conditions, such as any particular class or set of tasks (Tremblay, Goldberg, & Gardner, 1995). However, Dörnyei (2002, 2009) has proposed that the factors influencing task motivation may be more complex than the dual influences of trait and state motivation.

A weakness of the more traditional trait/state conceptualizing of task motivation is the static nature of the construct (Dörnyei, 2002), which doesn't clearly account for motivation maintained over longer periods of time. Drawing on past research by Heckhausen (1991), Heckhausen and Khul (1985), and Khul and Beckmann (1994); Dörnyei (2002) has recommended that a more dynamic understanding of the processes involved with task motivation might be more appropriate. Dörnyei proposes that there are three discernible motivational phases that a student will engage in. Very briefly, there is a pre-actional stage where motivation needs to be generated, an actionable stage where motivation needs to be maintained, and a post-actional stage where evaluations of past performance may affect future performance (Dörnyei & Otto, 1998; Dörnyei, 2002).

Though the current research acknowledges that the three phases may play an important role with motivation, the focus of this section of this thesis is specifically on the motivational effects that take place during the actionable stage. Dörnyei (2002, p.141) claims that the

85

actionable stage is characterized by five elements that influence motivation during task performance.

> Main motivational influences: Quality of the learning experience (pleasantness, need significance, coping potential, self and social image)….Teachers' and parents' influence….Classroom reward-and goal structure (e.g. competitive or cooperative)….Influence of the learner group….Knowledge and use of self-regulatory strategies (e.g. goal setting, learning and self-motivation strategies)

All five of these elements are speculated to have an effect on Dörnyei's (2002, 2003b, 2009, 2010) motivational task processing system, which is a tripartite model that hypothesizes the existence of three interacting mechanisms that a student would experience during the performance of a task.

The three-part process involves task execution, which is the students' actual engagement/ performance of the task through whatever form the task is presented to the student (in the case of this thesis, through the task instructions provided for the writing tasks); task appraisal, which is a continuous self-appraisal of performance in which students' process environmental elements and compare actual performance against anticipated performance; and action control, which is self-regulation where students' initiate mechanisms needed to "enhance, scaffold or protect learning specific action" (Dörnyei, 2003b, p.16) to maintain engagement during the task execution stage.

In terms of predictions for the motivational task processing system, Dörnyei (2009) claims that the quality and quantity of learner output will be the result of interaction between the

three interrelated dimensions that constitute the tripartite motivational task processing system. In a study that included task motivation as a variable, Dörnyei and Kormos (2000) stated that the results of task-motivated and task-unmotivated samples suggested that attitude was like a filter that can affect the regularity of task performance, with a raised filter leading to random performance. The negative effects of this filter were proposed to effect output variables such as accuracy, complexity and fluency. If the predictions by Dörnyei (2009) and Dörnyei and Kormos (2000) are realized in this thesis, then there may be some discernable correlation between learners' attitudes towards the relevance and expectancy of writing tasks and the production of mean length of T-units.

Finally, considering that much of this thesis associates the findings with the predictions of the Cognition Hypothesis (Robinson, 2001a, 2005), it might also be helpful to clarify how Dörnyei's predictions (2002, 2009, 2010) might relate to Robinsons' work.

2.5.5 Task Motivation, attention, and the Cognition Hypothesis

The Cognition Hypothesis views motivation as an affective element situated in the difficulty dimension, which relates to learner factors (individual differences) that affect task performance. The affective factor, motivation, is predicted to have an effect on attentional resource pools (and subsequently performance). Robinson (2001b) claims that higher levels of motivation could lead to temporary increases in the level of resource pools currently used to meet the needs of a pedagogical task. The fluctuation in attention caused by motivation may even affect what a participant pays attention to in task input (Manolopoulou-Sergi, 2004).

Extrapolating from this, it could be assumed then that the opposite might be true should the motivation be lower. When a task requires a lot of attention (like the complex tasks used in this thesis), it may be more susceptible to anything that negatively affects resources during performance. When considering both Dörnyes' hypothesis that task motivation may have an effect on output and Robinsons' speculation that motivation affects attentional resources, task motivation may have either a positive or negative effect on attention demanding complex tasks and the production of syntactic output, which is resource demanding to produce.

2.6 Literature review conclusion

A review of the literature on cognitive task complexity and different modalities has shown that, contrary to the notion that the written modality might be more conducive to the positive effects of increases in cognitive task complexity, there is no strong evidence to support the benefits of modality. In addition, research in the area of the written modality that utilises cognitive task complexity as an independent variable is a small group, with the results for syntactic complexity appearing to be neutral (meaning increases in cognitive task complexity appear to have no effect). This seems inconsistent with predictions made by both The Cognition Hypothesis and the Limited Attentional Capacity Model. This thesis proposes that these neutral results may be partly influenced by the inability to accurately measure cognitive task complexity between complex tasks, thus a baseline non-complex task might be required to track the effects of cognitive task complexity. Additionally, current measures of subordination may not be sensitive enough to measure the effects of cognitive task complexity.

As a result of these issues, RQ1 (research question 1) examines the effects of cognitive task complexity on writing, focusing specifically on number of elements and reasoning demands to expand on the small amount of research on these variables. In doing so, this research hopes to add to the small body of work on cognitive task complexity in the written modality, and contribute to an understanding of the relationship between modality and cognitive task complexity. In addition, a baseline non-complex task has been added in an effort to counter the issues regarding the inexact measurement of cognitive task complexity. Finally, a more sensitive measure of subordination is added in the hope of finding any effect for cognitive task complexity on syntax.

Pre-task planning time has shown to be a factor promoting positive outcomes for complex output in the oral modality, but with studies in the written modality appearing to have contradictory findings. However, it should be noted that many more studies have been performed in the oral modality, possibly skewing the results in favour of the oral modality and pre-task planning time. Furthermore, few of the pre-task planning and writing studies have dealt specifically with increases in cognitive task complexity and strategic planning time. Additionally, of the few that have, there are variations in the types of cognitive task complexity variables manipulated in the input. As a result, strong predictions based on past research and the particular variables used in this research are difficult. Finally, the issues mentioned above about levels of cognitive task complexity, baseline tasks, and measures of subordination also come into effect with pre-task planning studies as well.

In consideration of these issues, RQ2 (research question 2) focuses on adding to the small body of work that has used unguided 10 minutes pre-task planning time and increases in cognitive task complexity in the written modality. By utilizing the same independent

89

variables as RQ1 (number of elements and reasoning demands) this study hopes to create a small amount of consistency by using the same types of cognitively complex independent variables while adding the additional variable of unguided pre-task planning time. Furthermore, a baseline non-complex task and a more sensitive measure of syntactic complexity are added (as with RQ1) to address the problems mentioned earlier.

Self-editing time is an issue that does not appear to have been addressed in any studies that focus on the effects of cognitive task complexity and complex output. As a result, RQ3 (research question 3) spotlights the effects of 10 minutes post-task self-editing time, with the aim of discerning any effect for cognitive task complexity and self-editing time on syntactic and lexical complexity. As with the RQ1 and RQ2, the same issues regarding baseline tasks, and measures of subordination are accounted for.

Finally, RQ4 (research question 4) takes a different perspective than RQ1, RQ2, and RQ3, by focusing on task motivation, which is the relationship between a participants' attitudes towards the tasks and their performance. Unlike the previous three questions, this question looks more towards the effects of attitude, specifically how perceptions of cognitively complex tasks might affect performance. Taking a situated micro approach to the research on task motivation, RQ4 analyses how participant perceptions of task relevance and expectancy are correlated with syntactic output. Given the amounts of attention required by complex tasks, it might be expected that task-motivation elements (that are believed to affect attention) could be associated with variations in the performance of the syntactic output variables used in this research.

CHAPTER THREE

METHODOLOGY

3.1 Overview

This chapter presents the methodological approach and the methods used in this thesis to answer the research questions. Section 3.2 explains the normative methodological approach that underpins this research. A normative approach is viewed as the appropriate methodology for answering the research questions that have motivated this research. Combinations of both the experimental and associative dimensions of quantitative research are used to provide insights into the effects of cognitive task complexity on written complexity and the effects of task motivation on written complexity.

In the scope of the research (3.3), the aims of this thesis are contextualized by mentioning past research that has contributed to the inspiration of this thesis, and by explaining the way in which this thesis intends to build on previous research. Following this, the research questions (3.4) are stated along with an explanation of the additional issues that arose while investigating gaps in past research.

The target measures (3.5) section introduces the two measures of written complexity analysed in this thesis, including a variation on the standard syntactic measure. The measures are described and the reasons for using them are explained. Following this, the population and sample (3.6) are provided, presenting background information on the participants and the context in which the data was collected. The participants are presented

in three separate groups, with each group representing the research question they participated in.

The design section (3.7) clarifies and illustrates how the research was designed to answer the research questions. This is followed by a detailed description of the types of instruments (3.8) used to collect the data. The instruments are presented in conjunction with rationales as to why they were considered the appropriate instruments for answering RQ1, RQ2, RQ3, and RQ4. Examples of the instruments are also provided in this section. Subsequently, the data collection section (3.9) illustrates the procedures and different stages employed when administering the writing tasks, which were used to answer RQ1, RQ2, and RQ3, and the Likert scale questionnaire used to answer RQ4.

Finally, the data analysis (3.10) provides details of the data collected from the instruments, starting with clarification of coding issues related to T-units as well as presenting examples from the data, which makes clear how the dependent clauses were coded and analysed. This is followed by an explanation of the mean segmental type-token ratio, which is the second target measure. Examples of a mean segmental type-token ratio taken from the data are provided. Following this, the statistical procedures used to calculate the mean scores for the target measures are presented and clarified. Lastly, the process by which the Likert scale questionnaires were analysed is presented along with examples of the relevant scores and explanations of their meaning.

3.2 Methodological approach

Methodologically this research is grounded in a positivist/normative approach to research. Ellis and Barkhuizen (2005) describe normative research as investigations using the same methods as those used in natural science to identify cause and effect. Though this study acknowledges there are limitations to using only a normative method, the positivist/normative approach was viewed as the appropriate way to answer the four research questions.

The positivist/normative approach to research is typically quantitative (Ellis & Barkhuizen, 2005), which can be viewed as having two general types. One type seeks to establish causality between the variables, and one does not. Mackey and Gass (2005) describe these two categories of quantitative research as experimental and associative. This thesis has utilized both experimental and associative quantitative research methods in the process of answering this thesis' four research questions.

The main body of this thesis is experimental in which independent variables are manipulated with the express purpose of determining whether there are any effects on the dependent variables, in essence trying to establish causality between the two variables (Mackey & Gass, 2005). This experimental quantitative approach covers RQ1, RQ2, and RQ3, as these three research questions involve the manipulation of the independent variables (cognitive task complexity, pre-task planning time, and post-task editing time) in order to determine any noticeable effects on the dependent variables, syntactic and lexical complexity.

The second view of qualitative research is associational (Mackey & Gass, 2005). An associational approach seeks to determine what type of relationship might exist between variables, and if so, the strength of that relationship. Associational qualitative research does not make any claims about the causal relationship between the variables, only that a relationship may or may not exist. This approach is utilised in RQ4, which examines the strength of association between the participants' opinions of the pedagogic tasks used to collect data for RQ1, RQ2, and RQ3, and the occurrence of syntactic complexity in the output.

3.3 Scope of the research

This thesis expands on the small body of work focusing on the effects of cognitive task complexity on complex written output. To date, there are few studies that have utilised the same central independent variables (reasoning demands and number of elements) as this thesis on the production of lexical and syntactic complexity.

Past research in the written modality, utilising reasoning demands and number of elements (Kuiken et al., 2005; Kuiken & Vedder, 2007; Kuiken & Vedder, 2008; Sercu et al., 2006), has investigated both complexity and accuracy in the output. Regarding the complexity aspect of those studies, Kuiken and Vedder analysed both lexical and syntactic complexity in the output, something this thesis also does. However, this thesis does not include accuracy in the analyses.

This thesis differs to the work of Kuiken et al., (2005) Kuiken and Vedder (2007), Kuiken and Vedder (2008), and Sercu et al. (2006) in a number of other ways. Firstly, a low

complexity control task, against which more complex tasks can be compared in the

analyses, has been added to the design. Secondly, there are two extra independent variables

added to the study of reasoning demands and number of elements in the written modality.

These variables are 10 minutes pre-task planning time (in RQ2) and 10 minutes post-task

editing time (in RQ3).

In order to build on the findings of previous research, this thesis has utilised similar aspects.

Both this thesis and the previous studies utilise similar types of letter-writing tasks that

manipulate the same dimensions of cognitive task complexity in the task design.

Additionally, both studies analyse complex output using T-units and type-token ratios.

However, this research has also expanded on previous the work by using a mean segmental

type-token ratio, which measures lexical complexity while accounting for text length.

Additionally, traditional measures of T-units are used alongside a non-standard variation of

T-unit measurement. The non-standard measurement of T-units involves measuring all

dependent clauses separately, something not done in previous studies.

Task motivation is also addressed in this thesis, something not previously studied in

conjunction with cognitive task complexity and its affects on T-unit length. Task

motivation refers to the prosed effects tasks have on learners' motivations. To do this, a

Likert scale questionnaire was added to elicit the participants' perceptions of the tasks they

performed during the study. In doing so, it is the aim of this thesis to widen the scope of the

study by better understanding the relationship between task motivation and syntactic

complexity.

3.4 Research Questions

This study has four questions used to investigate the gaps in the research noted in the previous chapters.

RQ1: What are the effects of cognitive task complexity on written complexity?

Research question 1 explores the effects of increased cognitive task complexity (reasoning demands and number of elements) on the syntactic and lexical complexity of second language writing. Additionally, issues relating to the effects of modality, the effectiveness of standard measures of mean length T-units, and the utility of including a patently non-complex task for improved tracking of the effects of cognitive task complexity are discussed in the findings of this question.

RQ2: What are the effects of pre-task planning time combined with cognitive task complexity on written complexity?

Research question 2 explores the effects of 10 minutes unguided pre-task planning time in conjunction with increased cognitive task complexity (reasoning demands and number of elements) on the syntactic and lexical complexity of second language writing. As with RQ1, issues relating to the effects of modality, the effectiveness of standard measures of mean length of T-units, and the utility of including a patently non-complex task for improved tracking of the effects of cognitive task complexity are also explored in the findings of this question.

RQ3: What are the effects of post-task editing combined with cognitive task complexity on written complexity?

Research question 3 investigates the effects of 10 minutes post-task editing time in conjunction with increased task complexity (reasoning demands and number of elements) on the syntactic and lexical complexity of second language writing. As with RQ1 and RQ2, the additional issues of the effects of modality, the effectiveness of standard measures of mean length of T-units, and the utility of adding a patently non-complex tasks for improved tracking of the effects of cognitive task complexity are also explored in the findings for this question.

RQ4: What is the relationship between the participants' attitudes and complex written output?

Research question 4 investigates the area of task motivation by analysing Likert scale questionnaires for any correlation between the participants' attitudes towards task relevance (a task's perceived relevance to students' goals needs and values) and task expectancy (including perceptions of task difficulty and the effort required for completion) and their production of dependent clauses.

3.5 Target measures

The bulk of the data used in this research was taken from three letter-writing tasks (see section 3.8.2 below for definitions and examples), which were designed to answer RQ1, RQ2, RQ3, and RQ4. Two target measures were analysed in the letter-writing tasks.

The two target measures used in this thesis are measures of complex output, one of which targets syntactic complexity while the other targets lexical complexity. The first target measure is clauses per T-unit, which is a measure of syntactic complexity that analyses the ratio of dependent clauses to T-units. The second target measure is a mean segmental type-token ratio, which is a measure of lexical complexity. Although prior studies on cognitive task complexity and writing (Kormos, 2011; Ong & Zhang, 2010; Kuiken & Vedder, 2007, 2008; Sercu et al., 2006) have utilised a wider variety of target measures in their analyses of the effects of cognitive task complexity on complexity, accuracy, and fluency, it was felt that this thesis would focus specifically on the two measures of complexity.

The reasons for choosing these two specific target measures are as follows. Firstly, the clauses per T-unit (or mean length of T-units) and mean segmental type-token ratio are both used by other studies utilizing reasoning demands and number of elements, for example, Kuiken and Vedder (2008, 2012). Considering that one of the aims of this thesis is to build on previous work such as Kuiken and Vedders', it was considered prudent to use parity in the choice of target measures to add to the reliability of this thesis. Secondly, when examining the effects of cognitive task complexity on written complexity, it was important to use more than one type of complexity measure because language learners may use different means to express complexity in their writing (Ellis & Barkhuizen, 2005).

T-units are widely considered to be a good unit of measurement for writing (Ellis & Barkhuizen, 2005) and are viewed as a good indicator of developmental progress in writing ability (Hunt, 1965). It is believed that writer maturity correlates with an increase in the average length of T-units (Cooper, 1976). Hunts' (1965) description of a T-unit, which is

an independent clause (at its base level a subject and a finite verb) and any subordinate clauses (adverbial, nominal, adjectival) attached to the main clause, is the T-unit definition used for this research. Additionally, the T-units analysed in this study included both error free and error inclusive T-units. In Table 13 (see page 121), a T-unit and T-unit combinations with nominal, adverbial, and adjectival dependent clauses are illustrated.

As mentioned above, error free and error inclusive T-units were analysed together in the data. The reasoning behind this was the pervasiveness of errors in the data due to the participants' developmental stage, with errors occurring frequently in adult second language data (Gaies, 1980). To exclude error inclusive T-units would be to dismiss most of the data. Additionally, the focus of this research, as it regards T-units, is T-unit depth not T-unit accuracy. However, parameters for what should be considered an acceptable error inclusive T-unit needed to be set as there were occurrences of error inclusive T-units that were so flawed that they could not be analysed as positive outcomes for the effect of task complexity. The parameters for acceptable error inclusive T-units are illustrated in section 3.10.1.

Finally, the use of T-units is not without controversy. A fuller discussion on the problems with T-units and why they were considered appropriate for this study can be found in the literature review chapter in section 2.2.5.7.

As mentioned above, two measures of complexity are used in this study to account for the different ways in which a participant might express complexity in their written output. The second measure is a mean segmental type-token ratio, which is a measure of lexical variety.

In a review of lexical complexity measures, Wolfe-Quintero et al. (1998) state that measures of lexical variation appear to be related to the development of second language, particularly those measures that also account for text length. The type-token ratio measure of lexical variety used in this thesis is a measure of lexical variety that accounts for variations in text length.

3.6 Population and sample

The population from which the samples were obtained consisted of non-native speakers of English who were studying English in New Zealand. All the participants were classified as intermediate level language learners by the schools in which they were enrolled with IELTS levels ranging from 4.5 to 5.5. It was upon these IELTS scores that the assumptions of participant proficiency were initially made.

94 non-native speakers of English volunteered to take part in this study. There were a variety of nationalities in the group: Japanese (3), Chinese (40), Korean (17), Thai (5), Russian (4), Vietnamese (4), French (3), Burmese (2), Arabic (12), Indonesian (1), Portuguese (1), and Turkish (2), There were 49 female and 45 male participants, and though the ages ranged from 18 to 60, they were predominantly in the twenties age group with 66 of the 94 participants between 20 and 29 years of age.   The average amount of time each student had been studying English was 8 years. There was variation in the amount of time each student had been studying English in New Zealand, with the average time each student had studied in New Zealand (across all 94 participants) being 2.3 months. The 94 participants were separated into three separate groups (RQ1, RQ2, and RQ3), which each addressed a different research question.

The participants of RQ1, which focused only on changes in cognitive task complexity, varied in nationalities as follows: Japanese (1), Chinese (13), Korean (11), Thai (3), Russian (2), Vietnamese (2), and French (2). There were 21 female and 13 male participants, and though the ages ranged from 18 to 41, they were predominantly in the twenties age group with 26 of the 34 participants between 20 and 29 years of age. The average amount of time this group had been studying English was 9 years. The average amount of time this group had spent studying English in New Zealand was 2 months.

The participants of RQ2, which focused on changes in cognitive task complexity plus 10 minutes pre-task planning time, varied in nationalities as follows: Indonesian (1), Chinese (15), Korean (3), Thai (1), Russian (2), Vietnamese (2), Portuguese (1), Turkish (2), Arabic (2) and French (1). There were 18 female and 12 male participants, and though the ages ranged from 21 to 40, they were predominantly in the twenties age group with 24 of the 30 participants between 21 and 27 years of age. The average amount of time this group had been studying English was 10 years. The average amount of time this group had spent studying English in New Zealand was 3 months.

The participants of RQ3, which focused on changes in cognitive task complexity plus 10 minutes post-task editing time, varied in nationalities as follows: Japanese (2), Chinese (12), Korean (3), Thai (1), Burmese (2), and Arabic (10). There were 10 female and 20 male participants, and though the ages ranged from 18 to 60, they were predominantly in the late teens to twenties age group with 23 of the 30 participants between 18 and 27 years of age. The average amount of time this group had been studying English was 5 years. The

average amount of time this group had spent studying English in New Zealand was 2 months.

The learning contexts for the participants were four language schools situated in Auckland. The schools specialized in motivated students who had come to New Zealand with the expectation of actually improving their English skills. Of the 94 participants, 84 were interested in studying at University level in the English language and as such were motivated to improve their writing skills. The courses run by the four schools focused on raising students overall proficiency in reading, writing, speaking, and listening.

3.7 Design

This thesis utilised three separate letter-writing tasks (Task 1, low complexity; Task 2, medium complexity; Task 3, high complexity). Each task was designed to have different amounts of cognitive task complexity that were manipulated by increasing the complexity dimensions reasoning demands and number of elements in the task instructions.

As mentioned above, the participants were divided into three separate groups of around 30, to address the first three research questions. These are referred to as RQ1, RQ2, and RQ3. Each group performed the three different letter-writing tasks (Task 1, low complexity; Task 2, medium complexity; Task 3, high complexity); however, there were also additional variables added to RQ2 and RQ3. Whereas RQ1 was solely focused on increases in cognitive task complexity, RQ2 combined increases in cognitive task complexity with 10 minutes pre-task planning time, and RQ3 combined increases in task complexity with 10 minutes post-task editing time.

RQ1, RQ2, and RQ3 followed the same basic data-collecting procedure. The data

collection took place over two stages. Stage 1 was exactly the same for each of the three

research questions; however, the extra variables of pre-task planning time and post-task

editing time were initiated during stage 2 of the data collection process, thus it was stage 2

where the design was slightly different between RQ1, RQ2, and RQ3.

During stage 1, the researcher initially handed out a survey that was designed to gather

personal information about each participant. This was followed by Task 1 (low

complexity). When the task was finished, stage 1 was completed.

Table 2: The two steps of stage 1 for RQ1, RQ2, and RQ3

Stage 1 for RQ1, RQ2, and RQ3	
Step 1	Step 2
Participant survey: 10 minutes.	Task 1(low complexity) 30 minutes.

Stage 2 of the data collection process was mostly similar for each research question;

however, it was in stage 2 that additional variables were added to RQ2 and RQ3. While

RQ1-stage 2 focused solely on changes in cognitive task complexity, RQ2-stage 2

addressed cognitive task complexity in conjunction with 10 minutes pre-task planning time,

and RQ3-stage 2 addressed cognitive task complexity in conjunction with 10 minutes post-

task editing time.

The data collecting process for RQ 1-stage 2 involved the completion of two consecutive

letter-writing tasks of differing levels of cognitive task complexity (Task 2, medium

complexity; and Task 3, high complexity). The participants were allowed 30 minutes to complete each task. It should be noted that Task 2 and Task 3 were counterbalanced (see table 6) to account for participant fatigue. When both Task 2 and Task 3 were finished, the participants completed a Likert scale questionnaire. This completed the data collection for RQ1-stage 2.

Table 3: The three steps of RQ1, stage 2

Stage 2 for RQ1.		
Step 1	Step 2	Step 3
Task 2 or Task 3 depending on counterbalancing: 30 minutes.	Task 2 or Task 3 depending on counterbalancing: 30 minutes.	Likert scale questionnaire: 5 minutes.

The data collecting process for RQ2-stage 2 also involved the completion of Task 2 (medium complexity) and Task 3 (high complexity). These tasks were counterbalanced in the same way as RQ1-stage 2. Where RQ2-stage 2 differed was the inclusion of the variable 10 minutes pre-task planning time. 10 minutes planning time was allowed prior to the writing of both Task 2 and Task 3. Participants were encouraged to take notes and plan for the task during this time, but they were not allowed to begin writing the tasks until the 10 minutes planning time was completed. Subsequently, the participants had 30 minutes to complete each task, during which they were able to refer to the notes they had taken during the planning time. After the final task was completed, the participants filled out the Likert scale questionnaire.

Table 4: The three steps of RQ2, stage 2

Stage 2 for RQ 2.		
Step 1	Step 2	Step 3
10 minutes pre-task planning. Task 2 or Task 3 depending on counterbalancing: 30 minutes.	10 minutes pre-task planning. Task 2 or Task 3 depending on counterbalancing: 30 minutes.	Likert scale questionnaire: 5 minutes.

RQ3-stage 2 utilised the same counterbalanced approach to the task order of completion used in RQ1-stage 2 and RQ2-stage 2. RQ3-stage 2 had the added variable of 10 minutes post-task editing time at the completion of Task 2 and Task 3. At the end of the 30-minute writing time allowed for each task, participants were given a red pen, which was a different colour to the pens used in the writing of the tasks. Subsequently, the participants were instructed to edit the work that they had just completed. After Task 2 and Task 3 were completed, the participants answered the Likert scale questionnaire.

Table 5: The four steps for RQ3, stage 2

Stage two for RQ3		
Step 1	Step 2	Step 3
Task 2 or Task 3 depending on counterbalancing: 30 minutes. 10 minutes post-task editing time.	Task 2 or Task 3 depending on counterbalancing: 30 minutes. 10 minutes post-task editing time.	Likert scale questionnaire: 5 minutes.

Table 6 below shows how the tasks were counterbalanced by running two consecutive streams of participants (for each research question) during stage two of the data collection

process. Random selection was used to identify which participants would start with either

Task 2 or Task 3 first.

Table 6: Counterbalancing Task 2 and Task 3 during stage two

RQ1, RQ2, and RQ3.		1st task assigned to participants	2nd task assigned to participants
RQ1: 34 participants	17 participants	Task 2 (medium complexity)	Task 3 (high complexity)
	17 participants	Task3 (high complexity)	Task 2 (medium complexity)
RQ2: 30 participants	15 participants	Task 2 (medium complexity)	Task 3 (high complexity)
	15 participants	Task 3 (high complexity)	Task 2 (medium complexity)
RQ3: 30 participants	15 participants	Task 2 (medium complexity)	Task 3 (high complexity)
	15 participants	Task 3 (high complexity)	Task 2 (medium complexity)

3.8 Instruments

Three different types of instrument were utilised during the data collection process. The

first instrument was a survey, used to collect background information on the participants.

The second instrument type was the letter-writing, tasks and the third instrument type was a

Likert scale questionnaire.

Task 1 Task 2, and Task 3 were pedagogic letter-writing tasks, which were delivered in the

form of written prompts. The type of pedagogic letter-writing tasks used in this thesis were

expected to elicit learner language samples that reflected a participants' language

competence (Ellis & Barkhuizen, 2005). The Likert scale questionnaire was used to elicit a

participants' immediate response to the tasks themselves. The immediate responses were

expected to provide information regarding attitudes towards the tasks and the ability to

perform them.

The letter-writing tasks and the Likert scale questionnaire were selected as they are viewed as appropriate for collecting the types of data required for answering RQ1, RQ2, RQ3, and RQ4. The learner language samples elicited by the pedagogic letter-writing tasks are used to investigate the effects of cognitive task complexity on the participants' written output. Ellis and Barkhuizen (2005) have termed this type of learner language sample *expression,* and it is language that is used primarily as a way to infer performance competence.

The Likert scale questionnaire, used to answer RQ4, does not elicit learner language samples, as there are no resultant texts. The numerical data elicited by this instrument is expected to provide insights into the participants' point of view, which is then correlated with the syntactic complexity results from RQ1, RQ2, and RQ3. The Likert scale questionnaire was also chosen because of data collecting issues related to participant availability. Availability refers to participant fatigue levels, which may be a factor given that the Likert scale questionnaire was completed subsequent to the completion of the letter-writing tasks. Likert scale questionnaires are considered a less demanding option for collecting data than one in which participants were expected to write their answers down.

3.8.1 Survey

The background survey was designed to gather information about the participants. This included names, contact details, age, gender, ethnicity, first language, time spent studying English both at home and in New Zealand, interest in studying at University in an English speaking country, time spent practicing writing, time spent reading in English, what type of English was read, and self-perceptions of English ability.

Table 7: Example of participant survey

_____

Name
Email
Gender
Age
First language
How many years have you been studying English?
How old were you when you started studying English?
How long have you been studying English in New Zealand?
How long have you been studying English writing?
Do you want to study at University in English?
How long have you been studying academic writing?
How often do you do writing homework?
Do you write English outside of class-time (not homework)?
What type of English writing do you do outside of class?
How often do you write in English outside of class time including homework?
Are you currently studying at more than one language school?
If the answer to the last question was yes, how many schools, including this school, are you studying at?
How often do you read in English (not including school work)?
What type of things do you read in English outside of schoolwork e.g. English novels, English magazines etc.?
What level do you think your English writing is? Is it very poor, poor, average, above average, good, very good, or really good, or excellent.

_____

3.8.2 Letter-writing tasks

Three letter writing tasks: Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity) were used in this research. The letter writing tasks were designed to elicit samples that reflected the writing competency of the participants who were subjected to increases in cognitive task complexity, which is posited to place increasing demands on attentional resources.

Letter-writing tasks were chosen for two main reasons. Firstly, any task used to collect data should be authentic in that it reflects a real world activity. An argument could be made that the authenticity of letter-writing tasks as modern real-world activities may be undermined by the rise of emails in modern communication. However, Ellis (2003) points out that tasks

do not always correspond to the types of activities that a participant will do during their day-to-day lives. Furthermore, Skehan (1996) writes that a task may only need to have a relationship to the real world, which means that the types of behaviours that are elicited by tasks need only correspond to the types of behaviours that arise in real-world tasks. Though writing letters may not currently be as popular among students as the sending of emails, it remains a real-world activity, or for the participants, one that reflects communicative behaviour from the real world. The second reason for choosing letter-writing tasks is that this thesis aims to expand on previous studies (Kuiken & Vedder, 2007, 2008) using similar independent and dependent variables. As such, the use of similar types of tasks was expected to contribute to the validity and reliability of the research.

There were three variations of letter-writing task, each designed to initiate different levels of task complexity through the manipulation of cognitive task complexity in the task instructions. Task 1 (low complexity) was designed to apply the least amount of stress on the participants' attentional resources, whereas Task 2 (medium complexity) and Task 3 (high complexity) were both designed to apply higher levels of pressure on the participants' attentional resources while completing the tasks.

Task 1 (low complexity) consisted of a written handout that provided a situation and a set of instructions. The situation involved an English-speaking friend who was thinking of moving to New Zealand. The instructions directed the participant to write to this friend about New Zealand. Task 1 was considered a low complexity task for the following reasons. Firstly, the task instructions avoided phrases that clearly indicated that the writer should form any opinions or state any reasons why New Zealand might be worthwhile

moving to. The absence of any prompting to form opinions or give reasons was considered less cognitively stressful than Task 2 and Task 3 in which the prompting of opinions and reasons were part of the complexity variable reasoning demands. Furthermore, Task 1 did not supply additional information for the participants to use in the letter, unlike Task 2 and Task 3. As a result, Task 1 participants were expected to rely on their own resources, thus they were expected to produce language at a level they were comfortable with as opposed to being pushed to synthesize the extra information. The complexity variables used in Task 2 and Task 3 were expected to push the participants to process language at the edge of their abilities; whereas, the absence of these variables in Task 1 was predicted to allow processing that was not pushed and was thus, potentially less cognitively stressful.

Table 8: Example of task rubric for Task 1 (low complexity)

Situation.
You have a close English-speaking friend called Peter.
Peter is thinking about moving to New Zealand.

Instructions.
You have to write a letter to your friend of about 200 to 250 words.
In this letter you should write to Peter about New Zealand.
Start the letter below.

Task 2 (medium complexity) was made complex by the manipulation of reasoning demands and number of elements in the task instructions. This was expected to induce higher levels of attention demanding activity (Brown et al., 1984; Ellis, 2003; Robinson, 2005, 2007a) in order to complete the task.

Table 9: Example of task rubric for Task 2 (medium complex)

Situation.
1. Your friend John is coming to New Zealand for one weekend, and there are two restaurants he really wants to try.
2. There is only time to go to one restaurant. As a result, John wants you to choose one restaurant.
3. Neither of the restaurants you have checked is perfect for John and your requirements.

Instructions.
1. Look at John's requirements in list A.
2. Look at the restaurant information in list B.
3. Consider your own personal preferences.
4. Using the information from lists A and B and your own preferences, write John a letter of between 200 and 250 words telling him which restaurant you have chosen and why you choose it.
5. Start the letter below.

The cognitive task complexity dimension reasoning demands was incorporated into the task by using instructions that directed the participants to write to a friend who was coming to New Zealand and inform that friend which of the two restaurants they would visit upon the arrival. Information on the two restaurants was supplied with the task instructions. In addition, the choice of restaurant had to be justified based on information provided about the restaurant and the visiting friends' (John) information.

Table 10: John's information from Task 2 (medium complexity)

John's information:
1. He is arriving on Saturday morning and leaving on the following Monday afternoon.
2. Seafood and Pork are his favourite food.
3. He generally eats a lot.
4. He doesn't particularly love sweet food, but enjoys dessert some times.
5. He likes to drink a glass of wine with dinner.
6. He only speaks English.
7. He will be staying with you during his time here, so transportation will be your responsibility.

Table 11: Example of one restaurant from Task 2 (medium complexity)

Restaurant 1:
Opening times: 4:00 pm to 8:00 pm Monday to Saturday.
Prices: Main courses (main meal) cost around $20.
Availability: Usually the restaurant is very busy and bookings (reserve a table) are necessary to get a table.
Critic's review of food quality:
The seafood selection (what is available) is good, and is considered very high quality.
The beef is average quality.
The pork is average quality.
There are no desserts (ice cream etc.) at this restaurant.
The portions (size of meal) are average size.
Drink: The beer and wine is expensive. There is no BYO (bring your own drinks).
Staff: Some staff speak English, some staff only speak Japanese.
Service: The service is quick, but the staff does not appear friendly.
Entertainment: Karaoke after 7:00pm.
Location: In Auckland's central city.
Parking: Restaurant supplies no parking

The second cognitive task complexity dimension was number of elements. This dimension was added to the task instructions by supplying more elements that needed to be considered when completing the task than were used in Task 1 (low complexity).

Task 3 (high complexity) was similar to Task 2 (medium complexity) except there was a greater increase in the complexity dimension *number of elements*, thus making it a more complex task. This increased the number of elements that needed to be considered as well as increasing the reasoning demands. Instead of two restaurants to choose from, there were three. In addition, the participants were also expected to consider the information of two other friends (Marvin and Jerry) who would be coming to the restaurant as well as the person receiving the letter (Kate)

Table 12: Example of friends' information from Task 3 (high complexity)

Jerry's information:
1. He lives in the city centre in an apartment.
2. He eats anything, but really likes seafood.
3. He eats a lot.
4. He doesn't eat dessert.
5. He likes to drink beer.
6. He speaks Japanese and Korean, and his English is ok.
7. He has a car.
8. He enjoys Karaoke.

Marvin's information:
1. He lives on the North Shore.
2. He loves chicken and beef, but doesn't eat seafood.
3. He eats a lot.
4. He eats dessert, but doesn't really care about it.
5. He drinks, but it is not important to him if the restaurant has beer or wine.
6. His English is average, his Japanese is average, and his Chinese is good.
7. He has a car.
8. He doesn't care about entertainment at the restaurant.

3.8.3 Likert scale questionnaire

Whereas the letter-writing tasks produce data related to participant performance, the type of

information obtained by Likert scale questionnaires can represent social identity, styles of

learning, motivation, aptitude, and learning strategies (Ellis & Barkhuizen, 2005). For this

thesis, the Likert scale questionnaire was designed to obtain data on the dimension of

motivation.

The Likert scale questionnaire is a self-reporting, closed-response item (Brown & Rodgers,

2002), which was chosen for a number of reasons. Firstly, Likert scale questionnaires are

popular in language research due to their simplicity, versatility, and reliability plus the fact

that a lot of information can be collected efficiently in a small amount of time (Dörnyei,

2003a). The simplicity of the questionnaire was central to its selection considering that the

participants who took part in this research were likely fatigued after concurrently

completing Task 2 (medium complexity) and Task 3 (high complexity). The Likert scale questionnaire allowed for the collecting of data, with a minimal drain on the participants' already taxed attentional resources, directly after the completion of Task 2 and Task 3.

The Likert scale questionnaire was utilised directly after Task 2 and Task 3 to increase the likelihood that the data elicited was fresh in the short-term memory rather than being taken from the long-term memory. Mackey and Gass (2005) have suggested that data collected from the long-term memory could be negatively influenced by two factors: (a) the distance from the treatment, and (b) the possibility that participants are being influenced by what they perceive the teacher requires.

Secondly, the Likert scale questionnaire collects the type of data appropriate for RQ4. Dörnyei (2003a) categorizes Likert scale questionnaire data as behavioural data, which may represent a participants' actions, their lifestyles and habits, or their personal history. In the field of second language research, this data can represent language learning strategies, and attitudinal data reflecting a number of factors such as participant's attitudes, opinions, beliefs, interests and values.

In this thesis, the Likert scale questionnaire was used to elicit the participants' immediate responses to 11 items that reflected behavioural data concerning the participants' attitudes and opinions. The specific nature of the data concerned the participants' perceptions of the tasks and their perceptions of their performance of the tasks. Both of these perceptions are viewed as being related to issues of task motivation.

The Likert scale questionnaire presented the participants with eleven statements that encouraged the participants to immediately reflect on the treatments that they have just completed. The participants' answers were marked down on a six-point scale. It is Dörnyei's (2003a) contention that using a six-point scale is acceptable and that having even numbers as opposed to odd numbers in the scale does not affect the relative proportions of participants who are actually expressing an opinion, as opposed to those who just decide to choose one of the two middle options.

Furthermore, by assigning the scaled items from strongly disagree to strongly agree, this thesis echoes the work of Dörnyei (2003a), who has worked extensively with questionnaires in the field of second language research. This type of scale is described as a standard set of responses that can be used to measure degrees of agreement or disagreement within the range of target attitudes relevant to this thesis.

3.9 Data collection

In the following section, the procedures used for the application of the letter-writing tasks and the Likert scale questionnaires for RQ1, RQ2, RQ3, and RQ4 are illustrated.

It should be noted that the availability of participants, from the language schools where they studied, was such that all the participants in RQ1, RQ2, and RQ3 performed the letter-writing tasks and the Likert scale questionnaires in groups smaller than the individual sample sizes used for each research question. For example, RQ1 had 34 participants; however, the participants were not put through the data collecting process as one group. RQ1 comprised a number of smaller groups of between 3 and 10 students, depending on how many volunteers were available at any one time. Students were seen in this manner

until the requisite number for RQ1 was met; at which point, the requirements for RQ2 were implemented. The data collection for RQ1, RQ2, and RQ3 took place over the course of eighteen months.

3.9.1 Stage 1 for RQ1, RQ2, and RQ3

The *stage 1* data collection procedures for RQ1, RQ2, and RQ3 were all the same. Participants were approached at their schools and asked to participate in a study on writing. Those who were interested attended a meeting, which also served as stage 1 of the data collecting process. Initially, the participants were given a consent form to complete. After the consent form was completed, the date for stage 2 of the data collecting process was negotiated with the participants. In all cases, stage 2 was performed between three to five days after stage 1.

After the date for stage 2 was set, the participants were given a survey, which was completed in ten minutes. Subsequently, Task 1 (low complexity) was distributed. The participants were given two minutes to read the instructions and ask questions if they were unclear about any aspects of the instructions. The participants were informed that they were not allowed to use dictionaries or smart phones during the performance of the task. Once everyone had signalled that they understood, the task began. Participants had thirty minutes to complete the task, after which the researcher collected it, and the participants were free to leave.

3.9.2 Stage 2 for RQ1

Stage 2 for RQ1 required the application of two consecutive tasks of differing levels of cognitive task complexity (Task 2, medium complexity; and Task 3, high complexity).

116

Firstly, the schedule for stage two was explained to the participants, after which they were then given the first task as well as a 2-minute comprehension check of the task instructions. Clarifications were provided for any participant who had problems understanding the requirements.

Thirty minutes was provided for the completion of the first task. When all the participants had finished, they were allowed a five-minute break (in class) where they could relax but could not discuss the task they had just completed. After the five-minute break, the second task was presented to the participants. As with the first task, there was a comprehension check of the instructions followed by a thirty-minute time limit for completion. It should be noted that Task 2 and Task 3 were counterbalanced to account for participant fatigue. Counterbalancing involved half the participants being randomly selected to perform Task 2 first while the other half performed Task 3 at the same time.

Following the completion of the second task, the participants were provided with a Likert scale questionnaire. After a two-minute comprehension check, ten minutes was provided for the participants to complete the questionnaire. This completed the data collection process for RQ1-stage 2.

3.9.3 Stage 2 for RQ2

The data collecting procedure for RQ2-stage 2 was similar to RQ1-stage 2; however, the extra variable 10 minutes pre-task planning time was included before each task. After the comprehension check of the task instructions was completed for Task 2 and Task 3, the participants were informed that they would have ten minutes in which they could prepare.

They were told that they could take notes, but they could not start writing the letter. After the ten minutes (for both tasks), the participants were given thirty minutes completion time.

3.9.4 Stage 2 for RQ3

RQ3-stage 2 was essentially the same as stage 2 for the previous research questions except for the addition of 10 minutes post-task editing time. After the comprehension check of the task instructions, the participants were informed that on completion of each task, they would be given a red pen. It was explained that the red pen was to be used to edit the work that they had just finished. It was also explained that the editing process was to focus on work already completed, and if the letter was not finished during the thirty minutes provided, the editing component should focus only on the work already done and not completion of the task.

3.9.5 RQ4

Whereas RQ1, RQ2, and RQ3 each comprised different groups of learners, RQ4 utilised all the participants from RQ1, RQ2, and RQ3. The collection of the Likert scale questionnaire data was the same for each group. After all the participants of RQ1, RQ2, and RQ3 had finished their final tasks at the end of stage 2 of the data collecting process, the Likert scale questionnaires where handed out to the participants. After a brief explanation on how to complete the questionnaire, there was a two-minute comprehension check. Following this, the participants were given ten minutes completion time.

3.10 Data analysis

This section covers in detail the analyses of the data from Tasks 1, Tasks 2, Tasks 3 and the Likert scale questionnaires. The section begins with a clarification of the target measures,

clarifying issues related to the coding of T-units as well as presenting examples from the data, which makes clear the two ways in which subordinate clauses were coded and analysed. This is followed by a description of the second target measure, the mean segmental type-token ratio as well as an example of a mean segmental type-token ratio taken from the data. Following this, explanations of the statistical analyses used to calculate the mean scores for the target measures are provided. Finally, the methods by which the Likert-scale questionnaires were analysed are illustrated.

3.10.1 T-units

Hunt's (1965) description of a T-unit, which is an independent clause (at its base level a subject and a finite verb) and any subordinate clauses (adjectival, nominal, adverbial) attached to the main clause, is the T-unit definition used for this research.

The participants involved in the data collection process were all intermediate level learners of English as a second language; therefore, errors were a regular feature of the written texts. As a result, clear definitions of what constituted acceptable levels of error in the T-units needs to be explained. The list below constitutes the guidelines for viable (meaning acceptable for data analysis in this study) error inclusive T-units based on issues that arose during coding. This list was also used to guide the recoding of data for the purpose of interrater reliability.

1. Spelling mistakes are acceptable.
2. A T-unit should have both a subject and a finite verb; however, if the finite verb was the wrong tense, it was still accepted.

3. If the wrong word was used to introduce an adjectival dependent clause, it was accepted and still coded as an adjectival dependent clause.

4. Coordinated T-units, which were separated incorrectly, were still considered separate T-units.

5. T-units that contain incorrect words, but the meaning of the T-unit was still clear and understandable in context, were acceptable.

6. When an independent clause and a dependent clause were incorrectly separated with a period, they were still coded as one T-unit.

7. If an independent clause was lacking a subject or a verb it was not considered viable. However, if the nonviable independent clause had a viable dependent clause that contained a subject and finite verb, the dependent clause was coded as a T-unit without a dependent clause.

8. When only one part of a two-part verb was used, the T-unit was still considered viable.

9. When a T-unit contained sufficient syntactic or lexical errors as to render it incomprehensible, it was not considered viable.

10. When the second clause in a conjunction was without a subject, it was not considered viable.

11. When two independent clauses were joined by an incorrect conjunction, the two clauses were still coded as viable separate T-units.

12. Where commas were omitted between independent clauses, the independent clauses were still considered separate viable T-units.

13. Verb phrases resulting from ellipses were not accepted as viable T-units.

14. Wh-interrogatives such as "How are you" were considered viable T-units. The reason for this exception to the rule was that these types of sentences constitute a grammatically acceptable minimally terminable unit, which fits one description of a T-unit. It just so happens that the nature of this structure precludes a subject.

In Table 13, T-unit variations from the data are illustrated. T-units can have either no dependent clauses or combinations of nominal, adverbial, and adjectival dependent clauses.

Table 13: T-unit examples from the data (dependent clauses in italics)

_____

1. T-unit.	I have been here for three weeks.
2. T-unit with nominal dependent clause.	I think *this city is the best one in New Zealand*.
3. T-unit with adverbial dependent clause.	*When you come here*, I will show you some Kiwi culture.
4. T-unit with adjectival dependent clause.	We can go to a popular bar, *which is quite near to the second restaurant*.

_____

Table 14 below shows increasingly complex T-units. Number 1 is the least complex with no dependent clauses, number 2 is more complex with one dependent clause, and number 3 is the most complex with two dependent clauses.

Table 14: Increasing mean length of T-units from the data

_____

1. T-unit with no dependents clause.	I have been here for three weeks.
2. T-unit with one dependent clause	I think *this city is the best one in New Zealand*.
3. T-unit with multiple dependent clauses.	I was so happy *when I heard that you were coming to New Zealand*.

_____

Two different approaches were taken to analysing the T-units. Initially, the mean length of T-units was analysed across all dependent clause types (adjectival, nominal, and adverbial). This is sometimes referred to in the following chapters as the standard measure. Subsequently, the mean length of T-units was measured with all the dependent clauses being coded and analysed as separate items. This is sometimes referred to in the following chapters as the non-standard measure.

In Table 15 below, different ways the same piece of data might be analysed using standard and a non-standard measures of mean length of T-units are illustrated.

Table 15: Different ways to analyse the same data

Data sample:
I have been here for three weeks. I think this city is the best one in New Zealand. I was so happy when I heard that you were coming to New Zealand.

All dependent clauses analysed as one group. (standard measure)	Sample analyses: 3 T-units and 3 dependent clauses.
All dependent clauses analysed separately. (non-standard measure)	Sample analyses for *adjective clauses*: 3 T-units and 0 dependent clauses. Sample analyses for *nominal clauses*: 3 T-units and 2 dependent clauses. Sample analyses for *adverbial clauses*: 3 T-units and 1 dependent clause.

3.10.2 Mean segmental type-token ratio

The second measure of complexity used in this thesis is a mean segmental type-token ratio (see Table 16 below), which was employed to measure lexical variety while accounting for any problems associated with variations in text length.

Table 16: Mean segmental type-token ratio analysis

_____

Data sample: Segment=40 words

I have been here for three weeks. I think this city is the best one in New Zealand. I was so happy when I heard that you were coming to New Zealand. You will be very happy when you come to New Zealand.

_____

Analysis

Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq
I	4	weeks	1	best	1	happy	2	coming	1
have	1	think	1	one	1	when	2	to	2
been	1	this	1	in	1	heard	1	will	1
here	1	city	1	New Zealand	3	that	1	be	1
for	1	is	1	was	1	you	3	very	1
three	1	the	1	so	1	were	1	come	1

_____

Result

30 (word variations) divided by 40 (words in the segment) equals a type-token ratio of 75%. This process would be repeated based on the number of 40 word segments in a text. The type-token ratios would be added and divided by the number of segments in the text, giving the mean segmental type-token ratio for the text, adjusting for text length.

_____

3.10.3 Summary

In the results sections for RQ1, RQ2, and RQ3 (see chapter 4), the different methods of

analysing the mean length of T-units (see table 13 above), and the analysis of mean

segmental type-token ratio (see table 16 above) are clearly signposted by the italicized

phrases listed below:

1. *Mean length T-units and all dependent clause types* signifies that the relevant

   findings have analysed mean length of T-units by measuring all dependent clauses

   as one inclusive group.

2. *Mean length T-units and adjectival dependent clauses* signifies that the relevant findings have analysed mean length of T-units by measuring only adjectival dependent clauses but no other dependent clauses.

3. *Mean length T-units and nominal dependent clauses* signifies that the relevant findings have analysed mean length of T-units by measuring only nominal dependent clauses but no other dependent clauses.

4. *Mean length T-units and adverbial dependent clauses* signifies that the relevant findings have analysed mean length of T-units by measuring only adverbial dependent clauses but no other dependent clauses.

5. *Mean segmental type-token ratio* signifies that the relevant findings have been analysed with a mean segmental type-token ratio.

3.10.4 Statistical analyses for written complexity

SPSS version 17 was used to analyse the statistical data form the letter-writing tasks and the Likert scale questionnaires. To ensure parity with other research in the field, the commonly used confidence level of .05 ($p<.05$) was employed during analyses.

Most of the analyses of T-unit length were within group, meaning tasks were analysed within the groups that comprised RQ1, RQ2, and RQ3. This was done using a repeated measures ANOVA. Subsequently, a between group analysis of T-unit length was employed using a mixed between-within subjects ANOVA. This compared the overall performance between the groups that comprised RQ1, RQ2, and RQ3.

For the analyses of lexical complexity, a mean segmental type-token ratio was used. The benefit of this particular type-token ratio measure of lexical complexity is the accounting for variations in text lengths in the analyses. He results for the mean segmental type-token ratio were also analysed within groups using a repeated measures ANOVA and between groups using a mixed between-within subjects ANOVA.

The analysis of the mean length of T-units using both standard and non-standard measures required that all dependent clause types within the T-units be noted during coding. This allowed for the analyses of the dependent clauses as both one inclusive group (standard) and also as separate items (non-standard).

To ascertain the mean length of T-units in a standard measure across any one text, the number of dependent clauses and T-units in that text would be added together then divided by the number of T-units. This provided the ratio of T-units to dependent clauses (adverbial, nominal, and adjectival). Subsequently, to ascertain the mean length of T-units using a non-standard measure in the same text, the number of T-units would be added to the number of a particular dependent clause (for example the number of nominal dependent clauses in that text). The combined T-units and nominal dependent clauses total would then be divided by the number of T-units in order to find the ratio of T-units to nominal dependent clauses. This process would then be repeated using adverbial and adjectival dependent clauses.

When the mean length of T-units scores, for both the standard and non-standard measures, were collected from all the letter-writing tasks, the results were subjected to a repeated

measures ANOVA (confidence level .05). This analysis ascertained whether there were any variations in the mean scores between Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity) within each of the groups that comprised RQ1, RQ2, and RQ3.

Multivariate statistics were used in the analyses of all the data because multivariate tests do not require sphericity. Additionally, this thesis reported findings from the statistical analyses using Wilks' Lambda because this is the multivariate test most commonly reported (Pallant, 2002).

Before proceeding with the analyses, the data would be checked for issues of normality. If normality issues arose, the data would be subjected to a Friedman Test, which is a non-parametric analysis that is not subject to issues of normality. The Friedman Test is less sensitive and has less power in the analysis; however, by comparing the results of the Friedman Test to data from the repeated measures ANOVA, the comparison should indicate if normality issues had affected the findings from the repeated measures ANOVA.

When potential normality issues have been addressed and the data has been subjected to a repeated measures ANOVA, the findings would show whether or not significant variations in the mean length of T-units exist between tasks. If significant variations in the mean scores of T-units do exist between tasks, then a Bonferroni adjustment is used to establish where the statistically significant variations are situated between Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity).

As part of the analysis, the Bonferroni adjustment provides a pairwise comparison table that compares scores while highlighting those that are significant. Additionally the Bonferroni adjustment tests at a more stringent confidence level meaning that any scores of significance can be considered robust (Pallant, 2002).

To assess how the mean length of T-unit scores compared between the groups that constituted RQ1, RQ2, and RQ3, a mixed between-within subjects ANOVA was used. A mixed between-within subjects ANOVA utilizes the combined results of the within subjects analyses of variance (in this case the within subjects analysis is the repeated measures ANOVA performed on the data from RQ1, RQ2, and RQ3) and yields a result representing the between subjects effect (the subjects in this case being RQ1, RQ2, and RQ3). As with the repeated measures ANOVA, any significant results for the mixed between-within subjects ANOVA were followed up with a Bonferroni post hoc analysis.

A mean segmental type-token ratio was used to analyse lexical variation in the texts produced by the letter-writing tasks. Problems associated with type-token ratios that fail to acknowledge variations in text length are accounted for in this procedure by dividing the target text into segments of equal length.

Once divided into texts of equal length, each segment is analysed using a type-token ratio, which is the number of types (instances of different lexical types) divided by the number of tokens (total number of words). The more types there are in comparison to the number of tokens, then the greater the lexical variety. An overall mean score for a text is produced by adding all the type-token ratios from each segment, then dividing this number by the entire

number of segments that the text was divided into (Ellis & Barkhuizen, 2005; Yuan & Ellis, 2003). To avoid, as much as possible, wasting data, the texts from the letter-writing tasks were divided into segments of 40 words. This number was based on the size of the smallest sample provided during the data collection process.

Once the results from the mean segmental type-token ratio were gathered, they were analysed using a repeated measures ANOVA (confidence level .05) with a Bonferroni adjustment added to ascertain where any statistically significant differences in the scores might be placed.

3.10.5 Likert scale questionnaires

The Likert scale questionnaire was employed as a means to collect the data used for answering RQ4, which addressed the relationship between the participants' attitudes towards the letter-writing tasks and their syntactic complexity production from RQ1, RQ2, and RQ3.

There were eleven items in the Likert scale questionnaire designed to elicit responses providing some insight into the participant's attitudes towards Task 1, Task 2, and Task 3. Each item on the Likert scale questionnaire had a six-point scale (1=strongly disagree, 2=disagree, 3=slightly agree, 4=partly agree, 5=agree, 6=strongly agree) on which the participants marked the answer that most closely represented their attitude.

The answers from the six-point scale were correlated with the standard measure of mean length of T-units from RQ1, RQ2, and RQ3. A Pearson product-moment correlation was employed with the expectation that it would reveal any relationship between increases and

decreases in the mean length of T-units and the participants' degree of agreement, which ranged from strongly disagree to strongly agree.

Only the Likert items that had a significant correlation ($p<. 50$) across two or three tasks were analysed in detail because the advent of correlations across more than half of the tasks lent strength to the findings.

If normality issues arose, the correlational data would be subjected to a square root transformation, which is used to normalize the data. In addition, a secondary analysis using Spearman's rho would also be used as a means of double-checking the findings. Spearman's rho is similar to a Pearson product-moment correlation but it is a non-parametric analysis and not subject to issues of normality. As a result, the Spearman's rho has less power. Comparing the results from these two analyses with the findings from the Pearson product-moment would show whether any normality issues in the data had resulted in any effect on the findings.

3.10.6 Correlation

A Pearson product-moment correlation coefficient, simple bivariate, was used to determine whether any of the results from the 11 items in the Likert scale questionnaire had a significant positive or negative (see figure 1 and figure 2 below) correlation with the variation in mean length of T-units scores from RQ1, RQ2, and RQ3.

A positive relationship is one where increases in the two scales being correlated increase together, and a negative relationship would be one where the two scales move in different direction, with one increasing while the other decreases. In this thesis the two scales (as

they appear on the scatterplot graphs in chapter 4) represent variations in the mean scores of T-units (on the vertical axis) and the Likert six-point scale (on the horizontal axis).

An example of a positive correlation would be high mean length of T-units scores correlating with high number scores (5 and 6) on the Likert scale while low mean length of T-units scores correlated with low number scores (1 and 2) on the Likert scale.

Figure 1. Perfect positive correlation
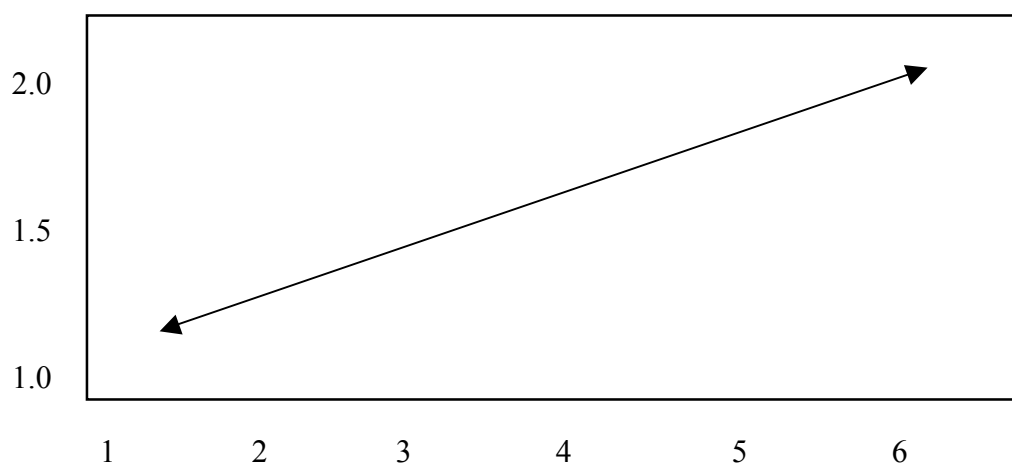
Figure 1 gives an example of a perfect positive correlation between increases in mean length of T-units (represented on the vertical axis) and the six-point Likert scale (represented on the horizontal axis). The arrow used in this figure shows the direction that would indicate a positive relationship in a scatterplot graph. The distribution density of points clustered around the area indicated by the arrow would show the strength of association.

An example of a negative correlation would be high mean length of T-units scores correlating with low number scores (1 and 2) on the Likert scale while low mean length of T-units scores correlated with high number scores (5 and 6) on the Likert scale.

Figure 2. Perfect negative correlation
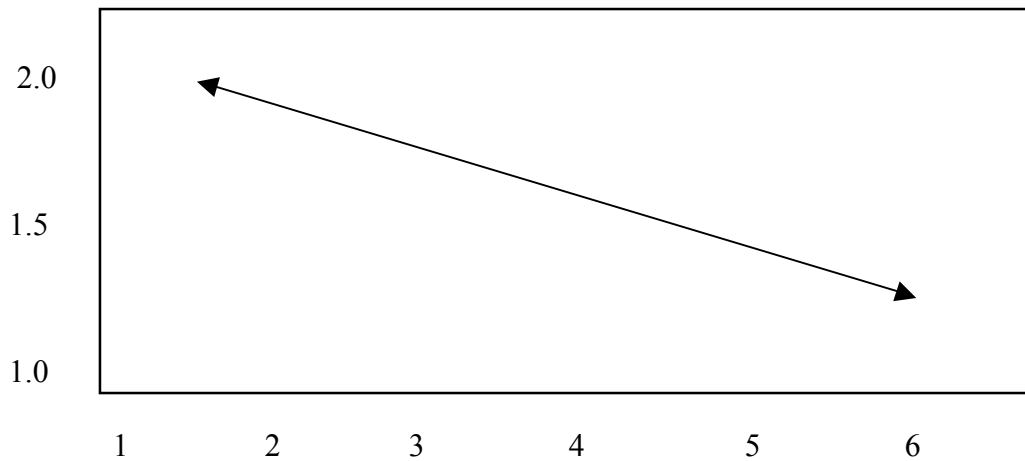
Figure 2 gives an example of a perfect negative correlation between increases in mean length of T-units scores (represented on the vertical axis) and decrease in the six-point Likert answer scale (represented on the horizontal axis). The arrow used in this figure shows the direction that would indicate a negative relationship in a scatterplot graph. The distribution density of points clustered around the area indicated by the arrow would show the strength of association.

3.10.7 Statistical analysis Likert scale questionnaires

Four elements from the Pearson product-moment correlation are reported in the analysis of the data in this thesis. These are SOA (strength of association), participants, the P score (significance), and COD (the coefficient of determination). Examples and explanations of these elements are illustrated below.

In table 17 below, two examples of results from the findings are illustrated, one denoting a positive relationship, and one denoting a negative relationship.

Table 17: Examples of Pearson product-moment correlation results

Items correlated	SOA	Participants	P score	COD
Task 2 and Likert Q8	$r = .255$	$n = 94$	$p = .013$	$R^2 = 0.062$
Task 3 and Likert Q6	$r = -.205$	$n = 94$	$p = .047$	$R^2 = 0.044$

SOA=strength of association. COD= coefficient of determination.

1. Strength of association is signified by r: This figure shows the strength of the relationship between the two variables (mean length of T-units scores and Likert six-point scale). 1.00 is a perfect correlation and 0.00 is no correlation at all. As mentioned above, the relationships can be positive or negative and are denoted as such.

A score of −1 would indicate a perfect negative correlation (much like Figure 2 above if all the scatter points were clustered tightly along the arrow). A score of +1 would indicate a perfect positive correlation (much like Figure 1 above if all the scatter points were clustered along the arrow). A score of 0 would indicate no relationship at all.

A method of grading the strength of the associations between −1 and +1 is required. Pallant (2002) suggests using Cohen's (1998) guidelines. Cohen provided the following scale to interpret the strength of the association: $r = .10$ to .29 is considered a small effect: $r = .30$ to .49 is considered a medium effect: $r = .50$ to 1.0 is considered a large effect.

2. Participants are signified by n: This letter illustrates the number of participants in the analysis. For RQ4, the number is 94, which is the total number of participants from RQ1 (34), RQ2 (30), and RQ3 (30).

3. Significance scores is signified by p: This letter illustrates whether the relationship between the two variables is statistically significant or not. This score was used to decide whether or not further analysis was needed for any of the correlations.

4. Coefficient of determination is signified as R^2: The coefficient of determination provides a number (reported as a per cent e.g. $R^2 = 0.062$ would be 6.2 %) that shows the shared variance between the two variables (as one variable changes so does the other). This makes the interpretation of the correlation easier by showing the extent to which the variables, as sets of numbers, vary together (Brown, 2003). Simply, the per cent represents the amount of the result that can be explained by the analysis (how much overlap there is between the two variables) and gives a clearer indication of how meaningful the result is.

3.10.8 Validity and Reliability

Throughout this thesis, attempts have been made to establish the validity and reliability of the research. Validity, which can broadly be described as the degree to which the research accurately addresses the concepts described in the research questions, is discussed first. Subsequently the issue of reliability, which is the extent to which the means used to test and measure the concepts in the research question can be replicated in other studies, is discussed.

To safeguard validity, this thesis was designed to ensure as much as possible that the results accurately reflected variations in cognitive task complexity and not other factors. For example, students were chosen who were considered to be of the appropriate proficiency level. In the literature review, it is stated that the analysis of dependent clauses is best suited to students who are at an intermediate level of proficiency. As a result, the participants were selected on the basis that each school provided intermediate level students, who were all supposed to have placed within an IELTS levels ranging from 4.5 to 5.5, which is supposed to represent skills commensurate with the requirements of this thesis.

As an extra way of ensuring that the students had a similar skill levels, Task 1 (which is the low complexity task) was also used as a gauge of how closely the participants performed in terms of mean length of T-units and mean segmental type-token ratio. A mixed between-within subjects ANOVA was used that enabled the observance of mean score variations between groups of participants who performed all tasks including Task 1.

A further method employed to ensure that the variables being testes were responsible for variations in the output was the use of counterbalancing during the data collecting process. Counterbalancing was used to rule out fatigue as a factor affecting output. This was important considering that the output was already being produced under cognitive stress and that during stage two of the data collecting procedure, two tasks were performed concurrently.

A pilot study was also performed. This allowed for the researcher to test the instruments on a number of participants who were supposed to be at the same level as the participants who

were part of the main study. The pilot study allowed for any necessary revising of the instruments; for example, the moderation of the amount of time needed to adequately perform the tasks and the adjustment of the task instructions to ensure that students were able to understand what was required of them. Additionally, the pilot study proved that the tasks produced the type of language samples needed for this thesis.

Furthermore, both the independent and dependent variables used in this thesis have been clearly investigated in the literature review. Prior research that explains the theoretical underpinnings of cognitive task complexity and pre-task planning time are reviewed, in addition to prior studies that have used these variables. The additional variable of post-task planning time was addressed; however, there was little literature on this subject. Additionally, the dependent variables were also reported in the methodology section, with extra reference in the literature review given to explaining T-units and subordination; moreover, explanations for the use of the new non-standard measure of subordination were also provided.

To contribute to the external validity of the results, which is the extent to which the findings can be generalized to a wider population than the sample group (Mackey & Gass, 2005), the participants were chosen from a range of language schools that were representative of the language schools in Auckland. Additionally, the participants chosen from the schools were from a variety of ethnicities, which represented the types of participants studying English in New Zealand and abroad. The survey completed by each student appeared to show that there was nothing about theses participants that made then noticeably different from each other or the group that they mostly represented.

In order to contribute to the reliability of this thesis, a number of steps were taken. Firstly, the types of tasks used in this study were letter-writing tasks similar to the ones used in prior research, this contributed to the validity of the research, but also the reliability by using task types that are easily replicated. Additionally, extensive details were provided in the methodology section explaining the design of the tasks as well as the way in which the data collecting process was performed.

To contribute to the reliability of this thesis, one of the PhD candidates from AUT University participated with the recoding of the data. This is referred to as interrrater reliability. Approximately eight months after the data was originally coded, the PhD candidate was provided with 20 per cent of the data to recode.

The target measures (mean length of T-units and mean segmental type-token ratio) were clearly explained to the PhD candidate; additionally, the list of potential problems and their solutions, which arose during the original coding, were also provided. These issues can be found in the methodology section.

The results of the recoding were subjected to a Pearson product-moment correlation, which Mackey and Gass (2005) state is one way to calculate interrater reliability. This shows the degree of association between both the original and the recoded data. For both mean length of T-units and mean segmental type-token ratio, the Pearson product-moment correlation returned positive correlations between the original data and recoded data. The mean length of T-units (for 20% of the participants) had a significant ($p=$. 000) positive correlation with a strength of association of .92, which is considered a large effect. The results for mean

segmental type-token ratio (for 20% of the participants) had a significant (p=. 000) positive

correlation with a strength of association of .96, which is considered a large effect.

After viewing the results for mean length of T-units, the recoded data was reviewed to

gauge why the result was not higher than .92. After a detailed evaluation of the data, it was

found that in some cases the PhD student who had recoded the data had not always

followed the instructions regarding the list of potential coding problems supplied from the

methodology section. These omissions were likely contributors to the final results.

CHAPTER FOUR

RESULTS

4.1 Overview

This chapter presents analyses of the data used to answer the four research questions posited in the previous chapter. The first section addresses research questions 1 (RQ1), research question 2 (RQ2), and research question 3 (RQ3), which use qualitative data that is part of the experimental dimension of the research (investigating causality between variables). The subsequent section addresses research question 4 (RQ4), using quantitative data for the associational dimension of the research (exploring the strength of association between different variables).

The analyses are listed by research question, starting with RQ1, which focuses solely on increases in task complexity. This is followed by RQ2, which combines increases in task complexity with 10 minutes pre-task planning time; and finally, RQ3, which combines increases in task complexity with 10 minutes post-task editing time.

For RQ1, RQ2, and RQ3, the three letter-writing tasks, which provide the bulk of the data for this research, are the foci of the analyses. Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity) each contain differing levels of cognitive task complexity, operationalized as increases in reasoning demands and the number of elements (see section 3.8.2 on letter-writing tasks). Each tasks' output is analysed for written complexity, specifically mean length of T-units and mean segmental type-token ratio (see Tables 13, 15, and 16).

RQ4 requires a quantitative/associational analysis of the data. The data for RQ4 was gathered using Likert scale questionnaires, which elicited data in the form of participants' immediate responses to two aspects: (a) the tasks themselves and, (b) their performance of the tasks. Each question from the questionnaires was coded and analysed so as to provide mean scores that illustrated where the respondents' answers fell on the six-point scale. Subsequently, this data is correlated with the mean length of T-unit data from RQ1, RQ2, and RQ3.

4.2 The results for RQ1, RQ2, and RQ3

The following analyses are of the three sets of data that were collected to answer the research questions regarding cognitive task complexity's effect on written complexity under different conditions. The first condition is one where there is no pre-task planning time added to the performance of the tasks. The second condition is one where 10 minutes pre-task planning time was added as an extra variable during the completion of the two complex tasks (Task 2 and Task 3), and the third condition was the introduction of 10 minutes post-task editing time added to the performance of the two complex tasks (Task 2 and Task 3).

The data was initially analysed using a repeated measures ANOVA. During the initial analyses, the data was checked for normality by reviewing the mean length of T-units results (from Task 1, Task 2, and Task 3 for all 94 participants) as histograms, which show distribution. The histograms revealed slight positive skews in the distribution.

Because of the slight skew found in the results, a Friedman Test (a non-parametric analysis which is not subject to issues of normality but is less sensitive and has less power in the

analysis) was applied to the data to investigate if the removal of normality as a factor affected the results. The results from the Friedman Test were not different enough to be of concern, thus the normality issues noticed in the initial analyses appeared to have a minimal impact on the results. Subsequently, the following analyses are presented using the initial repeated measures ANOVA data.

4.2.1 RQ1: What are the effects of cognitive task complexity on written complexity?

In the following sections, the results for RQ1 are presented starting with the findings for the mean length of T-units measured across all dependent clause types. This is followed by the individual results for adjectival, nominal, and adverbial dependent clauses. After a summary of the results for individual dependent clauses, the findings for lexical complexity, which used a mean segmental type-token ratio, are presented.

4.2.2 Research question 1, part 1: Syntactic complexity

RQ1 investigated whether manipulating elements in the design of writing tasks have any effect on the written outcomes. Divided into three parts, RQ1 part 1 investigated the effects of cognitive task complexity on mean length of T-units, measured across all dependent clause types (adjectival, nominal, and adverbial).

Analysing the mean length of T-units (see Tables 13, 14, and 15 in the previous chapter) required calculating the ratio of dependent clauses to independent clauses for each of the participants' three written texts (Task 1: low complexity, Task 2: medium complexity, Task 3: high complexity).

To find the mean length of T-units for RQ1-part 1, the number of T-units was added to the number of dependent clauses. Subsequently, this number was divided by the number of T-units to ascertain the mean length of dependent clauses. For each participant, this produced three mean scores that revealed variation in the mean length of T-units among the tasks.

Table 18 below shows the mean length and standard deviation of T-units, in which all dependent clause types were analysed for the 34 participants of RQ1, for each of the three tasks (see section 3.8.2 for the defining characteristics of each task type).

Table 18: Mean length T-units and all dependent clause types

Level of complexity	Mean	Standard deviation
Task 1: Low	1.41	0.20
Task 2: Medium	1.40	0.15
Task 3: High	1.35	0.14

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity on the mean length of T-units (across all dependent clause types) for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p > .05$), Wilks' Lambda $\Lambda = .84$, $F(2, 32) = 3.02$, $p = .06$, multivariate partial eta squared $= .159$.

There were no statistically significant mean scores variations among Task1, Task 2, and

Task 3 when testing for increases in cognitive task complexity on the mean length of T-

units across all dependent clause types.

Figure 3. Mean length T-units and all dependent clause types

Figure 3 illustrates the results for all dependent clauses with the horizontal axis showing

Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity), and

the vertical axis showing the mean length of T-units measured across all dependent clause

types (adjectival, nominal, and adverbial).

Though there was no significant mean scores variation among Task1, Task 2, and Task 3,

when testing for increases in task complexity on mean length of T-units across all

dependent clause types, the *p* value (*p*=. 06) was considered close enough to the .05 level of

significance that a more detailed analysis using a post hoc test was used to ensure that no significant result might be missed when using a more detailed analysis. Subsequently, a Bonferroni adjustment was performed on the findings; however, the post hoc test revealed no significant mean scores variations between Task 1 (low complexity) Task 2 (medium complexity) and Task 3 (high complexity).

In Figure 3 above, the largest drop in the mean length of T-units can be seen between Task 1 and Task 3; however, the difference is not statistically significant. Thus there is no significant effect for cognitive task complexity on the mean length of T-units measured across all dependent clause types.

4.2.3 Research question 1, part 2

RQ1-part 2 focuses on a more detailed analysis of complexity and the mean length of T-units by examining whether cognitive task complexity had any statistically significant effects on the individual types of dependent clauses. During the coding process for RQ1, three types of dependent clause were identified and separately coded; these were adjectival dependent clauses, nominal dependent clauses, and adverbial dependent clauses (see Table 13). The results for each dependent clause are presented separately in the following sections.

4.2.3.1 Adjectival dependent clauses

The mean length of adjectival dependent clauses was gauged by calculating the ratio of adjectival dependent clauses to independent clauses for each participants' text from Task 1, Task 2, and Task 3. Firstly the number of T-units was added to the number of adjectival dependent clauses. Subsequently, this number was divided by the number of T-units to

ascertain the mean length of adjectival dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of T-units measuring only adjectival dependent clauses among the tasks.

Table 19 below shows the mean length and standard deviation of T-units, in which only adjectival dependent clauses were analysed for the 34 participants of RQ1, for each of the three tasks.

Table 19: Mean length T-units and adjectival dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.05	0.06
Task 2: Medium	1.07	0.08
Task 3: High	1.07	0.08

A repeated measures ANOVA was conducted on the means to investigate the effect of task complexity on the mean length of T-units measuring adjectival dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda = .95$, F (2, 32) $=. 70$, $p= .50$, multivariate partial eta squared $= .042$.

No statistically significant mean scores variation was found among Task 1, Task 2, and Task 3 attributable to increases in task complexity on the mean length of T-units measuring adjectival dependent clauses.

Figure 4. Mean length T-units and adjectival dependent clauses

Figure 4 illustrates the results for adjectival dependent clauses with the horizontal axis

showing Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high

complexity), and the vertical axis showing the mean length of T-units for adjectival

dependent clauses.

Figure 4 above shows a statistically non-significant increase in the mean length of T-units

measuring adjectival dependent clauses. A close examination of the measurements on the

vertical axis reveals the increases to be exceedingly small, thus illustrating a minimal effect

for cognitive task complexity on the mean length of T-units measuring adjectival dependent

clauses.

4.2.3.2 Nominal dependent clauses

The mean length of nominal dependent clauses was found by calculating the ratio of nominal dependent clauses to independent clauses for Task 1, Task 2, and Task 3. As with the adjectival dependent clauses, the ratio of nominal dependent clauses to independent clauses was found by adding the number of T-units to the number of nominal dependent clauses. Subsequently, this number was divided by the number of T-units to ascertain the mean length of nominal dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of T-units measuring nominal dependent clauses among the tasks.

Table 20 below shows the mean length and standard deviation of T-units, in which only nominal dependent clauses were analysed for the 34 participants of RQ1, for each of the three tasks.

Table 20: Mean length T-units and nominal dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.18	0.10
Task 2: Medium	1.20	0.10
Task 3: High	1.17	0.08

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity on the mean length of T-units measuring nominal dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3

(high complexity). The result was not statistically significant (p>.05), Wilks' Lambda Λ = .87, F (2, 32) =2.21, p= .12, multivariate partial eta squared = .122.

There was no significant mean scores variation among Task 1, Task 2, and Task 3 when testing for increases in cognitive task complexity on the mean length of T-units measuring nominal dependent clauses, thus no post hoc test was performed.

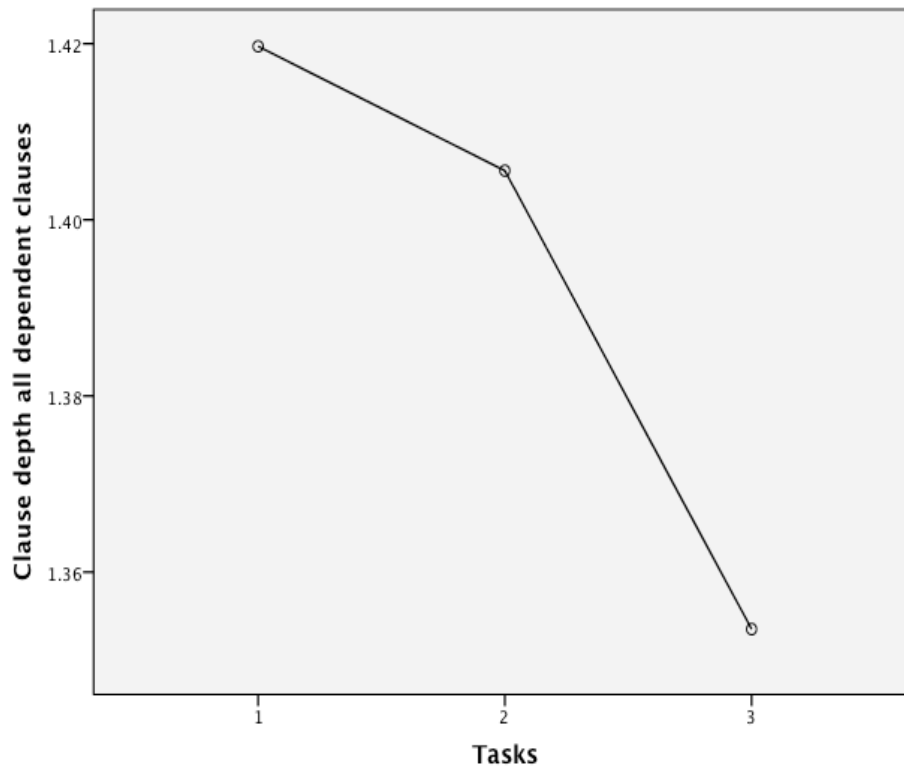**Figure 5. Mean length T-units and nominal dependent clauses**
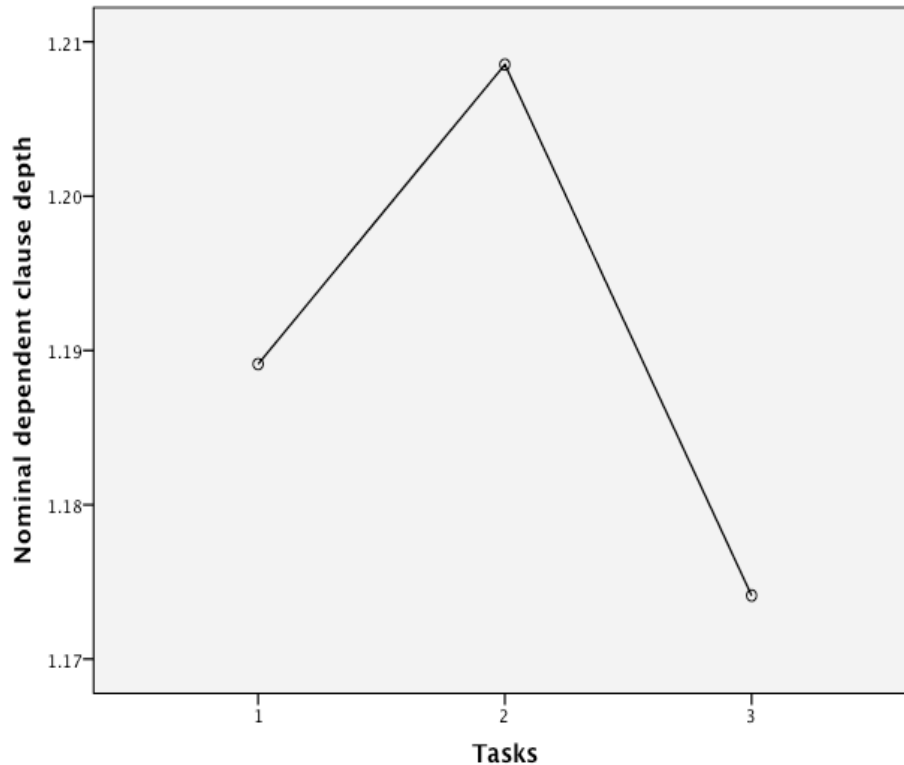
Figure 5 illustrates the results for nominal dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity), and the vertical axis showing the mean length of T-units for nominal dependent clauses.

Figure 5 above shows the non-significant variation in nominal dependent clauses, with a non-significant increase in T-units measuring nominal dependent clauses from Task 1 to Task 2, followed by a non-significant drop in mean nominal dependent clauses between Task 2 and Task 3.

4.2.3.3 Adverbial dependent clauses

The mean length of adverbial dependent clauses was found by calculating the ratio of adverbial dependent clauses to independent clauses for Task 1, Task 2, and Task 3. The ratio of adverbial dependent clauses to independent clauses was calculated by adding the number of T-units to the number of adverbial dependent clauses. Subsequently, this number was divided by the number of T-units to ascertain the mean length of adverbial dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of T-units measuring adverbial dependent clauses among the tasks.

Table 21 below shows the mean length and standard deviation of T-units, in which only adverbial dependent clauses were analysed, for the 34 participants of RQ1, for each of the three tasks.

Table 21: Mean length T-units and adverbial dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.16	0.11
Task 2: Medium	1.10	0.08
Task 3: High	1.09	0.08

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity on the mean length of T-units measuring nominal dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was statistically significant (p< .05), Wilks' Lambda Λ = .77, F (2, 32) =4.72, p= .016, multivariate partial eta squared = .228.

When a significant result is found in the initial analysis, it is prudent to consider the effect size (Pallant, 2002). Effect size indicates the strength of association between the independent variable (cognitive task complexity) and the dependent variable (in this case, the mean length of T-units measuring adverbial dependent clauses). This indicates that the result is not based on the size of the sample but is based on the relationship between the two variables (Pallant, 2002). As with the other multivariate results, effect size was taken from Wilks' Lambda. The appropriate value was found in the column labelled Partial Eta Squared. The effect size is .228 (which is considered a strong effect) meaning that the result is likely a product of the relationship between the two variables.

The results from the Bonferroni adjustment post hoc test revealed a statistically significant mean scores variation (*p*= .023) between Task 1 (low complexity) and Task 2 (medium complexity) and a significant mean scores variation (*p*= .020) between Task 1 (low complexity and Task 3 (high complexity). This indicates that Task 1, which has low task complexity, has a significantly higher incident of mean adverbial dependent clauses per T-unit than Task 2 and Task 3, which both have less complexity. Additionally, there appears to be no significant change in mean adverbial dependent clauses attributable to variations in task complexity between Task 2 and Task 3.

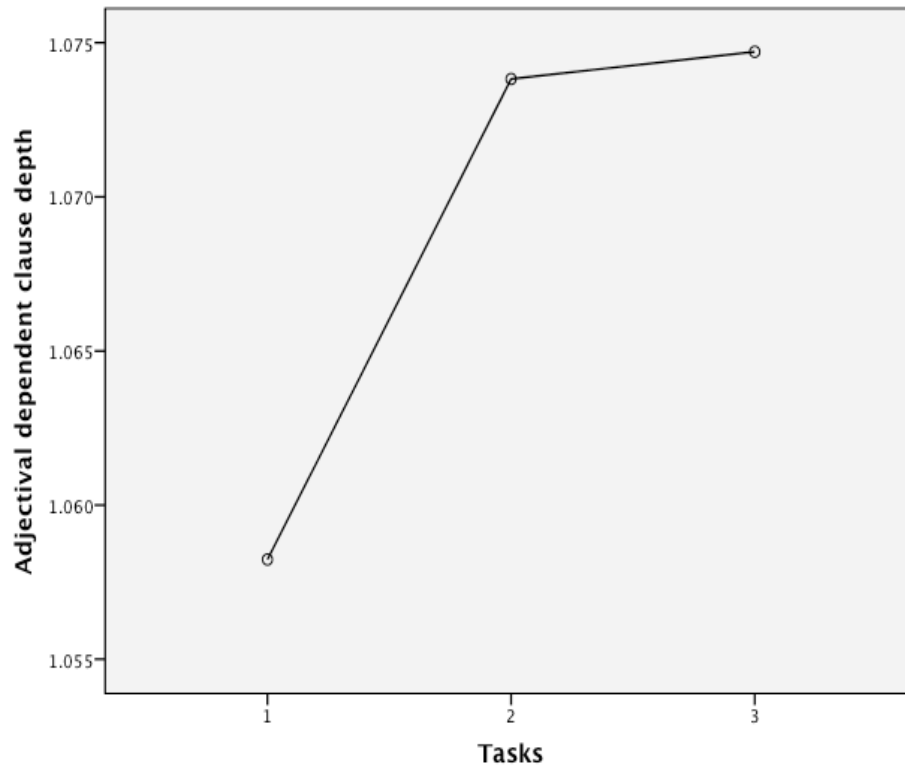**Figure 6. Mean length T-units and adverbial dependent clauses**
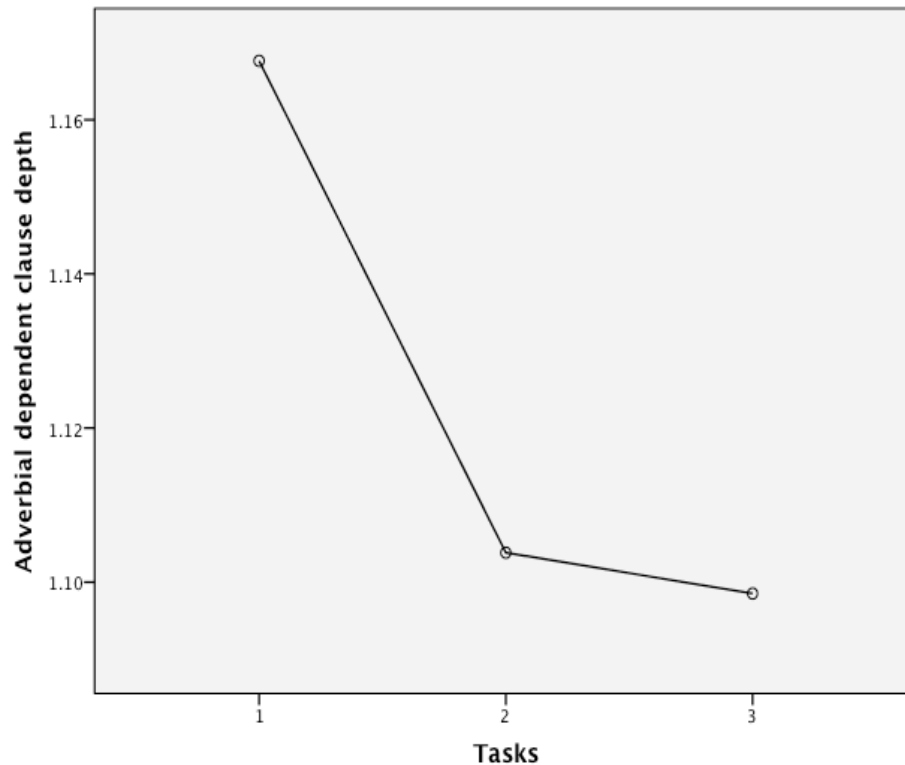
Figure 6 illustrates the results for adverbial dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity), and the vertical axis showing the mean length of T-units measuring adverbial dependent clauses.

In Figure 6 above, the statistically significant decrease in the mean length of T-units measuring adverbial dependent clause can be seen between Task 1 and both Tasks 2 and 3; however, no significant variation is noted between Task 2 and Task 3.

4.2.3.4 Summary of research question1, part 2

RQ1-part 2 has revealed that variations in the level of task complexity appear to have different effects on each dependent clause. In short, making the tasks more complex resulted in a statistically significant shortening of only the adverbial dependent clauses, with variations in mean nominal dependent clauses not reaching significance.

There was a small increase in mean adjectival dependent clauses (mean length of T-units in which only adjectival dependent clauses were analysed) among Task 1, Task 2, and Task 3; however, these increments were extremely small and not statistically significant ($p > .05$).

The mean length of T-units measuring nominal dependent clauses increased from Task 1 (low complexity) to Task 2 (medium complexity); however, the difference was not statistically significant ($p > .05$). There was a drop in mean nominal dependent clauses, with Task 3 (high complexity) having lower mean nominal dependent clauses than both Task 1 (low complexity) and Task 2 (medium complexity). The biggest drop in mean nominal dependent clauses was from Task 2 (medium complexity) to Task 3 (high complexity); however, none of the drops in complexity were statistically significant ($p > .05$).

There was no effect for task complexity on mean adverbial dependent clauses (the mean length of T-units in which only adverbial dependent clauses were analysed) between Task 2 (medium complexity) and Task 3 (high complexity). However, Task 1 (low complexity) had statistically higher mean adverbial dependent clauses than both Task 2 (medium complexity) and Task 3 (high complexity).

4.2.4 Research question 1, part 3: Lexical complexity

RQ1-part 3 investigated whether increases in cognitive task complexity in the task design had any statistically significant effects on the lexical variety (the variety of different words used in a text) of the written outcomes. Lexical variety was measured by using a mean segmental type-token ratio, which is used to discover the lexical variation in a piece of writing while taking into account the effect of text length.

To calculate the mean segmental type-token ratio, each text was divided into segments of equal length, which were analysed with a type-token ratio (the number of lexical types divided by the total number of words). Subsequently, the overall mean score for each text was found by adding the type-token ratios from each segment, then dividing this number by the entire number of segments in the text. For each participant, this produced three mean scores expected to show any instance of variation in word variety among the tasks.

Table 22 below shows the mean segmental type-token ratio and standard deviation for the 34 participants of RQ1 for each task.

Table 22: Mean segmental type-token ratio

Level of complexity	Mean	Standard deviation
Task 1: Low	82.00	3.53
Task 2: Medium	82.82	4.00
Task 3: High	83.81	3.52

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity on lexical variety for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was statistically significant ($p<.05$), Wilks' Lambda $\Lambda = .82$, $F (2, 32) = 3.40$, $p = .046$, multivariate partial eta squared $= .176$.

This result shows significant mean scores variation ($p = .046$) among Task 1, Task 2, and Task 3 for an increase in task complexity on the lexical variety of the written output. In addition, the effect size is .176, which is considered a strong effect. A post hoc analysis (Bonferroni adjustment) revealed a statistically significant difference ($p = .039$) between Task 1 (low complexity) and Task 3 (high complexity).

These findings revealed incremental increases in lexical variety as the tasks became more complex, with Task 1 having the least lexical variety, Task 2 having the medium amount of lexical variety, and Task 3 showing the most lexical variety. The increase in word variety between Task 1 and Task 2 was not statistically significant and the increase between Task 2 and Task 3 was also not significant. However, the increase between Task 1 (low complexity) and Task 3 (high complexity) was statistically significant ($p = .039$).

Figure 7 below shows increases in lexical variety as a result of increases in task complexity; however, the only statistically significant increase is between Task 1 and Task 3.

Figure 7. Mean segmental type-token ratio

Figure 7 illustrates the results for the mean segmental type-token ratio with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity), and the vertical axis showing variations in lexical variety.

4.2.5 RQ2: What are the effects of pre-task planning time combined with cognitive task complexity on written complexity?

In the following sections, the results for RQ2 are presented starting with the findings for syntactic complexity and the mean length of T-units measured across all dependent clause types. This is followed by the individual results for adjectival, nominal, and adverbial dependent clauses. After a summary of the results for individual dependent clauses, the findings for lexical complexity, which used a mean segmental type-token ratio, are presented.

4.2.6 Research question 2, part 1: Syntactic complexity

RQ2 investigated whether manipulating elements in the design of writing tasks in conjunction with pre-task planning time had any effect on the written outcomes. Divided into three parts, RQ2-part 1 investigated the effects of task complexity and 10 minutes pre-task planning time on the mean length of T-units, measured across all dependent clause types (adjectival, nominal, and adverbial).

Analysing the mean length of T-units for RQ2 was achieved in the same way as RQ1, by calculating the ratio of all dependent clauses (nominal, adverbial, and, adjective) to independent clauses for Task 1, Task 2, and Task 3. Initially, the number of T-units was added to the number of dependent clauses. Subsequently, this number was divided by the number of T-units to ascertain the mean length of dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of T-units among each task.

Table 23 below shows the mean length and standard deviation of T-units, in which all dependent clause types were analysed for the 30 participants of RQ2, for each of the three tasks.

Table 23: Mean length T-units and all dependent clause types

_____

Level of complexity	Mean	Standard deviation
Task 1: Low complexity	1.39	0.20
Task 2: Medium complexity	1.45	0.16
Task 3: High complexity	1.44	0.22

_____

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity and 10 minutes pre-task planning time on the mean length of T-units (across all dependent clause types) for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda = .93$, $F (2, 28) =1.00$, $p= .37$, multivariate partial eta squared = .067.

There were no statistically significant variations in the mean length of T-units among Task 1, Task 2, and Task 3 when testing for increases in cognitive task complexity plus 10 minutes pre-task planning time.

Figure 8 below shows an increase in the mean length of T-units between Task1 and Task 2; however, the increase is not statistically significant.

Figure 8. Mean length T-units and all dependent clause types

Figure 8 illustrates the results for all dependent clauses with the horizontal axis showing

Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes pre-task planning

time), and Task 3 (high complexity plus 10 minutes pre-task planning time), and the

vertical axis showing the mean length of T-units measured across all dependent clause

types (adjectival, nominal, and adverbial).

4.2.7 Research question 2, part 2

This part of RQ2 focuses on a more detailed analysis of complexity, pre-task planning time,

and the mean length of T-units by examining whether variations in cognitive task

complexity and the addition of 10 minutes pre-task planning time had any effect on the

individual types of dependent clauses produced across Task 1, Task 2, and Task 3. As with

RQ1, the coding process for RQ2 involved identifying and separately coding the three types of dependent clause. These were adjectival dependent clauses, nominal dependent clauses, and adverbial dependent clauses. The results for each dependent clause are presented separately in the following sections.

4.2.7.1 Adjectival dependent clauses

As with the adjectival dependent clause analysis in RQ1, the ratio of adjectival dependent clauses to independent clauses was calculated for each participants' text from Task 1, Task 2, and Task 3. Initially, the number of T-units was added to the number of adjectival dependent clauses. Subsequently, this number was divided by the number of T-units to determine the mean length of adjectival dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of T-units measuring adjectival dependent clauses among each task.

Table 24 below shows the mean length and standard deviation of T-units, in which only adjectival dependent clauses were analysed for the 30 participants of RQ2, for each of the three tasks.

Table 24: Mean length T-units and adjectival dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.07	0.09
Task 2: Medium	1.09	0.09
Task 3: High	1.12	0.11

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity and 10 minutes pre-task planning time on the mean length of T-units measuring adjectival dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was statistically significant ($p<.05$), Wilks' Lambda $\Lambda =.79$, $F (2, 28) =3.71$, $p=.037$, multivariate partial eta squared $=.210$.

The statically significant result, found among the mean scores of Task 1, Task 2, and Task 3, may be attributable to increases in cognitive task complexity in conjunction with 10 minutes pre-task planning time on mean adjectival dependent clauses. The effect size is .210 (which is considered a strong effect), meaning that the result is likely a product of the relationship between the independent and dependent variables.

The results from the Bonferroni adjustment post hoc test revealed a statistically significant mean scores variation ($p=.050$) between Task 1 (low complexity) and Task 3 (high complexity). This indicates that Task 1, which has low task complexity and no pre-task planning time, has a significantly lower mean adjectival dependent clause than Task 3, which has high task complexity and 10 minutes pre-task planning time. However, there appears to be no statistically significant variation in the mean length of T-units measuring adjectival dependent clauses attributable to variations in task complexity and 10 minutes pre-task planning time between Task 1 and Task 2, and between Task 2 and Task 3.

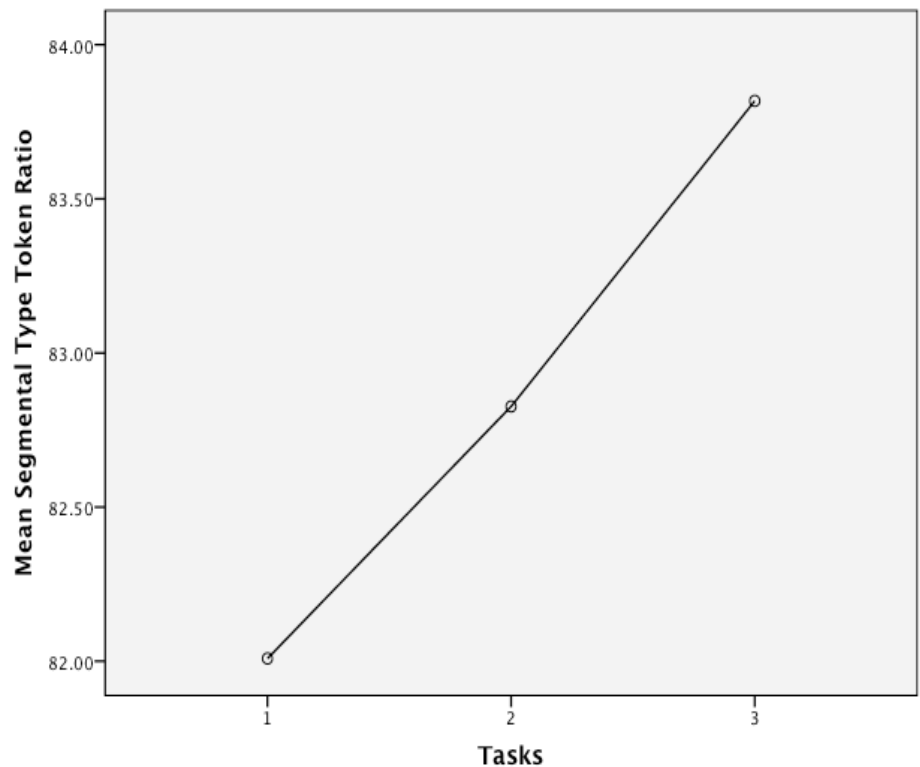**Figure 9. Mean length T-units and adjectival dependent clauses**
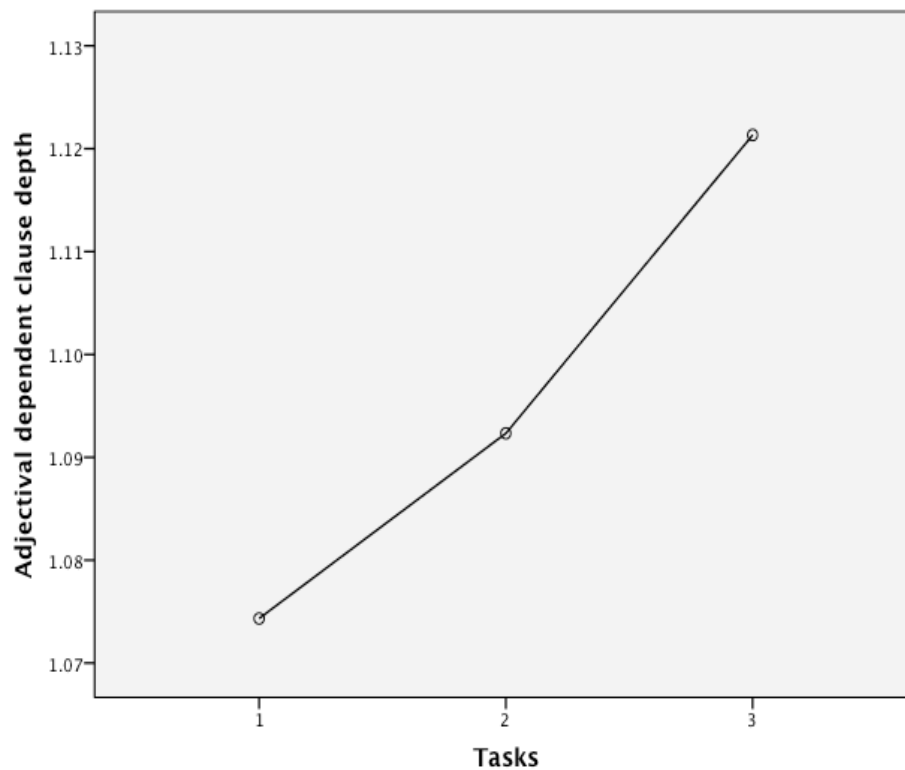
Figure 9 illustrates the results for adjectival dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes pre-task planning time), and Task 3 (high complexity plus 10 minutes pre-task planning time), and the vertical axis showing the mean length of T-units measuring adjectival dependent clauses.

Figure 9 above shows increases in the mean length of adjectival dependent clauses; however, the only statistically significant increase is between Task 1 and Task 3.

4.2.7.2 Nominal dependent clauses

The mean length of T-units measuring nominal dependent clauses was found by calculating the ratio of nominal dependent clauses to dependent clauses for Task 1, Task 2, and Task 3.

The ratio of nominal dependent clauses to independent clauses was found by adding the number of T-units to the number of nominal dependent clauses, then dividing this number by the number of T-units to ascertain the mean length of nominal dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of nominal dependent clauses among each task.

Table 25 below shows the mean length and standard deviation of T-units, in which only nominal dependent clauses were analysed for the 30 participants of RQ2, for each of the three tasks.

Table 25: Mean length T-units and nominal dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.16	0.07
Task 2: Medium	1.18	0.07
Task 3: High	1.16	0.08

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity and 10 minutes pre-task planning time on the mean length of T-units measuring nominal dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda =. 91$, $F (2, 28) =1.32$, $p=. 28$, multivariate partial eta squared $=. 087$.

No statistically significant mean scores variation was found among Task 1, Task 2, and

Task 3, attributable to increases in cognitive task complexity plus 10 minutes pre-task

planning time on the mean length of nominal dependent clauses.

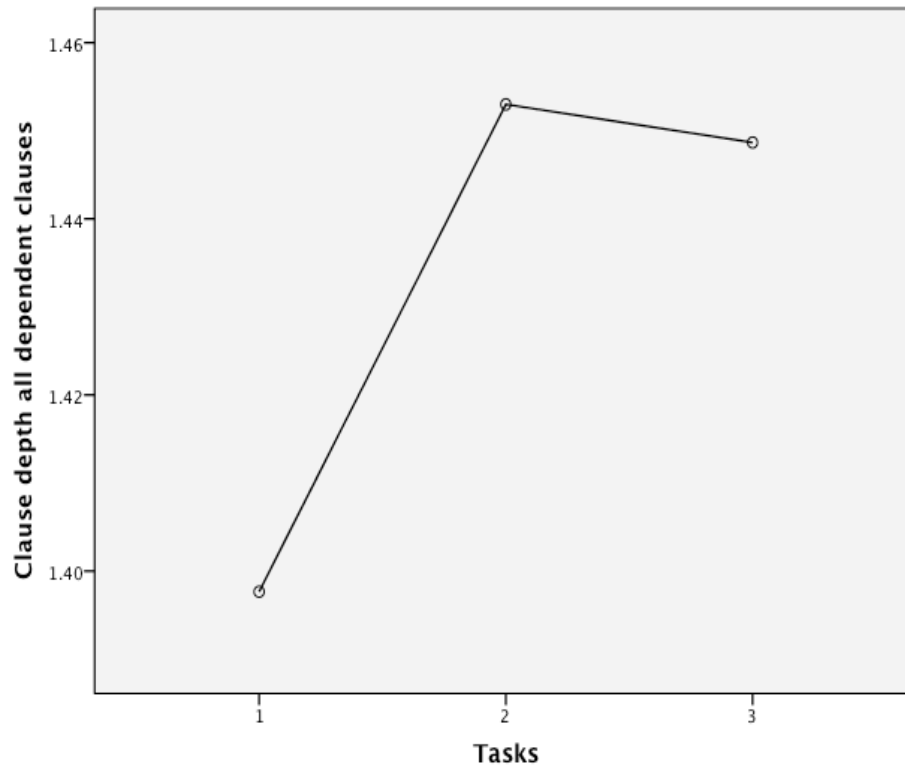**Figure 10. Mean length T-units and nominal dependent clauses**
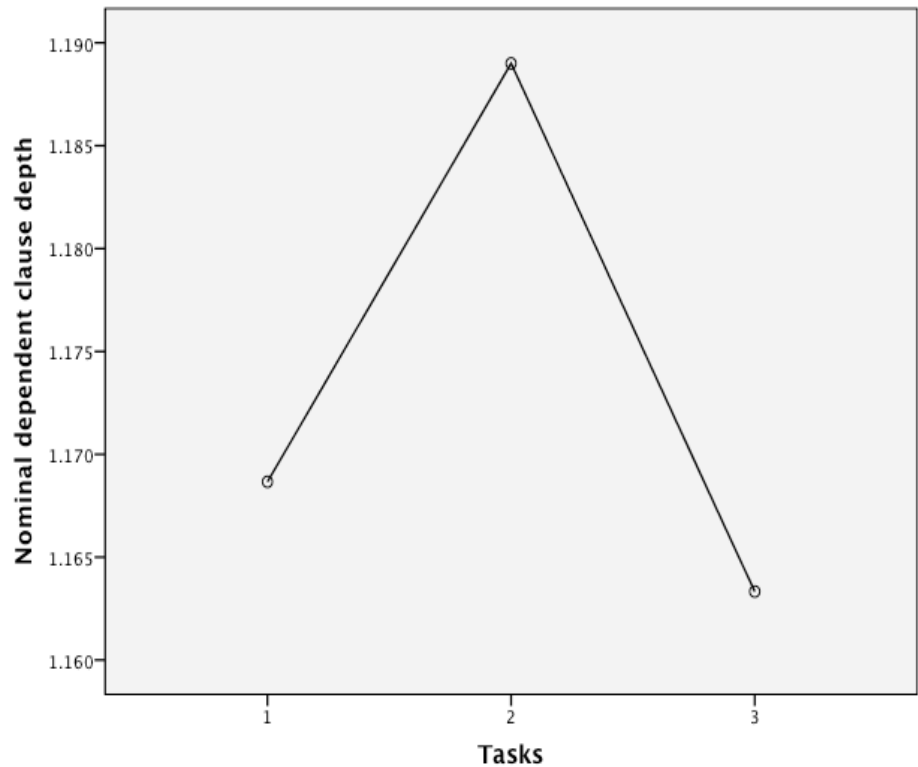
Figure 10 illustrates the results for nominal dependent clauses with the horizontal axis

showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes pre-task

planning time), and Task 3 (high complexity plus 10 minutes pre-task planning time), and

the vertical axis showing the mean length of T-units measuring nominal dependent clauses.

Figure 10 above shows increases and decreases in the mean length of nominal dependent clauses; however, a close look at the vertical axis reveals that these variations are small and not statistically significant.

4.2.7.3 Adverbial dependent clauses

The mean length of T-units measuring adverbial dependent clauses was found by calculating the ratio of adverbial dependent clauses to independent clauses for Task 1, Task 2, and Task 3. This ratio was found by adding the number of T-units to the number of adverbial dependent clauses. Subsequently, this number was divided by the number of T-units to determine the mean length of adverbial dependent clauses. For each participant, this produced three mean scores, designed to show instances of variation in the mean length of adverbial dependent clauses among each task.

Table 26 below shows the mean length and standard deviation of T-units, in which only adverbial dependent clauses were analysed for the 30 participants of RQ2, for each of the three tasks.

Table 26: Mean length T-units for adverbial dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.17	0.12
Task 2: Medium	1.15	0.07
Task 3: High	1.14	0.09

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity and 10 minutes pre-task planning time on the mean length of T-units measuring adverbial dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant (p>.05), Wilks' Lambda $\Lambda = .97$, F (2, 28) =. 36, p=. 69, multivariate partial eta squared =. 026.

No significant mean scores variation was found among Tasks 1, 2, and 3, attributable to increases in cognitive task complexity plus 10 minutes pre-task planning time.

Figure 11. Mean length T-units for adverbial dependent clauses

Figure 11 illustrates the results for adverbial dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes pre-task planning time), and Task 3 (high complexity plus 10 minutes pre-task planning time), and the vertical axis showing the mean length of T-units measuring adverbial dependent clauses.

In Figure 11 above, the mean length of adverbial clauses appears to decrease; however, a close inspection of the vertical axis reveals that the decreases are very small and not statistically significant.

4.2.7.4 Summary of Research question 2, part 2

RQ2-part 2 has revealed that variations in the level of cognitive task complexity plus 10 minutes pre-task planning time led to a statistically significant change in only one instance. For the mean length of adjectival dependent clauses, a significant mean scores variation ($p=.05$) was discovered between Task 1 (low complexity) and Task 3 (high complexity). This indicates that the combination of 10 minutes pre-task planning time and higher levels of cognitive task complexity potentially led to a statistically significant increase in the mean length of adjectival dependent clauses for Task 3 (high complexity). However, there were no statistically significant mean scores increases or decreases for nominal dependent clauses or adverbial dependent clauses between Task 1, Task 2, and Task 3.

4.2.8 Research question 2, part 3: Lexical complexity

RQ2-part 3 investigated whether increases in cognitive task complexity in conjunction with 10 minutes pre-task planning time had any statistically significant effects on the lexical

variety of the written outcomes. Lexical variety was measured by using a mean segmental type-token ratio.

As with RQ1-part 3, each text was divided into segments of equal length, which were analysed with a mean segmental type-token ratio (the number of lexical types divided by the total number of words). Subsequently, the overall mean score for each text was found by adding the type-token ratios from each segment, then dividing this number by the entire number of segments in the text. For each participant, this produced three mean scores, designed to show any instance of variation in word variety among the tasks.

Table 27 below shows the mean scores for the mean segmental type-token ratios for the 30 participants of RQ2, for each task.

Table 27: Mean segmental type-token ratio

Level of complexity	Mean	Standard deviation
Task 1: Low	81.82	2.99
Task 2: Medium	81.88	3.21
Task 3: High	83.16	2.97

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity and 10 minutes pre-task planning time on the lexical variety for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high

complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda =.84$, F (2, 28) =2.64, $p=.089$, multivariate partial eta squared =.159.

No statistically significant mean scores variation was found among Task 1, Task 2, and Task 3, attributable to increases in cognitive task complexity plus 10 minutes pre-task planning time on the lexical variety of the texts.

**Figure 12. Mean segmental type-token ratio**

Figure 12 illustrates the results for mean segmental type-token ratio with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes pre-task planning time), and Task 3 (high complexity plus 10 minutes pre-task planning time), and the vertical axis showing variations in lexical variety.

Figure 12 above shows apparent increases in lexical variety as tasks become more complex; however, the increases are not statistically significant.

4.2.9 RQ3. What are the effects of post-task editing combined with cognitive task complexity on written complexity?

In the following sections, the results for RQ3 are presented starting with the findings for syntactic complexity and the mean length of T-units measured across all dependent clause types. This is followed by the individual results for adjectival, nominal, and adverbial dependent clauses. After a summary of the results for individual dependent clauses, the findings for Lexical complexity, which used a mean segmental type-token ratio, are presented.

4.2.10 Research question 3, part 1: Syntactic complexity

RQ3 investigated whether manipulating cognitive task complexity in conjunction with the application of post-task editing time resulted in statistically significant effects in the written outcomes. Divided into three parts, RQ3-part 1 investigated whether varying levels of cognitive task complexity in conjunction with 10 minutes post-task editing time had any significant effects on the written output, measured across all dependent clause types (adjectival, nominal, and adverbial).

Analysing the mean length of T-units for RQ3 was achieved in the same way as RQ1 and RQ2, by calculating the ratio of dependent clauses (nominal, adverbial, and, adjective) to independent clauses for Task 1, Task 2, and Task 3. Firstly, the number of T-units was added to the number of dependent clauses. Subsequently, this number was divided by the number of T-units to determine the mean length of dependent clauses. For each participant,

this produced three mean scores expected to show any instance of variation in the mean length of T-units between each task.

Table 28 below shows the mean length and standard deviation of T-units, in which all dependent clause types were analysed for the 30 participants of RQ3, for each task.

Table 28: Mean length T-units and all dependent clause types

Level of complexity	Mean	Standard deviation
Task 1: Low	1.30	0.12
Task 2: Medium	1.27	0.13
Task 3: High	1.28	0.16

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity and 10 minutes post-task editing time on the mean length of T-units (across all dependent clause types) for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda = .92$, $F (2, 28) = 1.12$, $p = .33$, multivariate partial eta squared $= .074$.

There were no statistically significant mean scores variations among Task 1, Task 2, and Task 3 when testing for increases in cognitive task complexity plus 10 minutes post-task editing time on the mean length of T-units measuring dependent clauses across all clause types.

In Figure 13 below, a close inspection of the vertical axis shows that the variations in the

mean length of T-units are small and not statistically significant.

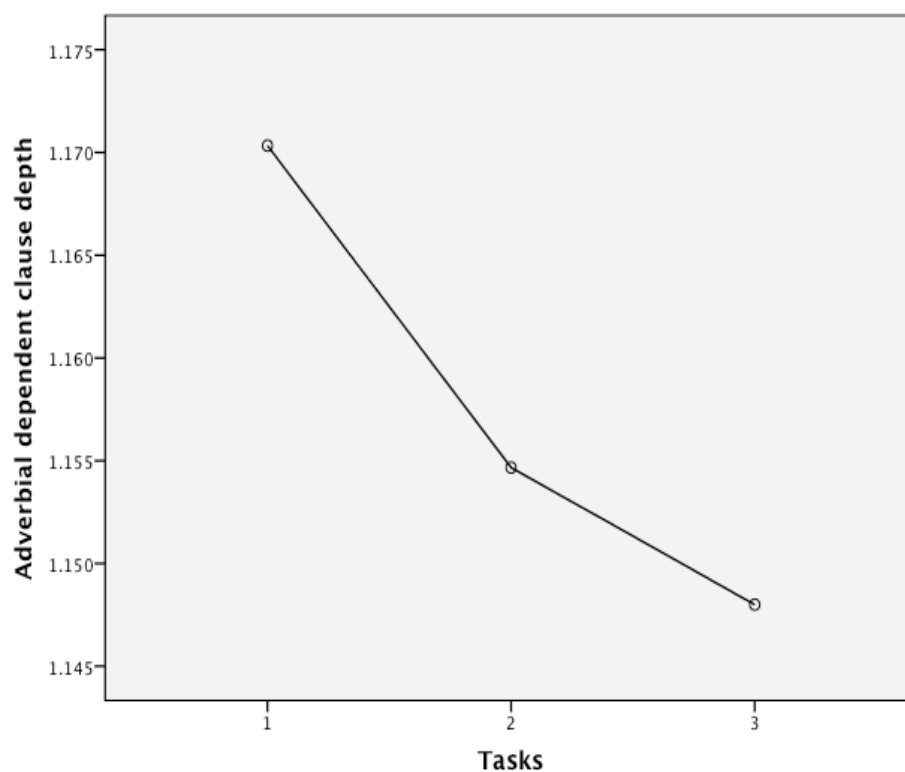**Figure 13. Mean length T-units and all dependent clause types**

Figure 13 illustrates the results for all dependent clauses with the horizontal axis showing

Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes post-task editing

Time), and Task 3 (high complexity plus 10 minutes post-task editing time), and the

vertical axis showing the mean length of T-units measured across all dependent clause

types (adjectival, nominal, and adverbial).

4.2.11 Research question 3, part 2

This part of RQ3 is a more detailed analysis of cognitive task complexity, editing time, and the mean length of T-units. It examines whether variations in cognitive task complexity and 10 minutes post-task editing time had any effect on the individual types of dependent clauses produced across Task 1, Task 2, and Task 3. As with RQ1 and RQ2, the coding process for RQ3 involved identifying and separately coding the three types of dependent clause. These were adjectival dependent clauses, nominal dependent clauses, and adverbial dependent clauses. The results for each dependent clause are presented separately in the following sections.

4.2.11.1 Adjectival dependent clauses

The mean length of T-units measuring adjectival dependent clauses was gauged by calculating the ratio of adjectival dependent clauses to independent clauses for Task 1, Task 2, and Task 3. To find this ratio, the number of T-units was added to the number of adjectival dependent clauses. Subsequently, this number was divided by the number of T-units to ascertain the mean length of adjectival dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of adjectival dependent clauses among the tasks.

Table 29 below shows the mean length and standard deviation of T-units, in which only adjectival dependent clauses were analysed for the 30 participants of RQ3, for each of the three tasks.

Table 29: Mean length T-units and adjectival dependent clauses

_____

Level of complexity	Mean	Standard deviation
Task 1: Low	1.05	0.09
Task 2: Medium	1.05	0.07
Task 3: High	1.06	0.12

_____

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity and 10 minutes post-task editing time on the mean length of T-units measuring adjectival dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant (p>.05), Wilks' Lambda Λ =. 98, F (2, 28) =. 17, p=. 84, multivariate partial eta squared =. 012.

No statistically significant mean scores variation was found among Task 1, Task 2, and Task 3, attributable to increases in task complexity and 10 minutes post-task editing time on the mean length of adjectival dependent clauses.

Figure 14. Mean length T-units and adjectival dependent clauses
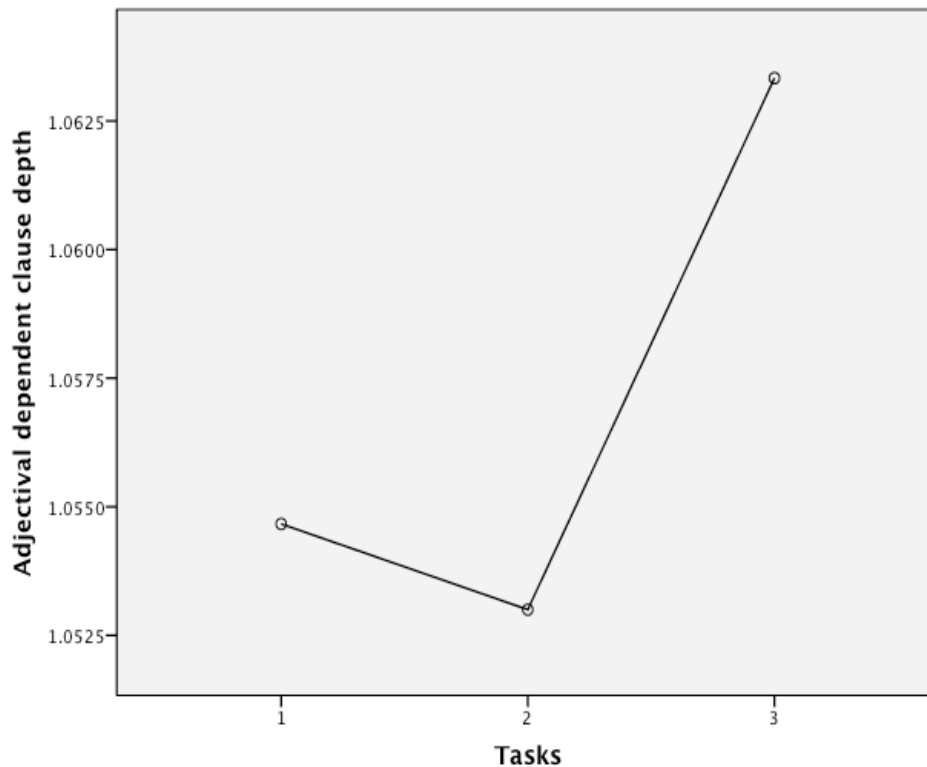
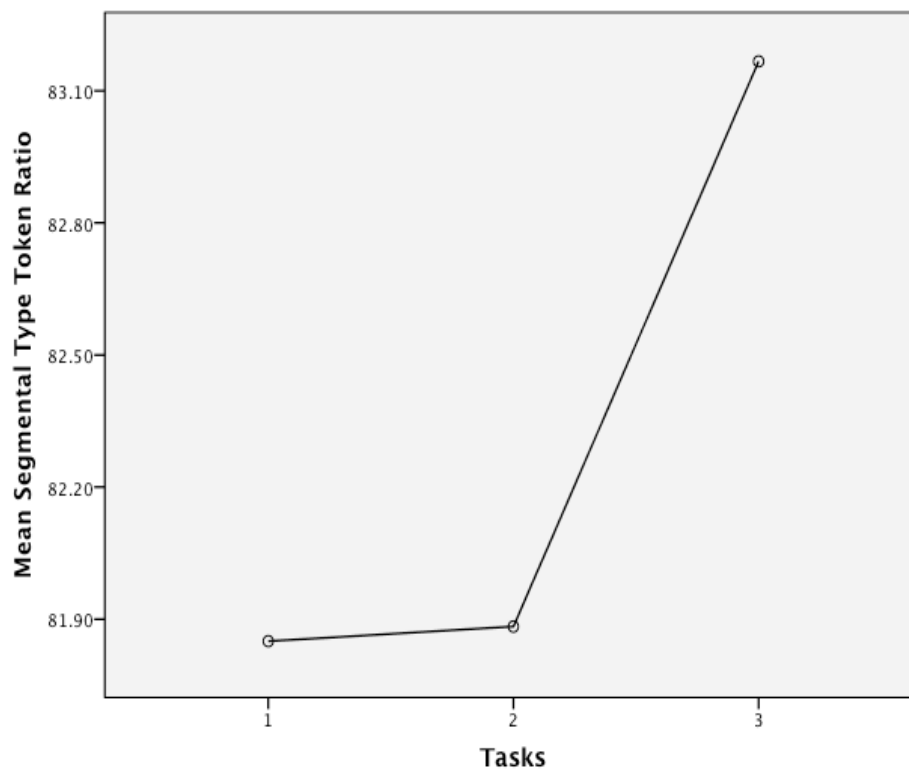This figure illustrates the results for adjectival dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes post-task editing Time), and Task 3 (high complexity plus 10 minutes post-task editing time), and the vertical axis showing the mean length of T-units measuring adjectival dependent clauses.

In Figure 14 above, a close inspection of the vertical axis shows that the variations in the mean length of adjectival dependent clauses are small and not statistically significant.

4.2.11.2 Nominal dependent clauses

To gauge the mean length of T-units measuring nominal dependent clauses, the ratio of nominal dependent clauses to independent clauses was calculated for Task 1, Task 2, and Task 3. Firstly, the number of T-units was added to the number of nominal dependent

clauses. Subsequently, this number was divided by the number of T-units to ascertain the mean length of nominal dependent clauses. For each participant, this produced three mean scores, designed to show any instance of variation in the mean length of nominal dependent clauses among the tasks.

Table 30 below shows the mean length and standard deviation of T-units, in which only nominal dependent clauses were analysed for the 30 participants of RQ3, for each task.

Table 30: Mean length T-units and nominal dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.11	0.08
Task 2: Medium	1.12	0.07
Task 3: High	1.10	0.07

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity and 10 minutes post-task editing time on the mean length of T-units measuring nominal dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda =. 97$, $F (2, 28) =. 38$, $p=. 68$, multivariate partial eta squared =. 027.

No statistically significant mean scores variation was found among Task 1, Task 2, and Task 3, attributable to increases in cognitive task complexity and 10 minutes post-task editing time on the mean length of nominal dependent clauses.

Figure 15. Mean length T-units and nominal dependent clauses
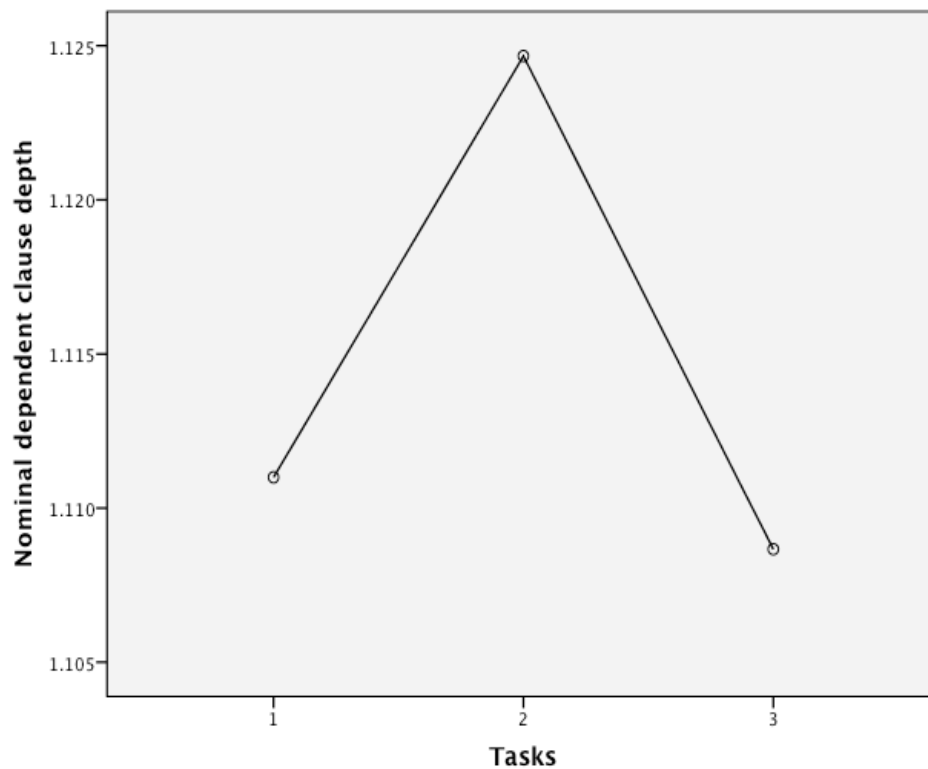
Figure 15 illustrates the results for nominal dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes post-task editing Time), and Task 3 (high complexity plus 10 minutes post-task editing time), and the vertical axis showing the mean length of T-units measuring nominal dependent clauses.

In Figure 15 above, the apparent variations in the mean length of nominal dependent clauses are small, and not statistically significant.

4.2.11.3 Adverbial dependent clauses

The mean length of T-units measuring adverbial dependent clauses was gauged by calculating the ratio of adverbial dependent clauses to independent clauses for Task 1, Task 2, and Task 3. This ratio was found by adding the number of T-units to the number of

adverbial dependent clauses then dividing this number by the number of T-units to ascertain the mean length of adverbial dependent clauses. For each participant, this produced three mean scores, which show instances of variation in the mean length of T-units measuring adverbial dependent clauses between each task.

Table 31 shows the mean length and standard deviation of T-units, in which only adverbial dependent clauses were analysed for the 30 participants of RQ3, for each task.

Table 31: Mean length T-units and adverbial dependent clauses

Level of complexity	Mean	Standard deviation
Task 1: Low	1.13	0.07
Task 2: Medium	1.09	0.07
Task 3: High	1.09	0.07

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in cognitive task complexity and 10 minutes post-task editing time on the mean length of T-units measuring adverbial dependent clauses for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p > .05$), Wilks' Lambda $\Lambda = .86$, $F (2, 28) = 2.28$, $p = .12$, multivariate partial eta squared $= .140$.

No statistically significant mean scores variation was found among Task 1, Task 2, and Task 3, attributable to increases in cognitive task complexity and 10 minutes post-task editing time on the mean length of adverbial dependent clauses.

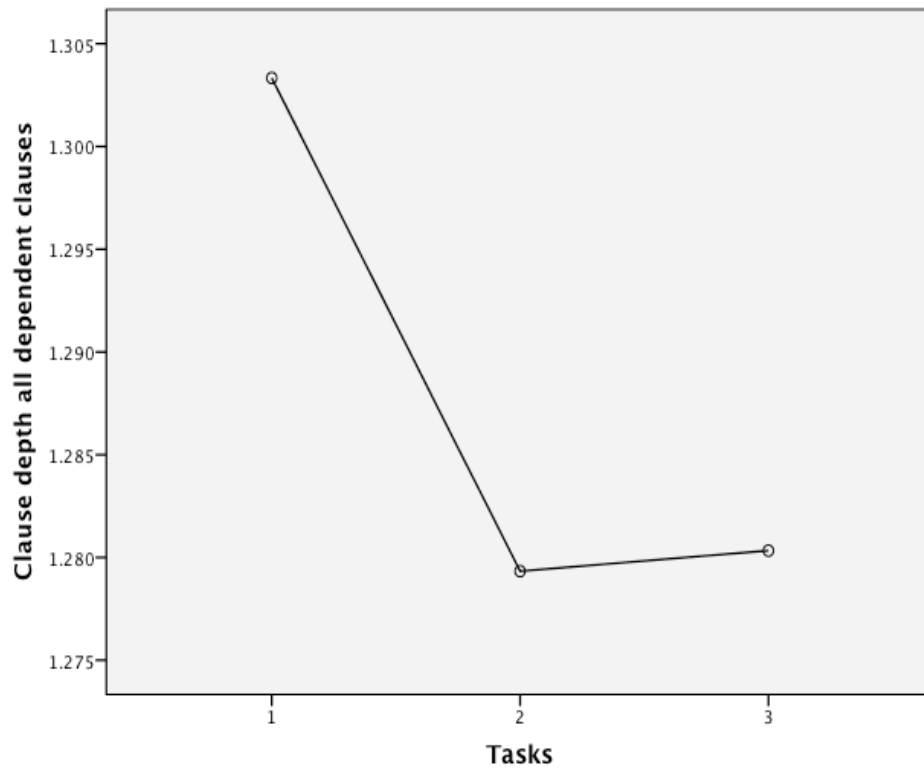**Figure 16. Mean length T-units and adverbial dependent clauses**
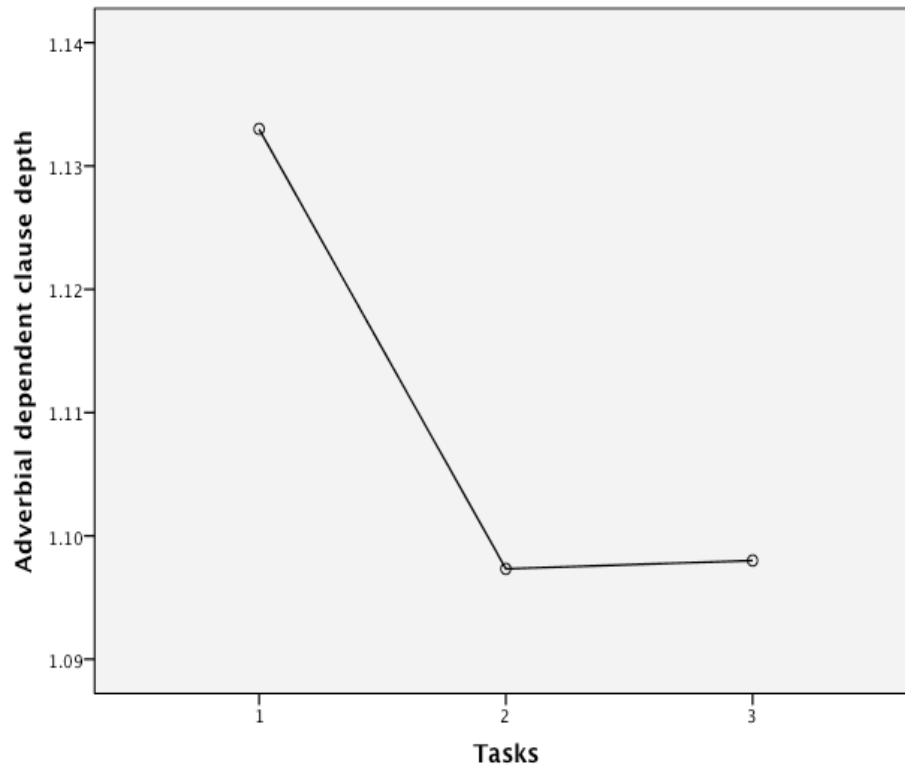
Figure 16 illustrates the results for adverbial dependent clauses with the horizontal axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes post-task editing Time), and Task 3 (high complexity plus 10 minutes post-task editing time), and the vertical axis showing the mean length of T-units measuring adverbial dependent clauses.

A close inspection of the vertical axis on Figure 16 shows that the variations in the mean length of adverbial dependent clauses are small and not statistically significant.

4.2.11.4 Summary of Research question 3, part 2

The three parts of Q3-part 2 have revealed that variations in the level of cognitive task complexity in conjunction with 10 minutes post-task editing time appear to have had no

statistically significant effect on the mean length of T-units measuring any dependent clauses.

4.2.12 Research question 3, part 3: Lexical complexity

RQ3-part 3 investigated whether increases in cognitive task complexity in conjunction with 10 minutes post-task planning time had any statistically significant effects on the lexical variety of the written outcomes. As with the previous lexical analyses, lexical variety was measured using a mean segmental type-token ratio.

Initially, each text was divided into segments of equal length, which were analysed with a mean segmental type-token ratio (the number of lexical types divided by the total number of words). Subsequently, the overall mean score for each text was found by adding the type-token ratios from each segment, then dividing this number by the entire number of segments in the text. For each participant, this produced three mean scores, designed to show any instance of variation in word variety among the tasks.

Table 32 below shows the mean score for the mean segmental type-token ratio of the 30 participants of RQ3, for each task.

Table 32: Mean segmental type-token ratio

Level of complexity	Mean	Standard deviation
Task 1: Low	79.07	3.96
Task 2: Medium	78.37	4.13
Task 3: High	78.85	4.16

A repeated measures ANOVA was conducted on the means to investigate the effect of increases in task complexity and 10 minutes post-task editing time on lexical variety for Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity). The result was not statistically significant ($p>.05$), Wilks' Lambda $\Lambda =. 97$, F (2, 28) =0.39, $p=.67$, multivariate partial eta squared =. 027.

No statistically significant mean scores variation was found among Task 1, Task 2, and Task 3, attributable to increases in task complexity and 10 minutes post-task editing time on the lexical variety of the texts.

In Figure 17 below, the apparent variations in the mean variety of lexical items are not statistically significant.

Figure 17. Mean segmental type-token ratio
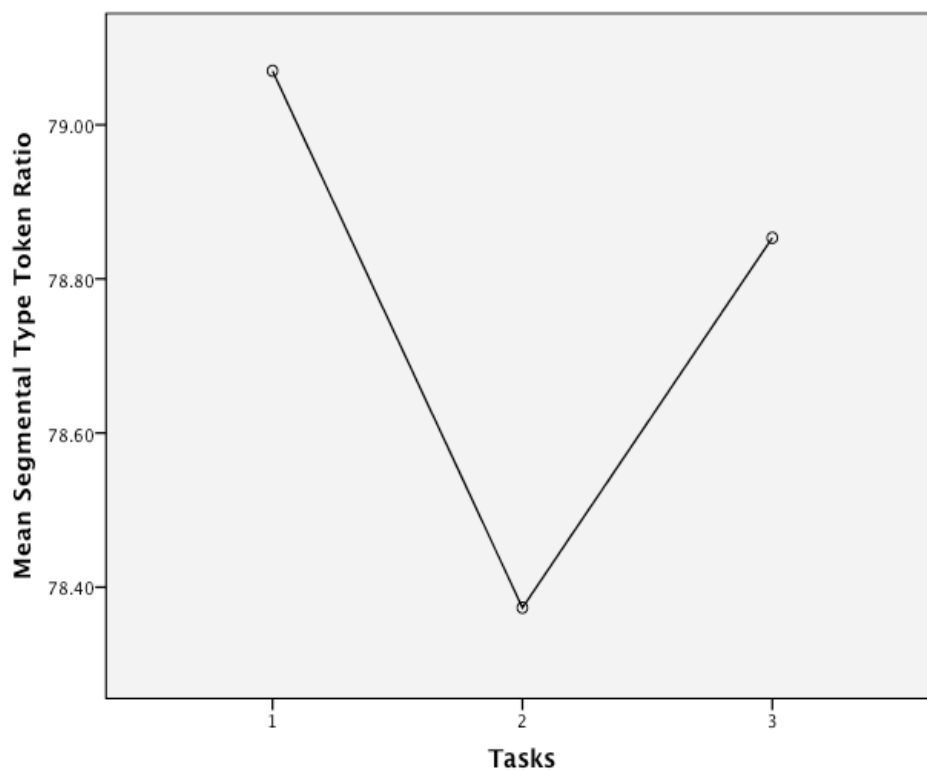
Figure 17 illustrates the results for the mean segmental type-token ratio with the horizontal

axis showing Task 1 (low complexity), Task 2 (medium complexity plus 10 minutes post-

task editing Time), and Task 3 (high complexity plus 10 minutes post-task editing time),

and the vertical axis showing variations in lexical variety.

4.2.13 Summary of within group results for RQ1, RQ2, RQ3.

Table 33: T-unit mean scores all dependent clauses for RQ1, RQ2, and RQ3

Level of complexity	RQ1mean	RQ2 mean	RQ3 mean
Task 1: Low	1.41	1.39	1.30
Task 2: Medium	1.40	1.45	1.27
Task 3: High	1.35	1.44	1.28

Table 34: Adjectival dependent clause means for RQ1, RQ2, and RQ3

Level of complexity	RQ1mean	RQ2 mean	RQ3 mean
Task 1: Low	1.05	1.07	1.05
Task 2: Medium	1.07	1.09	1.05
Task 3: High	1.07	1.12	1.06

Table 35: Nominal dependent clause means for RQ1, RQ2, and RQ3

Level of complexity	RQ1mean	RQ2 mean	RQ3 mean
Task 1: Low	1.18	1.16	1.11
Task 2: Medium	1.20	1.18	1.12
Task 3: High	1.17	1.16	1.10

Table 36: Adverbial dependent clause means for RQ1, RQ2, and RQ3

Level of complexity	RQ1mean	RQ2 mean	RQ3 mean
Task 1: Low	1.16	1.17	1.13
Task 2: Medium	1.10	1.15	1.09
Task 3: High	1.09	1.14	1.09

Table 37: Mean segmental type-token ratio for RQ1, RQ2, and RQ3

Level of complexity	RQ1mean	RQ2 mean	RQ3 mean
Task 1: Low	82.00	81.82	79.07
Task 2: Medium	82.82	81.88	78.37
Task 3: High	83.81	83.16	78.85

Table 38: Summary of significant mean scores within RQ1, RQ2, and RQ3

_____

RQ1	Task 1-Task 2	Task 1-Task 3	Task 2-Task 3
T-unit depth all dependent clauses	No	No	No
T-unit depth adjectival	No	No	No
T-unit depth nominal	No	No	No
T-unit depth adverbial	Yes -	Yes -	No
Mean segmental type-token ratio	No	Yes +	No

RQ2	Task 1-Task 2	Task 1-Task 3	Task 2-Task 3
T-unit depth all dependent clauses	No	No	No
T-unit depth adjectival	No	Yes +	No
T-unit depth nominal	No	No	No
T-unit depth adverbial	No	No	No
Mean segmental type-token ratio	No	No	No

RQ3	Task 1-Task 2	Task 1-Task 3	Task 2-Task 3
T-unit depth all dependent clauses	No	No	No
T-unit depth adjectival	No	No	No
T-unit depth nominal	No	No	No
T-unit depth adverbial	No	No	No
Mean segmental type-token ratio	No	No	No

Note: - decrease in mean scores, + increase in mean scores

4.2.14 Analysis of variance between groups for RQ1, RQ2, RQ3

4.2.14.1 Analysis of T-unit variance between groups

In order to create a clearer picture of how the results for mean length of T-units for each research question relate to each other, the findings for RQ1, RQ2 and RQ3 (where T-units were measured across all dependent clauses) were analysed together. This was done in order to clarify the overall relationship between the performances of findings from each research question, and also to show whether the performance of Task 1 (which was the same for each group) was the same for RQ1, RQ2, and RQ3.

Task 1, which was the lowest complexity task, operated as a base-line task in which the task was easy to perform and the conditions were the same for every participant. It was expected that the participants, who were all of the same proficiency level according to the schools from where they came from, should perform in a similar manner on Task 1.

Firstly, the within group mean length of T-units for RQ1, RQ2 and RQ3 were collected. These were the findings from sections 4.2.2, 4.2.6, and 4.2.10. Next, a mixed between-within subjects ANOVA was performed using these findings.

A mixed between-within subjects ANOVA can be used when analyses require the combining of a within subjects analyses of variance (in this case the within subjects analysis is the repeated measures ANOVA performed on the T-unit data from within RQ1, RQ2, and RQ3) with a between subjects analysis of variance (in this case, the between subjects analysis of variance is of the repeated measures ANOVA of T-unit data between RQ1, RQ2, and RQ3).

The findings regarding the within subjects effects of the independent variables on syntactic complexity have been extensively and separately analysed in the previous sections. This section of the analysis focuses only on the between subjects results. The between subjects results are expected to show if the mean length of T-units for each task are significantly different between RQ1, RQ2, and RQ3.

Table 39 below shows the mean length and standard deviation of T-units, in which all dependent clauses were analysed for the 94 participants across all three tasks, between RQ1, RQ2, and RQ3.

Table 39: Mean length of T-units between RQ1, RQ2, and RQ3
_____

Research Question	Mean	Standard deviation
RQ 1: Complexity alone	1.39	0.23
RQ 2: Pre-task planning time	1.43	0.25
RQ3: Post-task editing time	1.28	0.25

_____

A mixed between-within subjects ANOVA was conducted on the mean scores of RQ1, RQ2, and RQ3 where the mean length of T-units was measured across all dependent clause types. The results comparing the three different research questions was statistically significant ($p<.0005$), $F(2, 91)=9.235, p=.000$, partial eta squared $=.169$, meaning that there was a significant difference in performance between the three research questions. The effect size in this case is .169 (which is considered a large effect) meaning that the result is likely a product of the relationship between the groups and not the sample size.

The results from the Bonferroni adjustment post hoc test revealed two statistically significant mean score variations. These were (p= .008) between RQ1 and RQ3 and (p. = 000) between RQ2 and RQ3. This indicates that RQ3 had significantly lower performances of mean length of T-units than RQ1 and RQ2.

Figure 19. T-unit length between research questions
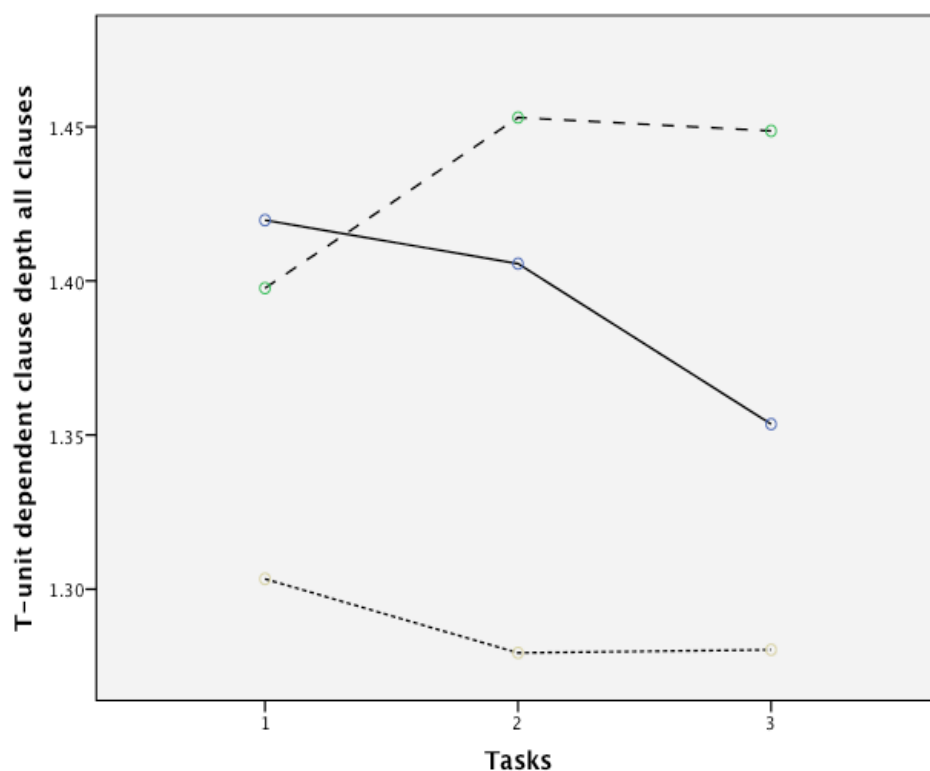
RQ1 is the solid line, RQ2 is the broken line, and RQ3 is the dotted line

Figure 19 illustrates the results for the mixed between-within subjects ANOVA with the horizontal axis showing Task 1(low complexity), Task 2 (medium complexity), and Task 3 (high complexity) performed by all the participants across RQ1, RQ2, and RQ3. The vertical axis shows the mean length of T-units measured across all dependent clause types (adjectival, nominal, and adverbial).

In Figure 19, the analysis shows that the mean length of T-unit production for the participants in RQ3 is significantly lower than both RQ1 and RQ2. Importantly, RQ3-Task 1, which had the same complexity and conditions as RQ1-Task 1 and RQ2-Task 1, performed lower than RQ1-Task 1 and RQ2-Task 1. Additionally, there appears to be very little within group variation between the three tasks for RQ3. For RQ1 and RQ2, there is only a minor variation between the performances of Task 1. However, there are non-significant divergences in the performances of Tasks 2 and Tasks 3, with RQ1 decreasing as RQ2 increases in mean length of T-units.

4.2.14.2 Analysis of mean segmental type-token ratio variance between groups

As with mean length of T-units, in order to create a clearer picture of how the results for lexical variety for each research question relate to each other, the mean segmental type-token ratio findings for RQ1, RQ2 and RQ3 were analysed together. This was done in order to clarify the overall relationship between the performances of findings from each research question, and also to show whether the performance of Task 1 (which was the same for each group) was the same for RQ1, RQ2, and RQ3.

Firstly, the within group mean segmental type-token ratios for RQ1, RQ2 and RQ3 were collected. These findings came from sections 4.2.4, 4.2.8, and 4.2.12. Subsequently, a mixed between-within subjects ANOVA was performed using these findings. A mixed between-within subjects ANOVA combines a within subjects analyses of variance (in this case the within subjects analysis is the repeated measures ANOVA performed on the lexical data from within RQ1, RQ2, and RQ3) with a between subjects analysis of variance (in this case, the between subjects analysis of variance is of the repeated measures ANOVA of lexical data between RQ1, RQ2, and RQ3).

187

The findings regarding the within subjects effects of the independent variables on lexical complexity has been extensively and separately analysed in the previous sections. This section of the analysis focuses only on the between subjects results. The between subjects results are expected to show if the mean lexical complexity of each task is significantly different between RQ1, RQ2, and RQ3.

Table 40 below shows the mean length and standard deviation of the mean segmental type-token ratio for the 94 participants, across all three tasks between RQ1, RQ2, and RQ3.

Table 40: Mean segmental type-token ratio between RQ1, RQ2, and RQ3

_____

Research Question	Mean	Standard deviation
RQ 1: Complexity alone	82.88	0.48
RQ 2: Pre-task planning time	82.30	0.51
RQ3: Post-task editing time	78.76	0.51

_____

A mixed between-within subjects ANOVA was conducted on the mean segmental type-token ratio scores of RQ1, RQ2, and RQ3. The result comparing the three different research questions was statistically significant ($p<.0005$), $F(2, 91)=19.386$, $p=.000$, partial eta squared $=.299$, meaning that there was a significant difference in performance between the three research questions. The effect size in this case is .299 (which is considered a large effect) meaning that the result is likely a product of the relationship between the groups and not the sample size.

The results from the Bonferroni adjustment post hoc test revealed two statistically

significant mean score variations. These were (p= .000) between RQ1 and RQ3 and ($p.$ =

000) between RQ2 and RQ3. This indicates that RQ3 had significantly lower mean

segmental type-token ratios than both RQ1 and RQ2.

Figure 20. Lexical variation between research questions
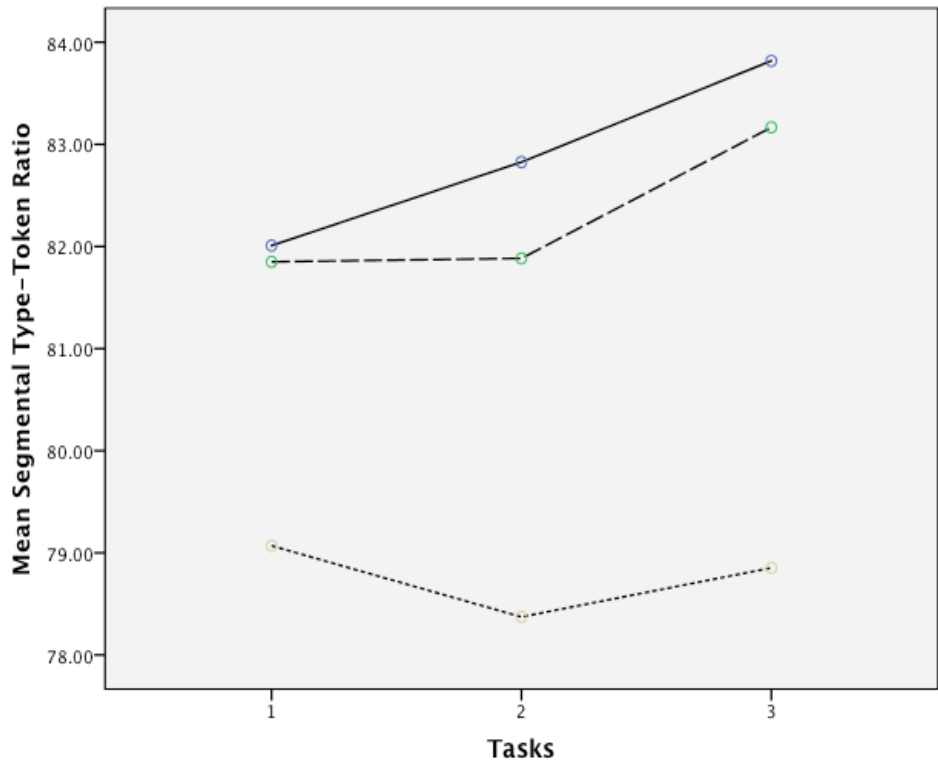
RQ1 is the solid line, RQ2= is the broken line, RQ3= is the dotted line

Figure 20 illustrates the results for the mixed between-within subjects ANOVA with the

horizontal axis showing Task 1(low complexity), Task 2 (medium complexity), and Task 3

(high complexity) performed by all the participants across RQ1, RQ2, and RQ3. The

vertical axis shows the mean segmental type-token ratio.

Figure 20 shows that the mean segmental type-token ratio for the participants in RQ3 is significantly lower than both RQ1 and RQ2. Importantly, RQ3-Task 1, which had the same complexity and conditions as RQ1-Task 1 and RQ2-Task 1, performed lower than both RQ1-Task 1 and RQ2-Task 1. Additionally, there appears to be very little within group variation between the three tasks for RQ3. For RQ1 and RQ2, there is little variation between the performances of Task 1. There is a non-significant divergence in the performance of Tasks 2 and Tasks 3; however, both tasks maintain a similar trajectory indicating increases in lexical complexity (of different significances) following increases in cognitive task complexity.

4.3 RQ4. What is the relationship between the participants' attitudes and complex written output?

Research question 4 investigated whether there were any correlations between the production of complex language in Task 1, Task 2, and Task 3 and the answers provided on the Likert scale questionnaires, in which participants answered questions about their attitudes to the tasks. The Likert scale questionnaire included questions that addressed the participants' attitudes towards the relevance and expectancy of the tasks. This correlation was analysed with a Pearson product-moment correlation (see section 3.10.6).

4.3.1 Overview

A Pearson product-moment correlation was performed between the complexity variable mean length of T-unit and the 11 items from the questionnaire, with the strength of association between these two variables provided along with an indication which of these associations are significant at .05 (see table 41 below). In addition, these results were

subjected to additional tests (see Spearman's rho and square root transformation in tables 42 and 43 below) to account for normality issues. From these results, a clearer view of which items appear to have a significant strength of association across two or more tasks is provided; subsequently, a more detailed analysis of each correlation is given to more clearly illustrate the strength of association between each variable. These analyses include the coefficient of determination (see section 3.10.7), which provides a percentage illustrating the amount of overlap between variables.

4.3.2 Correlation

Finding a correlation between the two variables (written complexity and questionnaire answers) involved utilising the mean length of T-units across all dependent clause types (from RQ1, RQ2, and RQ3) and applying a Pearson product-moment correlation to investigate the strength of association with the Likert scale questionnaire answers.

The Pearson product-moment correlation shows if there is a linear relationship between these variables. Simply put, this doesn't show that one variable definitely has a causal effect on another, but instead illustrates if there is a statistically significant relationship between the way the scores for each variable move in relation to each other. This relationship can be viewed as positive or negative. A positive relationship (denoted by a + sign in the analysis) would be one where movement along the scales being correlated increase or decrease together. A negative relationship (denoted by a – sign in the analysis) would be where movement along the scales being correlated is opposite, with one scale increasing while the other decreases.

For example, the scales in this analysis are increases or decreases in T-unit depth correlated with increases or decreases in the questionnaire scores marked on a scale of 1 to 6. If T-unit depth increases while the questionnaire scores move up towards 6, this would be a positive (+) relationship; however, if T-unit depth decreased while questionnaire scores move up towards 6, this would be a negative (-) relationship (see section 3.10.6).

In addition to the direction of the relationship (positive or negative), the strength of that relationship is also analysed. This is called strength of effect and can be characterized as small (r= .10 to.29), medium (r=. 30 to .49), or strong (r=. 50 to 1.0). A clarification of the Pearson product-moment correlation can be found in section 3.10.6.

Finally, the coefficient of determination is also analysed. This result shows shared variance between the two variables, or how one variable changes with the other. This makes the interpretation of the correlation easier by showing the per cent of the variables that overlap, thus giving a clearer indication of how meaningful the result is.

4.3.3 Strength of association

In this section, the results from the Pearson product moment correlations, showing the strength of association between mean length of T-units and the Likert scale questionnaire for all 94 participants from RQ1, RQ2, and RQ3, are presented.

The complexity variable, mean length of T-units across all dependent clause types from RQ1, RQ2, and RQ3, was correlated with all 11 items from the Likert scale questionnaire. The scores from the analyses are presented in table 41 below. Both significant and non-

significant results are provided, with any significant strength of association scores being underscored.

During the initial correlation, the analysis was checked for normality. This was done by reviewing the mean length of T units results as histograms, which illustrate distribution. The histograms revealed slight positive skews in the distribution. Subsequently, two alternatives, which were suitable for dealing with normality issues (Pallant, 2002), were applied to the data.

The first option was to use Spearman's rho (a nonparametric test), which is not subject to issues of normality; however, this option is less sensitive and has less power in the analysis. The second option was to transform the data using mathematical techniques that correct the normality issues; in this case the square root transformation was the technique that most closely represented the correction model for level of positive skew. To be as robust as possible with the results, both the Spearman's rho and the square root transformation were applied and the results were compared. Tables 42 and 43 below show the strength of association scores from both analyses.

4.3.4 Analyses

The initial analysis of the data used a Pearson product-moment correlation in which the issues of normality had not been accounted for or adjusted. The results from the Pearson product-moment correlation (in Table 42 below) revealed significant strength of associations across two or more tasks between Task 1, Task 2, and Task 3, and question 4; Task 2, and Task 3, and question 6; and Task 1, Task 2, and Task 3, and question 8.

Table 42 below shows the Pearson product-moment correlation scores for all 94 participants, correlating 11 Likert scale items with mean length of T-units from RQ1, RQ2, and RQ3.

Table 41: Pearson's strength of association for T-units and 11 Likert answers

Level of complexity	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Task 1: Low	-.076	-.174	-.172	-.376*	-.150	-.106	-.141	.210*	-.057	.023	-.239*
Task 2: Medium	-.303*	-.151	-.083	-.285*	-.168	-.238*	-.052	.249*	-.159	.104	-.172
Task 3: High	-.096	-.163	.016	-.307*	.002	-.210*	.037	261*	-.099	.054	-.136

Note: *=significant results at .05

The scores were subsequently analysed using a Spearman's rho and a Pearson product-moment with a square root transformation. These results were compared against the initial findings in table 41 above, and also against with each other, to gauge the effects of the normality issues.

Table 42: Spearman's strength of association for T-units and 11 Likert answers

Level of complexity	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Task 1: Low	-.041	-.143	-.176	-.244*	-.065	-.052	-.100	.223*	-.004	-.011	-.212*
Task 2: Medium	-.290*	-.180	-.060	-.164	-.152	-.198	.063	.284*	-.164	.103	-.153
Task 3: High	-.086	.119	.050	-.175	.015	-.170	.066	.252*	-.147	.060	-.120

Note: *=significant results at .05

Table 43: Transformed Pearson data for T-units and 11 Likert answers

_____

Level of complexity	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Task 1 SRT: Low	-.072	-.172	-.172	-.371*	-.143	-.100	-.141	.214*	-.051	.024	-.236*
Task 2 SRT: Medium	-.302*	-.156	-.082	-.282*	-.173	-.238*	-.044	.255*	-.159	.105	-.166
Task 3 SRT: High	-.097	-.158	.027	-.297*	.007	-.205*	.044	.263*	-.105	.053	-.130

_____

Note: *=significant at .05, SRT=square root transformation.

An inspection of the data in tables 42 and 43 found that the results, when adjusted for normality, were not different enough to be of concern. The Spearman's rho showed some minor losses in significance for Task 3 Q4, and Tasks 2 and 3 Q6, which might be expected as a result of the less powerful analyses of Spearman's rho. The Pearson product-moment correlation, with a square root adjustment, showed results that were not materially different from the original Pearson analysis (see table 41 above).

A review of the three different analyses revealed that the normality issues noticed in the initial results appeared to have a minimal impact. Subsequently, the following analyses are presented using the untransformed data from the original Pearson product-moment analysis. This data was chosen because the analyses has shown that the normality issues have minimal impact on the results, and by using the untransformed data, issues associated with transformed data are avoided.

4.3.5 Analyses of Correlations

Table 37 above illustrates the significant strength of associations across two or more tasks between Task 1, Task 2, Task 3, and question 4; Task 2, Task 3, and question 6; and Task

1, Task 2, Task 3, and question 8. In the following sections, more detailed analyses of the findings are presented.

All 94 participants from RQ1, RQ2, and RQ3 were included in each analysis. The scores for mean length of T-units across all dependent clauses from RQ1, RQ2, and RQ3 were subjected to a Pearson product-moment correlation. This included an analysis of both the strength of association and the coefficient of determination.

4.3.5.1 Likert Question 4

Likert question 4 stated: "I understood how these tasks were supposed to help me learn English". Likert question 4 can be viewed as addressing issues of relevance (Dörnyei, 1994), which is the degree to which class instruction is perceived to be related to a students' goals, needs and values. The results showed a significant strength of associations across Task 1, Task 2, and Task 3, indicating correlations between the mean length of T-units and question 4.

4.3.5.2 Task 1, Likert Question 4

The strength of association between mean length of T-units and question 4 was explored using the Pearson product-moment correlation. The results for Task 1 and question 4 showed a medium negative correlation between the two variables, $r = -.376$, $n = 94$, $p = .000$, $R^2 = 0.141$. The coefficient of determination ($R^2 = 0.141$) is 14.1%, which means there is 14.1 % overlap in the two variables' scores.

Figure 21. Task 1-mean length of T-units correlation with question 4

Figure 21 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 1 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 4.

The decreases in the mean length of T-units occur in tandem with increases in perception of the relevance of Task 1 (as expressed by the Likert scoring scale). The results of the correlation do not infer any causal relationship between the movements of the variables; it shows only that the correlation occurs, and that it is statistically significant. In addition, the coefficient of determination shows that the variation between the two variables is only the case for 14.1% of the data, and that the other 85.9 % of the data remains unexplained by the analysis.

4.3.5.3 Task 2, Likert Question 4

The strength of association between mean length of T-units and question 4 was explored using the Pearson product-moment correlation. The results for Task 2 and question 4 showed a small negative correlation between the two variables, $r = -.285, n = 94, p = .005,$ $R^2 = 0.081$.The coefficient of determination ($R^2 = 0.081$) is 08.1%, which means there is 08.1 % overlap in the two variables' scores.

Figure 22. Task 2- mean length of T-units correlation with question 4

Figure 22 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 2 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 4.

As with Task 1, the decreases in the mean length of T-units occur in tandem with increases in perception of the relevance of Task 2 (as expressed by the Likert scoring scale). The results of the correlation do not infer any causal relationship between the movements of the variables; it shows only that the correlation occurs, and that it is statistically significant. Furthermore, the coefficient of determination shows that the variation between the two variables is only the case for 08.1% of the data, and that the other 91.9 % of the data remains unexplained by the analysis.

4.3.5.4 Task 3, Likert Question 4

The strength of association between mean length of T-units and question 4 was explored using the Pearson product-moment correlation. The results for Task 3 and question 4 showed a medium negative correlation between the two variables, $r = -.307$, $n = 94$, $p = .003$, $R^2 = 0.094$. The coefficient of determination ($R^2 = 0.094$) is 09.4%, which means there is 09.4 % overlap in the two variables' scores.

Figure 23. Task 3-mean length of T-units correlation with question 4
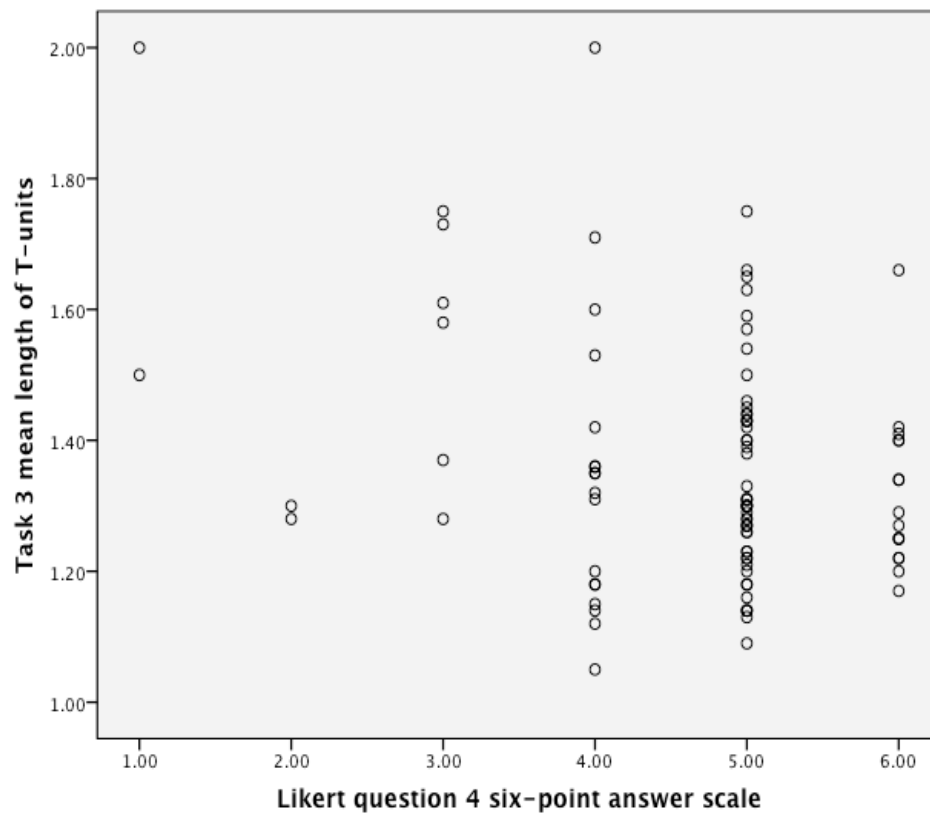
Figure 23 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 3 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 4.

Decreases in the mean length of T-units occur in tandem with increases in perception of the relevance of Task 3 (as expressed by the Likert scoring scale). The results of the correlation do not infer any causal relationship between the movements of the variables; it shows only that the correlation occurs, and that it is statistically significant. Importantly, the coefficient of determination shows that the variation between the two variables is only the case for 09.4% of the data, and that the other 90.6 % of the data remains unexplained by the analysis.

4.3.5.5 Summary of Tasks 1, 2, and 3 for Likert Question 4

Task 1, Task 2, and Task 3 all exhibited statistically significant strength of associations between decreases in the mean length of T-units, and perceived increases in perceptions of the relevance of the tasks as learning instruments (as represented by Likert questionnaire six-point scale). As mentioned above, the occurrence of this association does not imply that there is any causal relationship between the two variables; however, it may be worth noting that these associations exist across all three tasks. Though no causal relationship is inferred, the existence of the correlation suggests that further study may reveal a stronger relationship between participants who had the worst performance and their perception that they understood the tasks relevance.

However, it should be noted that in all three tasks, the percentage of the data that appeared to be explained by the correlation was low, between 14.1% and 8.1%, meaning that any relationship between performance and the perception of task utility was not addressed by the correlation for most of the data.

4.3.5.6 Likert Question 6

Likert question 6 stated: "When the tasks became difficult, I lost interest in completing them". Likert question 6 can be viewed as addressing issues of expectancy (Dörnyei, 1994), which can be viewed as the students' perception of success in completing a task when considering a tasks' difficulty and the perceived amounts of effort required for completion. The findings revealed a significant strength of associations across Task 2, and Task 3, indicating correlations between the mean length of T-units and question 6 for two of the three tasks. A more detailed account of the results for Task 2 and Task 3 are shown below.

4.3.5.7 Task 2, Likert Question 6

The strength of association between mean length of T-units and question 6 was explored using the Pearson product-moment correlation. The results for Task 2 and question 6 showed a small negative correlation between the two variables, $r = -.238$, $n = 94$, $p = .021$, $R^2 = 0.056$. The coefficient of determination ($R^2 = 0.056$) is 05.6%, which means there is 05.6 % overlap in the two variables' scores.

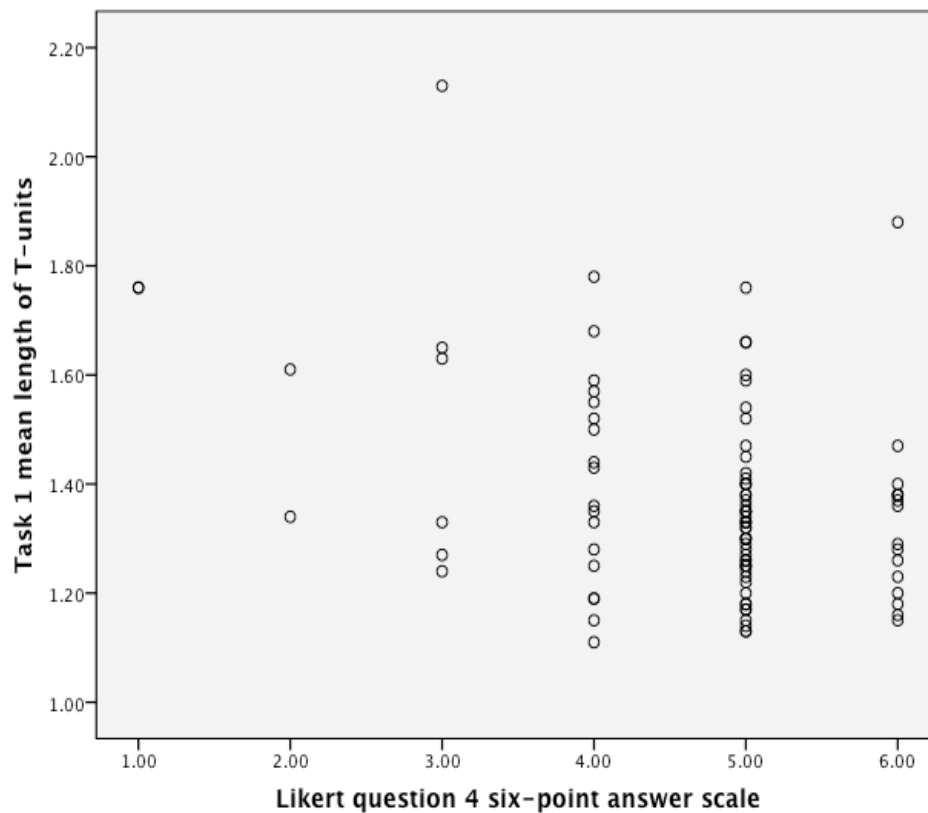**Figure 24. Task 2-mean length of T-units correlation with question 6**

Figure 24 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 2 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 6.

Decreases in the mean length of T-units occur in tandem with increases in loss of interest in completing Task 2. The loss of interest is related to the expectancy issue of perceived difficulty (as expressed by the Likert scoring scale). There is no inference of causality between the variables; only acknowledgement that the correlation exists, and that it is statistically significant. Additionally, the coefficient of determination shows that the variation between the two variables is only the case for 05.6% of the data, and that the other 94.4 % of the data remains unexplained by the analysis.

4.3.5.8 Task 3, Likert Question 6

The strength of association between mean length of T-units and question 6 was explored using the Pearson product-moment correlation. The results for Task 3 and question 6 showed a small negative correlation between the two variables, $r = -.210$, $n = 94$, $p = .043$, $R^2 = 0.044$. The coefficient of determination ($R^2 = 0.044$) is 04.4%, which means there is 04.4 % overlap in the two variables' scores.

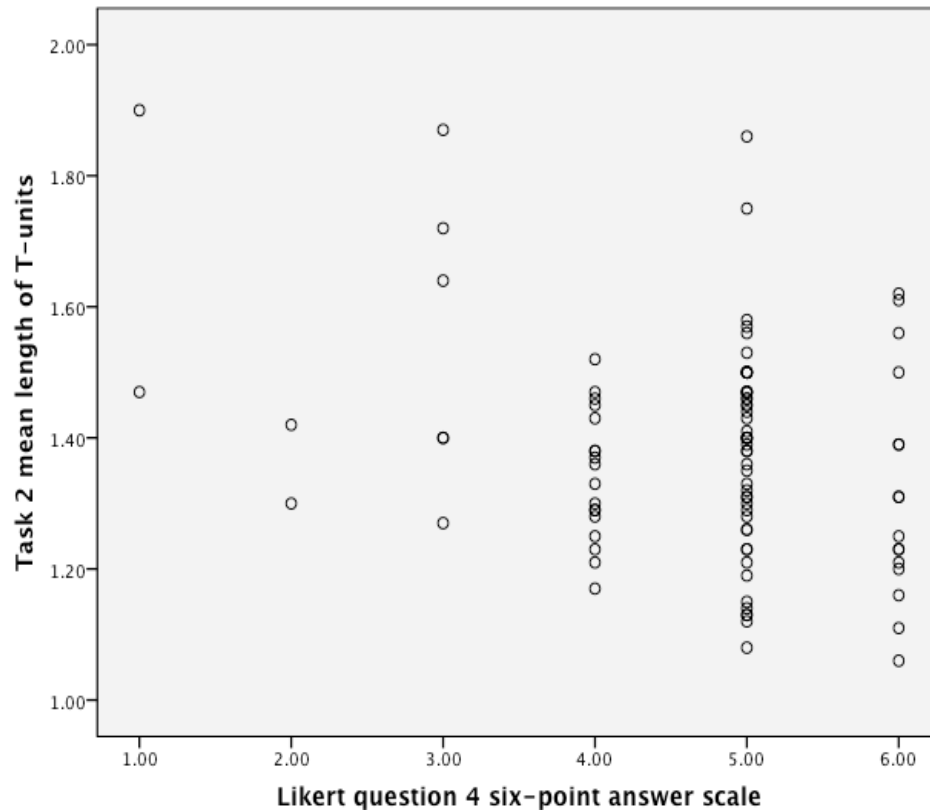**Figure 25. Task 3-mean length of T-units correlation with question 6**

Figure 25 illustrates the results of the Pearson product-moment correlation between the

mean length of T-units for Task 3 for all 94 participants taken from RQ1, RQ2, and RQ3,

and the answers from Likert question 6.

Decreases in the mean length of T-units occur in tandem with increases in loss of interest in

completing Task 3. The loss of interest is related to the expectancy issue of perceived

difficulty (as expressed by the Likert scoring scale). There is no inference of causality

between the variables; only acknowledgement that the correlation exists, and that it is

statistically significant. Additionally, the coefficient of determination shows that the

variation between the two variables is only the case for 04.4% of the data, and that the other 95.6 % of the data remains unexplained by the analysis.

4.3.5.9 Summary of Tasks 2 and 3 for Likert Question 6

Task 2, and Task 3 both exhibited statistically significant strength of associations between decreases in the mean length of T-units, and the loss of interest in Task completion due to perceived difficulty (as represented by Likert questionnaire six-point scale). No causality is implied between these two variables; however, the occurrence of this association across two tasks may be worth noting if it infers that task degradation and lack of interest due to perceived difficulty occur together. However, it should be noted that much like the correlations for question 4, the percentage of the data that appeared to be explained by the correlation was low. In this case, between 05.6% and 04.4%, meaning that any relationship between task degradation and lack of interest due to perceived difficulty was not addressed by the correlation for most of the data.

4.3.5.10 Likert Question 8

Likert question 8 was "These tasks had too many things to concentrate on", and it can be understood as addressing issues of expectancy (Dörnyei, 1994), which can be viewed as the students' perception of success in completing a task when considering a tasks' difficulty and the perceived amounts of effort required for completion. The findings revealed significant strength of associations across Task 1, Task 2, and Task 3 indicating correlations between the mean length of T-units and question 8 for all three tasks. A more detailed account of the results for Task 1, Task 2, and Task 3 are shown below.

4.3.5.11 Task 1, Likert Question 8

The strength of association between mean length of T-units and question 8 was explored using the Pearson product-moment correlation. The results for Task 1 and question 8 showed a small positive correlation between the two variables, $r = .210, n = 94, p = .042,$ $R^2 = 0.044$. The coefficient of determination ($R^2 = 0.044$) is 04.4%, which means there is 04.4 % overlap in the two variables' scores.

Figure 26. Task 1-mean length of T-units correlation with question 8
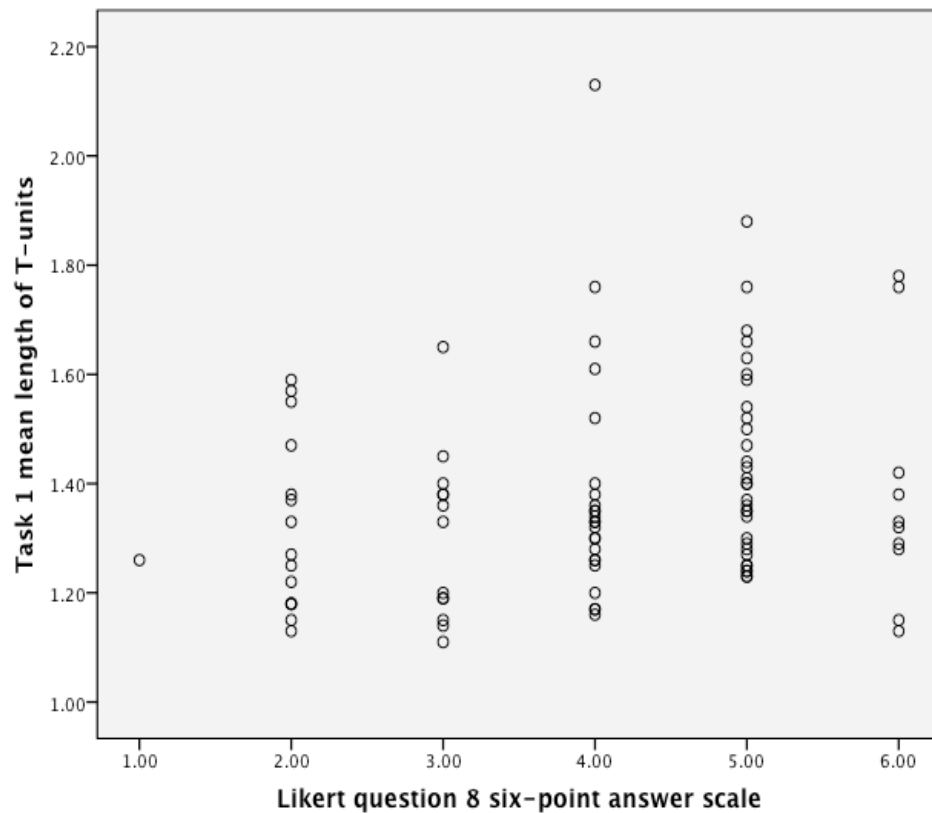
Figure 26 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 1 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 8.

Increases in the mean length of T-units occurred together with the increased perception that Task 1 had too many items to concentrate on (as expressed by the Likert scoring scale). There is no inference of causality between the variables; only acknowledgement that the correlation exists, and that it is statistically significant. Additionally, the coefficient of determination shows that the variation between the two variables is only the case for 04.4% of the data, and that the other 95.6 % of the data remains unexplained by the analysis.

4.3.5.12 Task 2, Likert Question 8

The strength of association between mean length of T-units and question 8 was explored using the Pearson product-moment correlation. The results for Task 2 and question 8 showed a small positive correlation between the two variables, $r = .249$, $n = 94$, $p = .016$, $R^2 = 0.062$. The coefficient of determination ($R^2 = 0.062$) is 06.2%, which means there is 06.2 % overlap in the two variables' scores.

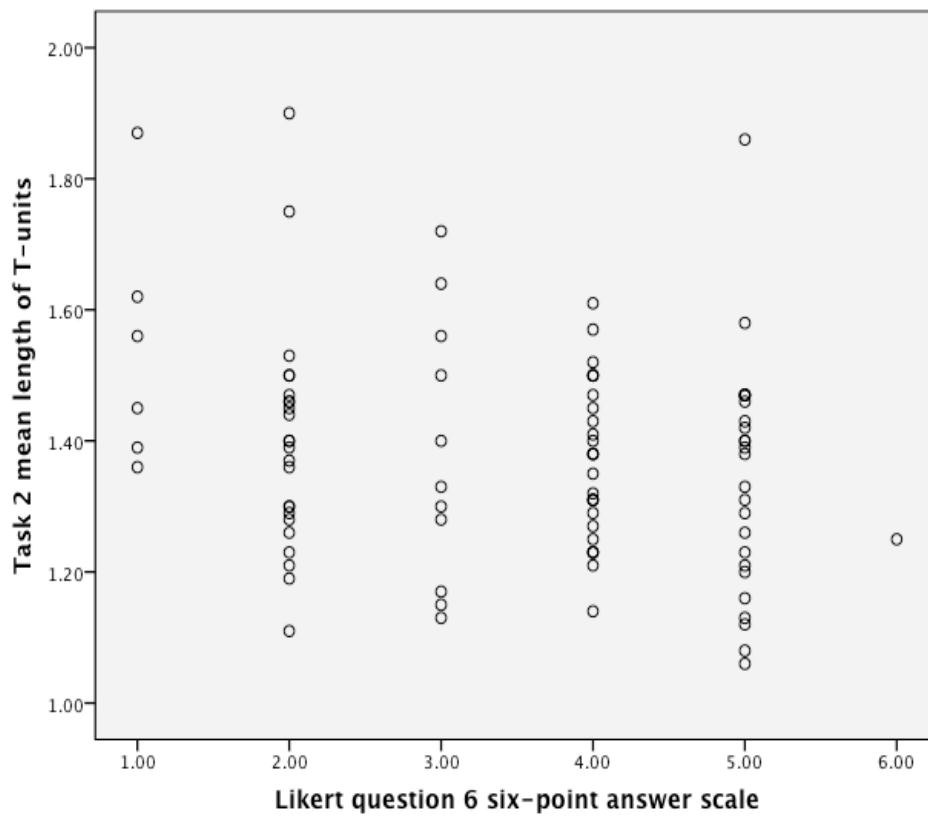**Figure 27. Task 2-mean length of T-units correlation with question 8**
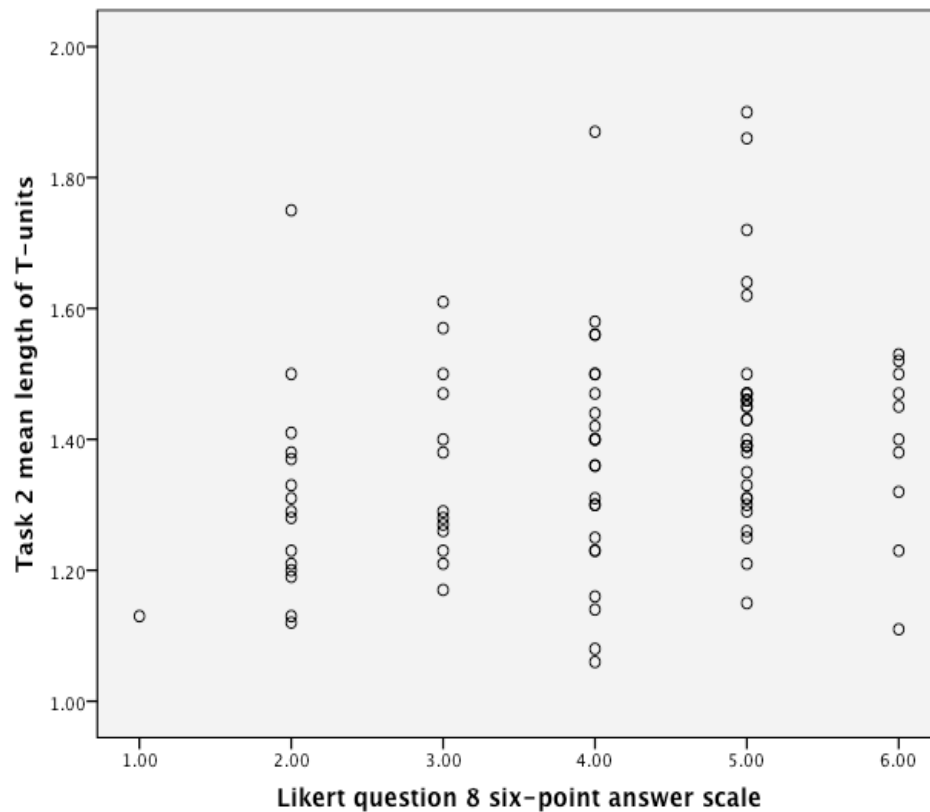
Figure 27 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 2 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 8.

Increases in the mean length of T-units occurred in tandem with the increased perception that Task 2 had too many items to concentrate on (as expressed by the Likert scoring scale). There is no inference of causality between the variables; only acknowledgement that the correlation exists, and that it is statistically significant. Additionally, the coefficient of determination shows that the variation between the two variables is only the case for 06.2% of the data, and that the other 93.8 % of the data remains unexplained by the analysis.

4.3.5.13 Task 3, Likert Question 8

The strength of association between mean length of T-units and question 8 was explored using the Pearson product-moment correlation. The results for Task 3 and question 8 showed a small positive correlation between the two variables, $r = .261$, $n = 94$, $p = .011$, $R^2 = 0.068$. The coefficient of determination ($R^2 = 0.068$) is 06.8%, which means there is 06.8 % overlap in the two variables' scores.

Figure 28. Task 3-mean length of T-units correlation with question 8
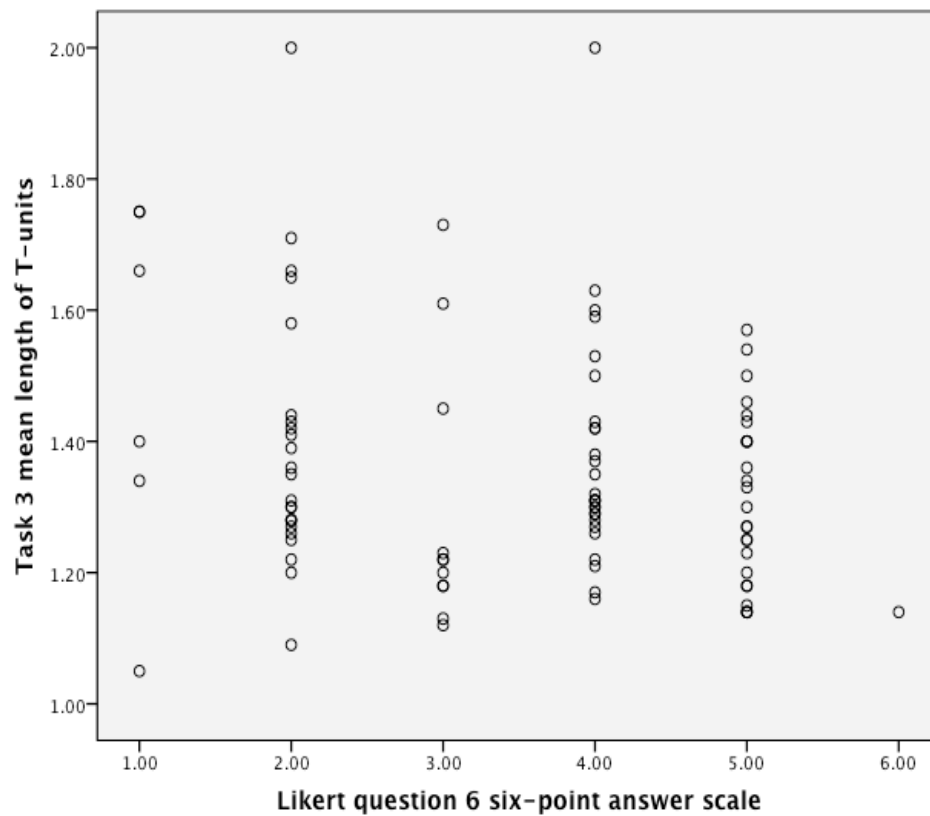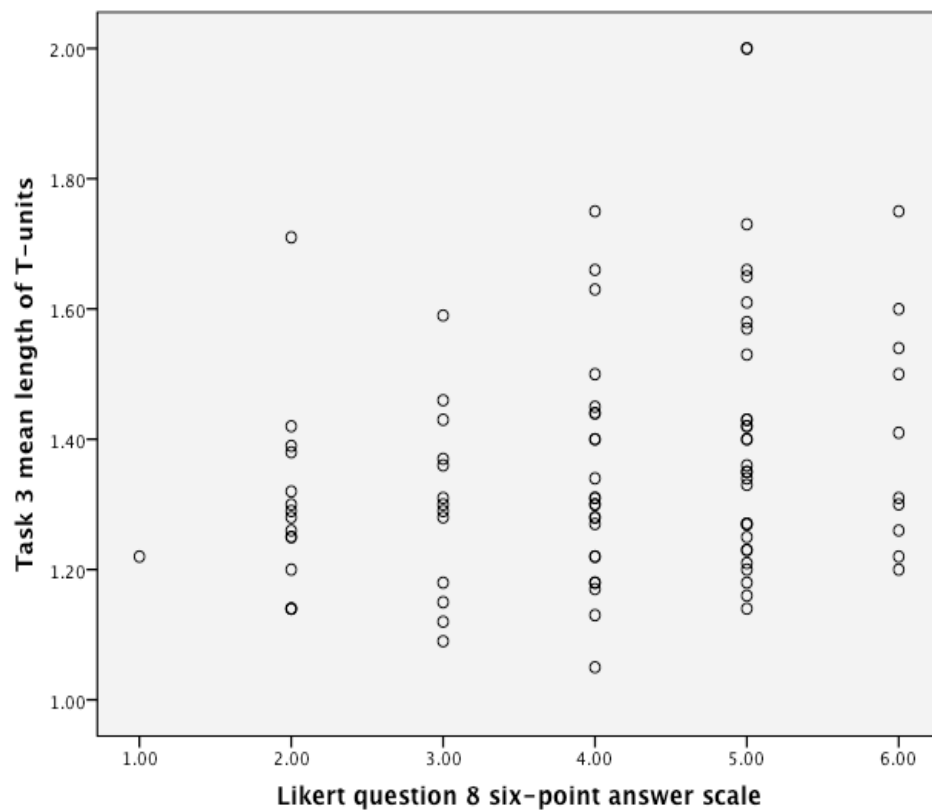

Figure 28 illustrates the results of the Pearson product-moment correlation between the mean length of T-units for Task 3 for all 94 participants taken from RQ1, RQ2, and RQ3, and the answers from Likert question 8.

Increases in the mean length of T-units occurred in tandem with the increased perception that Task 3 had too many items to concentrate on (as expressed by the Likert scoring scale). There is no inference of causality between the variables; only acknowledgement that the correlation exists, and that it is statistically significant. Additionally, the coefficient of determination shows that the variation between the two variables is only the case for 06.8% of the data, and that the other 93.2 % of the data remains unexplained by the analysis.

4.3.5.14 Summary of Tasks 1, 2, and 3 for Likert Question 8

Task 1, Task 2, and Task 3 all exhibited statistically significant strength of associations between increases in the mean length of T-units, and the increased perception that the tasks had too many items to concentrate on (as represented by Likert questionnaire six-point scale). Despite the lack of causality in the correlation analysis between these two variables, the occurrence of this association across all three tasks may be worth noting if it infers that increases in the mean length of T-units and the perception that the task was too difficulty occur together.

However, as with the previous correlations for questions 4 and 6, the percentage of the data that appeared to be explained by the correlation was low. In this case, between 04.4% and 06.8%, meaning that any possible relationship between increases in the mean length of T-units and perceptions that the task was too difficult due to the number of items, was not addressed by the correlation for most of the data.

4.3.6 Summary of findings for RQ4

Table 44*:* Summary of Pearson's strength of association and coefficient of determination

Level of complexity	Likert Q4		Likert Q6		Likert Q8	
	SOA	COD	SOA	COD	SOA	COD
Task 1: Low	-.376	14.1%			+.210	04.4%
Task 2: Medium	-.285	08.1%	-.238	05.6%	+.249	06.2%
Task 3: High	-.307	09.4%	-.210	04.4%	+.261	06.8%

Note: SOA= Strength of association, COD= Coefficient of determination as per cent,
- = Negative relationship, +=Positive relationship

A review of the results from the Pearson product-moment correlation shows that only 3 of the 11 questions showed statistically significant associations across two or three of the tasks.

For question 4, which asked if participants understood how the tasks were supposed to help them learn English, all three task showed decreases in the mean length of T-units that correlated with a stronger understanding of the tasks' utility for English learning. However, the amount of data that appeared to be explained by the correlation (as illustrated by the coefficient of determination) was low, between 14.1% and 8.1%, meaning that any relationship between performance and the perception of task utility was not addressed by the correlation for most of the data.

For question 6, which asked the participants if they lost interest in completing the tasks when they became difficult, tasks 2 and 3 showed decreases in the mean length of T-units that correlated with an increasing loss of interest as the tasks became more difficult.

211

However, the per cent of data that was explained by this analysis was also not large, in this case, between 05.6% and 04.4%.

For question 8, which asked if the tasks had too many things to concentrate on, all three tasks showed increases in the mean length of T-units that correlated with the increasing belief that the task contained too many items to concentrate on. As with the question 4 and 6, the amount of data that appeared to be explained by the correlation was low, in this case, between 04.4% and 06.8%.

CHAPTER FIVE

DISCUSSION OF RESULTS

5.1 Overview

Chapter five discusses the findings from chapter four, in which the meanings and significance of the results are reported. The findings are discussed in relation to the research questions and the supplementary issues mentioned in chapter one, which were considered important additional factors improving the scope of this research.

Sections 5.1 to 5.2.6 address the findings for RQ1. The research question is discussed along with the supplementary issues regarding the use of standard and non-standard measures of mean length of T-units, the inclusion of a patently low complexity task, Robinsons' predictions for resource-directing elements, and the effects of modality. This section finishes with 5.2.7, which summarizes the discussion of RQ1.

Section 5.3 to 5.3.6 discusses the results for RQ2. The section addresses the research question and also accounts for the supplementary issues regarding the use of standard and non-standard measures of mean length of T-units, and the inclusion on a patently low complexity task. The discussion of RQ2 also investigates how the findings relate to Robinsons' predictions for resource-directing and resource-dispersing elements, as well as the potential effects of modality. This section finishes with 5.3.7 where the discussion for RQ2 is summarised.

Section 5.4 to 5.4.3 discusses the results of RQ3. The research question is discussed in conjunction with the supplementary issues regarding the use of standard and non-standard measures of mean length of T-units, and the inclusion on a patently low complexity task. The discussion of RQ3 also investigates how the findings relate to Robinsons' predictions for resource-directing and resource-dispersing elements, as well as the potential effects of modality. This section finishes with 5.4.4 where the discussion for RQ3 is summarised.

Section 5.5 to 5.5.4 discusses the findings from RQ4, where the correlation between the Likert scale questionnaires and the production of mean length of T-units is discussed in conjunction with task motivation. The section finishes in section 5.5.5 where the discussion for RQ4 is summarised. Finally, section 5.6 concludes this chapter with a summary of the entire chapter.

5.2 RQ1: What are the effects of cognitive task complexity on written complexity?

5.2.1 Overview

RQ1 investigated the relationship between increases in cognitive task complexity (reasoning demands and number of elements) and complex written output, measured as syntactic and lexical complexity. Syntactic complexity was analysed as the mean length of T-units, which was found by calculating the ratio of T-units to dependent clauses (adverbial, nominal, and adjectival), and the results were analysed using a repeated measures ANOVA. Lexical complexity was investigated with a mean segmental type-token ratio, which finds lexical variety by dividing the texts into segments of equal length, whereupon the number of types (lexical types) is divided by the number of tokens (total number of words). These results were also subjected to a repeated measures ANOVA.

Further to the exploration of the relationship between cognitive task complexity and complex output, there were additional issues of interest investigated in the findings of RQ1. Sections 5.2.2 to 5.2.2.3 explore the two different approaches to analysing syntactic complexity (dependent clauses measured as one group, and dependent clauses measured separately). Subsequently the different approaches are discussed in terms of whether or not they produced differences in the findings that are relevant to this thesis. Following this, section 5.2.3 discusses the findings for cognitive task complexity on lexical complexity.

Section 5.2.4 investigates whether the inclusion of a patently lower complexity task (Task 1) helped to clarify the effects of cognitive task complexity on both syntactic and lexical complexity. Subsequently, section 5.2.5 reviews the results for lexical and syntactic complexity in terms of Robinsons' predictions for resource-directing variables, which are predicted to have positive effects on the written output. Next, section 5.2.6 assesses whether modality was found to have an effect on the results. Finally, section 5.2.7 concludes this section with a summary of the discussion of RQ1.

5.2.2 The effects of cognitive task complexity on the mean length of T-units

In this thesis, complex output has been investigated using two types of measures: syntactic and lexical. The first measurement addressed below is the mean length of T-units, which is a measure of syntactic complexity used to analyse the mean length of dependent clauses in the written output produced by the participants of RQ1.

As mentioned earlier, previous work (Kuiken & Vedder, 2008, 20011, 2012; Sercu et al., 2006), using similar task types and independent and dependent variables as this thesis, revealed non-significant results for increases in cognitive task complexity on syntactic

215

complexity. In fact, the results were not only statistically non-significant, there appeared to be very little variation in syntactic complexity resulting from increases in cognitive task complexity.

One of the assertions of this thesis is that the standard measure of mean length of T-units that incorporates all subordinate clauses into one group may not be a sensitive enough measure. This standard measure is used by Kuiken and Vedder (2008, 2011, 2012) and in Sercu et al. (2006). This thesis investigated the findings using both the standard measure, and a more detailed approach that separated the dependent clauses, in order to ascertain whether the two different approaches suggest something different about the effects of cognitive task complexity.

5.2.2.1 Mean length of T-units across all dependent clauses

The results for RQ1, where the mean length of T-units was measured with all subordinate clauses as one group, showed that increases in cognitive task complexity appeared to result in no statistically significant change ($p.$ =06) in the mean length of T-units. Taken solely on the basis of statistical significance, this result appears consistent with the research of Kuiken and Vedder (2008, 2011, 2012) and Sercu et al. (2006), which also showed no significant result for studies that used the same independent and dependent variables as this thesis.

However, if the current findings are considered regardless of statistical significance, then there is a more noticeable decrease in dependent clauses than is found in the work of Kuiken and Vedder (2008, 2011, 2012) and Sercu et al. (2006), which showed very little

noticeable variation. The non-significant result of this thesis may still provide some useful insights.

It could be argued that the difference between the current study and the previous work of Kuiken and Vedder (2008, 2011, 2012) and Sercu et al. (2006) is just a matter of degrees of cognitive complexity applied between tasks. Furthermore, had the cognitive task complexity levels used in this thesis been increased between tasks, the non-significant results for this thesis may have become statistically significant. Problematically, there is no way to prove this given that there is no adequate way in which to accurately measure cognitive task complexity at this time.

Ostensibly, the non-significant results could be interpreted as supporting predictions by the Limited Attentional Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001). In this case, increases in cognitive task complexity may have over-taxed the participant's limited attentional resources, leading to decreases in dependent clause production. Furthermore, the over-taxing of the attentional resources may have been exacerbated by dependent clauses being linguistic items viewed as intrinsically mentally demanding to produce (Cheung & Kemper; 1992; Lord, 2002; Mendhalson, 1983).

It is, however, not possible in this case to claim that non-significant decreases in complex output are solely the result of increased pressure on attentional resources induced by increased cognitive task complexity. Additionally, the variations in mean length of T-units noted above are based on the cumulative production of each individual dependent clause; therefore, any result is affected by whatever elements come to bear on each individual

dependent clause type. As a result, it may be more informative to assess the results at the level of individual dependent clause level.

5.2.2.2 Mean length of T-units with dependent clauses analysed separately

The results for mean length of T-unit in RQ1, in which dependent clauses (adjectival, nominal, and adverbial) were measured separately, revealed that increases in cognitive task complexity might have been one contributor to the findings, which showed numerous variations in results both between and within each dependent clause.

Whereas subordinate clauses measured as one group showed no significant results, measuring the subordinate clauses separately revealed more detailed variations, which included one significant result. The individual dependent clauses measures revealed statistically significant findings for decreases in adverbial dependent clauses, non-significant increases and decreases for nominal dependent clauses, and almost no variation across tasks for adjectival dependent clauses.

Adjectival dependent clauses: Adjectival dependent clauses had no significant change ($p=.50$) in mean length between the three tasks and were found to have the lowest occurrence of all dependent clause types. There are number of possible explanations for the findings for adjectival dependent clauses.

The inability to measure and apply cognitive task complexity with any degree of delicacy has been suggested, in this thesis, as a problem for complex language production because, if resource limits are a factor, one can never be sure when task input has induced those limits thus resulting in decrements in complex, resource demanding output.

218

However, given that the findings for RQ1 were different for each dependent clause, with nominal and adverbial dependent clauses showing different results to adjectival dependent clauses for the same amount of cognitive task complexity, it is likely that other factors, possibly in conjunction with levels of cognitive task complexity, have influenced the different findings for each dependent clause.

Another potential factor contributing to the low levels of adjectival dependent clause production across each task is the addition of procedural skill levels working in conjunction with cognitive task complexity. In this research, procedural skill level refers to the degree to which a language item is automatized and can be viewed as a constituent part of a learners' proficiency level. Ellis (2003) refers to fully procedural knowledge as being able to be used without having to think about it (thus requiring little in the way of attentional resources).

The Limited Attentional Capacity Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001) posits that the production of partially proceduralized language is subject to attentional resource limits. Furthermore, Cheung and Kemper (1992) state that subordination is demanding on memory capacity. Considering these two points, it can be seen how the level of procedural skill and the inherent memory demands of subordination (the effects of which are posited by Norris and Ortega (2009) to be noticeable on intermediate level learners) might easily be overloaded by the demands of cognitive task complexity. All three factors (procedural skill level, the inherent memory demands of subordination, and level of cognitive task complexity) may be interacting to create the finding for adjectival dependent clauses.

Given the similarly low levels of adjectival dependent clause production from Task 1 (low complexity) to Task 3 (high complexity), the procedural skill levels of adjectival depended clauses might be extremely low across the entire participating group of participants, which is to say that developmentally, many participants might not have reached a level where they were able to utilize the adjective clause structure in their writing.

While the nexus of procedural skill level, cognitive task complexity level and the processing requirements of subordination may seem like a clear explanation for the findings for adjectival dependent clauses, there are other potential elements that might be affecting the findings such as the pragmatic requirements of the task. The pragmatic requirements of the task refer to the tasks' intrinsically generated meaning expression and the obligatory language features required to express that meaning.

Bygate (1999), using argument tasks with similar complexity demands (giving reasons and justifications) as this thesis, noted that different types of tasks produced different frequencies of subordination. Subsequently, it is possible that low adjectival dependent clause production is not a result of increases in complexity, but a by-product of the task-type, one where adjectival dependent clauses are not frequently produced as a means of fulfilling the pragmatic requirements of the task.

If the pragmatic requirements of task-type were the central issue affecting the findings for RQ 1, then it might be expected that low adjectival dependent clause production would be the same for each task across RQ1, RQ2, and RQ3. However, RQ2 showed higher adjectival dependent clause production in Task 3 (which also used pre-task planning time) suggesting that variables other than task-type may be affecting the results.

Two further factors that could have affected the production of adjectival dependent clauses are participant choice (Pallotti, 2009) and social factors (Larsen-Freeman, 2002). Pallotti (2009) claims that variations in the production of accuracy, complexity and fluency cannot be exclusively attributed to psycholinguistic factors (e.g. memory, automaticity, or cognitive efficiency). Furthermore, Pallotti suggests that the elements of production may sometimes just be a matter of choice.

Two points could be made about Pallotis' assertions. Firstly, separating a participants' personal choices from psycholinguistic factors might be a difficult proposition. It is not inconceivable that choices are always made based, in part, on memory. If memory and the factors that affect memory do not have some influence on choice, then it is hard to imagine what does. Secondly, claims that the elements of language production may be based only on individual choices (separate from the psycholinguistic elements of the task) weaken as more participants in any group make the same choices. The notion that a large group of people simultaneously make the same choices devoid of other influences seems less likely than the notion that there is a common causal factor or factors influencing choices. As a result, this thesis is cautious about ascribing decreases in the mean length of dependent clause length to any choice that is viewed as separate from task influence or memory.

Larsen-Freeman (2002) states that language is not only a cognitive resource, but also one that is social in nature. As such, language used as a social action in social contexts can be subject to influences from a learners' identity, goals, or affective states (Larsen-Freeman, 2006). However, while this study acknowledges that there are social aspects to language that may have contributed to the findings for adjectival dependent clauses, those elements

have not been analysed in this thesis; therefore, it is difficult to make claims regarding their effects.

Finally, the complexity of the tasks used in RQ1 may have been increased in a manner not conducive to increased dependent clause production by the omission of planning time. Robinsons (2001a, 2001b, 2005, 2007a) refers to the omission of planning-time as resource-dispersing complexity, meaning that it is believed to have negative effects on output. Given that the results for RQ2 (in which planning time was added) show increased adjectival dependent clause production, there is evidence that this could be a contributing factor.

Nominal dependent clauses: Nominal dependent clauses also showed no statistically significant variation ($p= .12$) between tasks; however, there was more production and variation of nominal dependent clauses than adjectival dependent clauses. The findings revealed that increases in cognitive task complexity might have caused a non-significant increase in nominal dependent clause length from Task 1 (low complexity) to Task 2 (medium complexity) followed by a non-significant decrease in nominal dependent clause length from Task 2 (medium complexity) to Task 3 (high complexity).

The results for nominal dependent clause length, though non-significant, could be interpreted as an example of the important effects that inexact measuring of cognitive task complexity can have on the findings. The non-significant rise in mean nominal clause length between Task 1 and Task 2 might be the result of increases in cognitive task complexity; however, the levels of cognitive task complexity may not have been high enough to effect a significant variation.

The increase in nominal dependent clauses between Task 1 and Task 2 might cautiously be viewed as supporting Robinson's (2005, 2007a) claim that increases in cognitive task complexity induce a syntactic mode of expression required to meet the increased pragmatic complexity requirement of the task. Problematically, the requisite variations in levels of cognitive task complexity needed to achieve this may require a precision that is not possible at this time. Furthermore, the subsequent drop in nominal depended clause length between Task 2 and Task 3 (which appears to support the Limited Attentional Capacity Model) might be viewed as resulting from the same problem, with levels of cognitive task complexity exceeding the processing abilities of the participants, leading to decreases in nominal dependent clauses.

Across all three tasks, potential confirmation of both the Limited Attentional Capacity Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001) and Robinson's (2005, 2007a) Cognition Hypothesis seems to be in effect; however, the support is based on the assumption that the accurate application of cognitive task complexity is an essential factor.

In addition, a synergistic relationship between levels of cognitive task complexity and the participants' current procedural skill level of nominal dependent clauses could be mediating levels of task output, and thus influencing whether the Cognitive Hypothesis or the Limited Attentional Capacity Model most accurately explains the processes involved in complex output. Had the overall procedural skill level for nominal dependent clauses been higher, it is possible that the current level of cognitive task complexity could have led to a statistically significant increase in nominal depend clauses between Task 1 and Task 2 as well as facilitating a more positive outcome for Task 2 and Task 3.

Further to the suggestion that a delicate synergistic relationship between procedural and cognitive complexity levels may be required to elicit nominal dependent clauses, the inherent pragmatic requirements of any given task should also be considered as an additional causal effect promoting the non-significant increase between Task 1 and Task 2.

Bygate (1999) found that argument tasks (with similar complexity elements as this thesis) elicited an increased use of nominal dependent clauses as a result of participants using stock phrases to give opinions (I believe that…. and I suggest that….). If that were the case here, the increases between Task 1 and Task 2 might not be related to creative complex processing alone, but may also be connected to issues of task-type if Task 1 is viewed as a different task-type.

To design a task with noticeably lower cognitive task complexity than Tasks 2 and 3, Task 1 was designed without the elements that made Tasks 2 and 3 cognitively complex. As a result, Task 1 (in which participants use their own resources and the task instructions avoid eliciting any opinions or reasoning demands) might be viewed as a different task-type than Task 2 and Task 3 (in which participants manipulate information supplied in the instructions, which also elicit opinions and reasoning demands)

If Task 1 is a different task-type than Tasks 2 and 3, then pragmatic requirements could explain the increases in nominal dependent clauses between Task1 and Task 2; however, task-type alone would not explain why Task 3 had fewer nominal dependent clauses than both Task 1 and Task 2.

If the elements affecting the output between Task 1 and Tasks 2 and 3 were only one of task-type, it might be expected that both Task 2 and Task 3 would have higher nominal dependent clauses (even if the amount was non-significant). Furthermore, the largest non-significant variation in nominal clauses is a drop from Task 2 (medium complexity) to Task 3 (high complexity), both of which would be considered the same task-type.

Between Task 2 and Task 3, task-type is not an effective explanation for the drop in nominal dependent clauses, thus leading back to the possibility of the synergistic relationship between procedural and cognitive complexity levels.

It is possible that task-type was a contributing factor to the minor increase in nominal dependent clauses between Task 1 and Task 2, and increased complexity might not have been the central variable affecting change; however, the fact that Task 3 had less nominal dependent clauses than Task 1 suggests that influences other than task-type may have affected the output. Given that the only variable modified between tasks was cognitive task complexity, it is possible that cognitive task complexity (in conjunction with procedural level) was the reason for the variations.

Finally, the potential effects of personal choice and social influences on nominal dependent clauses are not addressed. As mentioned earlier, caution is needed when personal choice appears to happen spontaneously across large groups and the social nature of language is not covered in this part of the thesis.

Adverbial dependent clauses: The analysis of adverbial dependent clauses showed that overall there was significant variation in mean clause length ($p= .016$). The results from the

post hoc test revealed a statistically significant mean scores variation ($p= .023$) between Task 1 and Task 2, with Task 1 (low complexity) having more adverbial clauses than Task 2 (medium complexity). Additionally, there was a statistically significant mean scores variation ($p= .020$) between Task 1 and Task 3, with Task 1 (low complexity) having more adverbial dependent clauses than Task 3 (high complexity). There was, however, very little variation between Task 2 and Task 3.

Unlike the decreases in nominal dependent clauses, the decreases in adverbial dependent clauses are statistically significant. Additionally, the significant decreases in adverbial dependent clauses clearly separate Task 1 (low complexity) from both Task 2 (medium complexity) and Task 3 (high complexity), with Task 2 and Task 3 having little variation between them.

Once again, this might appear to support the Limited Attentional Capacity Model with increased levels of cognitive task complexity resulting in the degradation of adverbial dependent clauses as task became more complex. However, as with the findings for adjectival and nominal dependent clauses, there may be multiple factors responsible for the various levels of variation between tasks.

Regarding the lack of variation between Task 2 and Task 3, procedural skill level alone is not likely the only cause as higher levels of adverbial dependent clauses were found in Task 1. The lack of variation between Task 2 and Task 3 could be the result of two factors. Firstly there is the notion that the overloading of complexity leads to a minimal production level regardless of amount of cognitive task complexity; and secondly, the possibility that Task 2 and Task 3 are of a similar task-type that is not as conducive to adverbial dependent

clause production as Task 1.

It could be that over a certain level of cognitive task complexity (once past the participants' ability to process all the information required by the task) variations in cognitive task complexity cease to affect the output. It might be viewed as a type of terminal velocity where the participants operate at their minimum level allowed by the combination of procedural skill level and cognitive task complexity. In short, past a certain limit, too much pressure induces an overload that elicits the minimum amount of output needed to meet a task's requirements, and no increase in pressure over that limit has any effect.

Alternatively, task-type might be the reason for the low level of adverbial dependent clause production for Task 2 and Task 3. As noted above, Tasks 2 and 3 might be viewed as having different inherent pragmatic requirements than Task 1. The clear difference in production between Task 1 and Tasks 2 and 3 certainly infers that this is a possibility. Furthermore, if Tasks 2 and 3 are not of a task-type conducive to the elicitation of adverbial dependent clauses, then variations in cognitive task complexity may not have much impact.

Of note are the results for RQ 2 (which also utilized pre-task planning time) that showed a similar decrease in adverbial dependent clauses. Though the decrease was not statistically significant, it did follow the same downward trend with Task 1 having the highest advent of adverbial dependent clauses and Task 2 and 3 having the lowest. This repeating trend could make case for task-type being a contributing factor to the findings for adverbial dependent clause length between Task 1 and Tasks 2 and 3.

Once again, personal choice is not considered as a factor given the previous speculation

that similar spontaneous language choices across large groups seems like a difficult position to defend, and the social nature of language is not analysed here.

5.2.2.3 Considering the old and new measures of mean length of T-units

One of the aims of this thesis was to explore whether increases in cognitive task complexity had any effect on the mean length of T-units (a measure of subordination), using both the standard measure where all the dependent clauses are measured as one group, and using a non-standard measure where dependent clauses were analysed separately. This was done because the results in a standard measure are the sum of the variations of each dependent clause; however, viewing the dependent clauses as one group may obscure the details of what happens at dependent clause level.

Initially, the standard measure revealed a non-significant result for cognitive task complexity on mean length of T-units, which appeared to be consistent with the non-significant findings produced by Kuiken and Vedder (2008, 2011, 2012), when using many of the same elements (e.g., similar independent and dependent variables, and similar letter writing tasks).

However, a closer look at the results showed that the findings for the current study revealed a more noticeable (though non-significant) decrease in the mean length of T-units. A subsequent analysis of the findings using the non-standard measure revealed one significant variation that was not observable when clustered together with the other dependent clauses. Additionally, the other results, though not significant, may have revealed something about the combination of elements working in tandem with cognitive task complexity.

As a result of analysing the dependent clauses separately, this study has found a number of different and sometimes conflicting results (see above), which could be interpreted as showing that both The Cognition Hypothesis and The Limited Attentional Capacity Model may be credible explanations for what happens when cognitive task complexity is applied.

Because of the current finding, it is the position of this thesis that using the non-standard measure of mean length of T-units provides a clearer picture of the multifaceted effects of cognitive task complexity and thus provides a clear picture for analyses.

5.2.3 The effects of cognitive task complexity on lexical complexity

The second complexity measure used in this thesis was a mean segmental type-token ratio, which is a measure of lexical complexity. Contrary to the decreases in production noted in the analysis of syntactic complexity, there appeared to be a gradual increase in lexical complexity from Task 1 up to Task 3. Overall, the findings for lexical complexity showed significant increases $(p=0.46)$. The results from the post hoc test showed a significant difference $(p=.036)$ between Task 1 (low complexity) and Task 3 (high complexity).

Studies that have used similar independent variables and types of letter-writing tasks (Kuiken & Vedder, 2008, 2011, 2012; Sercu et al., 2006) have had limited success when analysing the effects of cognitive task complexity on lexical complexity. The results from these studies showed contradictory results when lexical complexity was measured as lexical sophistication, with some groups of students showing increases in complexity, but other groups showing decreases. When lexical complexity was measured using a type-token ratio, results were positive for measures that did not account for text length, but negative

when text length was accounted for. Kuiken and Vedder (2008) claim there was no significant effect at all for cognitive task complexity on lexical variation.

Contrary to the previous findings of Kuiken and Vedder, the results of this thesis have shown a statistically significant increase in lexical complexity using a mean segmental type-token ratio that accounts for text length. The increases in lexical complexity, which follows the increases in cognitive task complexity, appears to support Robinsons' assertion that increased complexity along resource-directing variables promotes increases in complex language.

However, though the results from this thesis seem to support Robinsons' approach to increases in cognitive complexity, it is possible that these findings also lend support to Skehans' notion that trade-offs in performance occur under cognitive duress.

Taking into account that writers can retrieve information from the task instructions or their memory (Kormos, 2011) when they are formulating what they will write, and that the task instructions in this thesis provided increasing lexical variation as each task became increasingly more complex, it is possible that the increases in lexical variety were the result of items being taken straight from the instructions.

The retrieval of readily available lexical units from the task instructions could be viewed as less resource-demanding than using the memory to produce dependent clauses, which have been described in the current research as resource-demanding language elements (subordination) produced under resource demanding-conditions (increasing cognitive task complexity).

Ellis and Yuan (2004) state that when a learner is formulating what they will write during the writing process, it is possible that they are prioritizing the search for lexical items over grammatical formulation in much the same way as stipulated in Levelts' (1989) psycholinguistic speaking model. This prioritizing may also explain the increases in lexical variety, as the mean length of T-units seemed to decrease under cognitive pressure, with participants using limited resources to focus on the easier means of meeting the pragmatic requirements of the tasks.

While the lexical results ostensibly support Robinsons' Cognition Hypothesis, it is possible the increase in lexical complexity are the result of trade-offs between syntactic and lexical complexity, meaning that the results lean towards Skehans' Limited Attentional Capacity Model. It is possible that trade-offs in performance are not viewed only in terms of variations within lexical and syntactic measures, but also in the participants' choices between them.

5.2.4 The inclusion of a low complexity task

Another issue of interest for the current thesis are the potential benefits of using a patently low complexity task (Task 1) in conjunction with more complex tasks (Task 2 and Task 3) when testing for the effects of cognitive task complexity on complex output.

The main motivating factor for adding a low complexity task was the concern that cognitive task complexity cannot be accurately measured when increasing and decreasing complexity. As a result, it is the position of this thesis that any findings might be obscured if the tasks were either too complex or not complex enough. For example, too little complexity between tasks might not push students enough to perform at a level where

syntactic processing is required.

Conversely, over a certain (and unknowable) amount of increased cognitive task complexity, the attentional demands might overload a participant, resulting in the participant operating at either a peak performance, or baseline performance (with peak and baseline performances determined by what their procedural skill levels in combination with other factors affecting production will allow). Once past cognitive task complexity overload, the output may not change regardless of the variations in cognitive task complexity.

It is possible that both Task 2 and Task 3, though containing different amounts of cognitive task complexity, might have exceeded the ability of the participants to process, thus eliciting peak performance, which would be the same or similar for both tasks. As a result, a task that could be considered patently less cognitively complex than the two higher complex tasks was added.

Looking at the findings for both syntactic and lexical complexity, there appears to be evidence that the inclusion of a patently lower complexity task might be helpful in clarifying the effects of cognitive task complexity on written output.

In the case of nominal dependent clause length, adverbial dependent clause length, and mean segmental type-token ratio, the inclusion of the low complexity task (Task 1) appears to have added clearer indications of the effects of cognitive task complexity. This is especially the case for adverbial dependent clause length and mean segmental type-token ratio where all statistically significant results include the relationship between Task 1 (low

complexity) and one or both of the other more complex tasks.

However, a qualification needs to be made regarding the type of low complexity task used in this thesis. To create a task of patently less complexity than the two higher complexity tasks, Task 1 was designed to get the participants to use their own resources (as opposed to using information supplied in the instruction as with Tasks 2 and 3). Additionally, the task instructions avoided eliciting any opinions or reasoning demands.

Problematically, an argument could be made that the low complexity task may be a different task-type than the two more complex tasks, and that it might have had different pragmatic requirements. For example, in the case of the adverbial dependent clause length, where variation was only clearly found along a possible task-type distinction, it might be argued that the inclusion of a low complexity task was actually the inclusion of a different task-type. As a result, it is difficult to say whether the findings for adverbial dependent clauses might be the result of task-type instead of complexity, or a combination of task-type and complexity.

In conclusion, the addition of a low complexity task appears to have generated some extra findings related to cognitive task complexity that would not have been noticeable had there only been the two more complex tasks; however, the results must be viewed with caution. In future research, it might be advisable to find a way to exclude any concerns about task-type while ensuring that the variation in cognitive task complexity between tasks is noticeably low enough to create the type of effect sought in this thesis.

5.2.5 The findings and predictions about resource-directing elements

The independent variables, number of elements and reasoning demands, used in the current thesis have been termed resource-directing variables (Robinson, 2001a, 2001b, 2005, 2007a). It is Robinsons' argument that increases in cognitive task complexity using resource-directing variables should lead to increases in complex language.

When viewing the findings from this thesis through the lens of Robinsons' Cognition Hypothesis, the findings seem to show limited support, with lexical complexity appearing to be consistent with the predictions made for resource-directing variables, but only a non-significant increase in nominal dependent clause length found between Task 1 and Task 2.

In regards to lexical complexity, the results above have been described in this thesis as possibly resulting from participants taking the easier option of lexical complexity over syntactic complexity when addressing the increasingly complex pragmatic demands of the tasks. Taking the lexical items directly from the instruction and thus choosing a readily available lexical means of expression at the expense of a syntactic mode may or may not support Robinsons' notion of resource directing.

Increasing complexity along resource-directing lines is posited to involve expending mental effort, which is required to direct attentional resources (Robinson & Gilabert, 2007) for participants to meet increasingly complex concepts with increasingly complex language. Increasing mental effort (elicited by cognitive task complexity) is predicted to result in increased complex output.

This thesis questions whether the findings for lexical complexity fit this description if the

use of lexical items can be viewed as the less complex means of expression (meaning less mental effort was expended) as opposed to the syntactic means of expression (which takes comparatively more mental effort than the lexical production). If the increased lexical complexity in the output resulted from choosing the simpler means of expression, then it might conform more to the Limited Capacity Attentional Model (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001), but with a positive complex output.

Conversely, if resource-directing only requires that increased complexity in the task instructions results in increases in complex output, regardless of whether those increases have been the result of increases in mental effort, then it could be the case that the predictions of the Cognition Hypothesis have been met in this thesis.

For the most part, the findings for syntactic complexity mostly appear counter to Robinsons' notion of resource-directing; however, the slight rise in nominal dependent clauses (between Task 1 and Task 2) could be viewed as an indicator that resource-directing might happen only when cognitive task complexity is applied with a precision not available at this time.

In conclusion, there is no clear evidence supporting resource-directing variables on syntactic complexity in these findings; however, there does appear to be evidence that supports the effects of resource-directing variables on lexical complexity, though this evidence is considered with caution, as it might be the result of taking the less attention-demanding option over the more mentally-demanding syntactic option.

5.2.6 The effects of modality

The final aspect of interest for RQ1 is whether or not modality provides better conditions under which complex output might flourish when cognitive task complexity is increased. The results of this thesis appear consistent with prior studies in which there appears to be some positive results for cognitive task complexity on lexical complexity regardless of mode. Furthermore, the results for syntactic complexity do not appear to favour the proposed benefits of the written modality on complex output, with the speculated easing of processing stress having no noticeable effect.

It has been speculated that the increases in the production of lexical complexity found in the results may not have resulted from the increased mental effort needed to meet the requirements of the complex tasks. If taking the lexical items from the rubric is a cognitively stress free option, then it may be an option that does not require the proposed processing benefits of the written modality to promote increases in lexical complexity.

Additionally, it has been speculated in this thesis that the lexical option may have been prioritized over the syntactic option as a less cognitively stressful means of meeting the pragmatic requirement of the increasingly complex tasks. The prioritizing of lexical production over grammar formulation could be similar to Levelt's (1989) psycholinguistic speaking model, suggesting that the cognitive processes leading to the current findings may not be dissimilar to the processes at work in the oral modality. In which case, no benefit for the written modality would be predicted in this study if the process of lexical prioritizing appears to the same for both modalities.

Finally, whether mode has contributed or not is difficult to speculate. In the small body of work utilizing cognitive task complexity on output that includes complexity, positive effects for lexical complexity have been found across modalities (Kormos, 2011; Michel et al., 2007). However, other positive results (Kuiken & Vedder, 2007; Sercu et al., 2006) in the written modality have used type-token ratios that did not account for variation in text length. Based on the finding of these pieces of research, caution would be advised when attributing the findings of this thesis to the proposed beneficial effects of the written modality.

Regarding the effect of mode on syntactic complexity, the findings for RQ1 do not appear to show any support for mode. One explanation could be that the proposed additional processing benefits predicted for writing may not be sufficient in this case to allow for the increased mental effort needed to produce subordination under cognitive duress. It should be noted that some of the previous studies in the written modality that did produce positive effects for syntactic complexity and the written modality (Ellis & Yuan, 2004; Ishikawa, 2006) did so in conjunction with pre-task planning time.

It is possible that mode by itself is insufficient (at intermediate level) to meet the requirement of creating syntactically more complex language, and that the addition of pre-task planning time might be included as an element affecting the production of syntactic complexity.

5.2.7 Summary of the discussion for RQ1

The discussion on RQ1 addressed cognitive task complexity's effects on syntactic and lexical complexity. It also discussed relevant issues, such as mode, which had no noticeable

influence on complex output, and the patently lower complexity task, which highlighted variations in complex output and procedural skill levels.

The effects of increased cognitive task complexity on the standard measure of mean length of T-units revealed non-significant decreases as tasks became more complex, while the non-standard measure revealed more detailed variations for each dependent clause, including one statistically significant result. These findings were interpreted as largely resulting from the interplay of procedural skill levels, the inability to accurately measure cognitive task complexity, and pragmatic task requirements.

Furthermore, these findings did not appear to support the Cognition hypothesis, with increases along resource-directing dimensions leading to decreases in syntactic complexity. Conversely, lexical complexity showed statistically significant increases; however, these were attributed to increases in lexical complexity in the task instructions, as opposed to any extra mental effort induced by cognitive task complexity, so it may not have supported the predictions of the Cognition hypothesis.

The results appeared to favour the Limited Attentional Capacity Model; however, this thesis suggests that the Cognition Hypothesis may still be a viable process within the confines of limited attentional capacities if factors such as the procedural skill levels, amounts of cognitive task complexity, and the appropriate pragmatic requirements of the task are appropriately aligned.

5.3 RQ2: What are the effects of pre-task planning time combined with cognitive task complexity on written complexity?

5.3.1 Overview

RQ2 investigated the relationship between increases in cognitive task complexity (reasoning demands and number of elements), and 10 minutes unguided pre-task planning time on complex written output, measured as syntactic and lexical complexity.

Syntactic and lexical complexity were analysed in the same way as in RQ1 with both sets of findings subjected to a repeated measures ANOVA (which analyses the performance of the tasks within each research question) as well post hoc tests whenever further in-depth analyses were required. Additionally, the results from the mixed between-within subjects ANOVA (which analyses the performance of the tasks between research questions) are used whenever comparisons are made between RQ1 and RQ2.

As with RQ1, RQ2 explores a number of additional issues in relation to the effects resulting from the relationship between the main variables. Firstly, the thesis explores whether or not the two different approaches to syntactic complexity (subordination measured as one group, and subordinate clauses measured separately) showed noticeable difference in the findings when the additional variable of pre- task planning time was added to increases in cognitive task complexity. The results are first explored separately, followed by a discussion of the utility of the two different approaches to measuring T-unit length. Secondly, the findings for lexical complexity are examined.

In the following section, the findings are investigated to ascertain whether the inclusion of a patently low-complexity task was helpful in tracking the effects of cognitive task complexity and pre-task planning time on both syntactic and lexical complexity. Subsequently, the results for lexical and syntactic complexity are considered in terms of Robinsons' predictions for the manipulation of both resource-directing variables (reasoning demands and number of elements) and resource-dispersing variables (pre-task planning time) on written complexity.

Next, the results will be considered in terms of modality, meaning that the findings for pre-task planning time on syntactic complexity and lexical complexity will be considered in relation to previous work in both the oral and written modalities that utilise pre-task planning time. Finally, the last section provides a summary of the main points discussed about the results for RQ2.

5.3.2 The effects of pre-task planning time and cognitive task complexity on the mean length of T-units

RQ2 investigated complex written output using two types of measures, syntactic and lexical. The first measurement addressed is the measure of syntactic complexity, which is the mean length of T-units. The following discussion of the findings addresses the two ways in which syntactic complexity was measured. Firstly, the effects for pre-task planning time and cognitive task complexity on subordinate clauses measured as one group are discussed. Secondly, the effects of pre-task planning time and cognitive task complexity on each individual subordinate clause are considered.

5.3.2.1 Mean length of T-units across all dependent clauses

The results for RQ2 revealed that increases in cognitive task complexity in conjunction with 10 minutes pre-task planning time resulted in no significant variation ($p= .37$) in the mean length of T-units when dependant clauses were analysed as one group.

If this result for RQ2 were viewed solely in terms of statistical significance, they would appear to be consistent with both the initial analysis of RQ1 and past work by Kuiken and Vedder, (2008, 20011, 2012), and Sercu et al. (2006). RQ1 had non-significant findings for cognitive task complexity (without pre-task planning time) on dependent clauses analysed as one group. The work of Kuiken and Vedder, (2008, 2011, 2012), and Sercu et al. (2006), which had some of the same independent and dependent variables as this thesis (without pre-task planning time), also had non-significant findings for cognitive task complexity on dependent clauses analysed as one group.

Additionally, the initial statistically non-significant result for RQ2 does not appear to support previous studies where planning time and complexity in the written modality did produce significant positive findings for syntactic complexity (Ellis & Yuan, 2004; Ishikawa, 2006).

However, it is important to note that even though the initial result for RQ2 is statistically non-significant, this does not mean that it does not provide some insight into the effects of pre-task planning time and cognitive task complexity. Much like RQ1, the measuring of T-units across all dependent clauses may obscure the more meaningful interactions of pre-task planning time and complexity on individual dependent clauses. As such a more detailed

241

analysis will be performed in the following section that concentrates on the individual dependent clauses.

Before analysing the effects of pre-task planning time and cognitive task complexity on individual dependent clauses, the non-significant results of RQ2 are viewed in relation to the non-significant result for RQ1. Some insights might be found from RQ2's results by comparing the findings between groups, as opposed to just within groups. Within groups results refers to the analyses of findings where the relationship between Task 1, Task 2, and Task 3 is only considered within one group (one group referring to the participants who performed either in RQ1, RQ2 or RQ3). Between group results refers to the comparison of results between the different groups, meaning a comparison of results between RQ1, RQ2, and RQ3.

In Figure 19 (see section 4.2.14.1), the results of the mixed between-within subjects ANOVA are illustrated as a set of lines depicting the way each research question performed in relation to the others. In Figure 19, the performance on Task 1 (low complexity) is similar for both RQ1 and RQ2, with the variation between the two groups being quite small. However, the figure shows an increasing, but statistically non-significant, divergent direction in T-unit length for Tasks 2 and 3, with RQ1 dropping in T-unit length and RQ2 increasing in T-unit length.

Viewed separately, the statistically non-significant variation in these research questions might be ignored; however, looking at the trend for both RQ1 and RQ2 to move in opposite directions suggests that pre-task planning time might have had a positive effect on

reversing the degradation of T-units seen in RQ1 and creating the conditions for increases in T-units seen in RQ2.

In the following section, the findings for RQ2 (where dependent clauses are analysed separately) are analysed with the expectation that these results might provide more details regarding both the within groups results for RQ2, as well as the divergence in results between RQ1 and RQ2 mentioned above in the between groups analysis.

5.3.2.2 Mean length of T-units with dependent clauses analysed separately

The results for syntactic complexity in RQ2 are explored by separating all the dependent clauses in order to get a more detailed account of the effects of pre-task planning time and cognitive task complexity on T-unit length. The findings for RQ2, in which adjectival, nominal and adverbial dependent clauses were measured separately, revealed that increases in cognitive task complexity in conjunction with pre-task planning time appeared to have different effects on each dependent clause.

In general, this result is similar to RQ1 in that the separating of dependent clauses revealed a number of variations not seen using the traditional measure of T-unit depth; however, the specific variations in dependent clauses were different between RQ1 and RQ2.

Whereas subordinate clauses measured as one group in RQ2 showed a statistically non-significant increase in mean length of T-units, measuring the subordinate clauses separately revealed a variety of effects with significant findings for increases in adjectival dependent clauses, statistically non-significant increases and decreases in nominal dependent clauses, and statistically non-significant decreases in adverbial dependent clauses.

Before the analyses of the individual dependent clauses, it should be stipulated that the following results be understood with the proposed effects of participant choice (Pallotti, 2009) and social factors (Larsen-Freeman, 2002) being considered as having the same relationship to the findings as in RQ1. Additionally, because the pre-task planning time in this study was unguided, it is not known exactly how the participants planned prior to writing, thus no comment can be made on how the type of planning might have affected the findings.

Adjectival dependent clauses: Adjectival dependent clauses showed a significant change ($p=.037$) in the depth of adjectival dependent clauses between with Task 1 (low complexity) having the lowest mean length of adjectival dependent clauses and Task 3 (high complexity) having the highest mean length of adjectival dependent clauses. The mean length of adjectival dependent clauses rises from Task 1, to Task 2, to Task 3.

Ostensibly, this results shows that, for Task 3 (high complexity), the combination of increases in cognitive task complexity and 10 minutes pre-task planning time elicited more complexity in the form of greater amounts of adjectival dependant clauses as a response to the increasingly complex pragmatic requirements of the tasks. This appears to support Robinsons' claims that increases along resource-directing elements (number of elements and reasoning demands) in conjunction with decreases in resource-dispersing elements (no planning time) provide the conditions for participants to express complexity in the output to match the needs of the increasingly complex tasks.

When trying to isolate causes explaining why Task 3 in RQ2 increased in one aspect of T-unit production (as opposed to Tasks, 1, 2, and 3 in RQ1, and Tasks 1 and 2 in RQ2, which

did not) it may be prudent to initially review the central factors affecting RQ1 and clarify how they might correspond to, or have affects the results of, RQ2.

In the discussion of the findings of RQ1, the low level of adjectival dependent clause production and variation across all three tasks was speculated as being the result of a number of elements. Two of these were procedural skill level and the pragmatic requirement of the tasks.

Previously, it was suggested that Task 1 and Tasks 2 and 3 might be different task-types with different pragmatic requirements. Subsequently, it might be easier to firstly discuss only the results of Task 2 and Task 3 for both RQ1 and RQ2 because Task 2 and 3 are easily identifiable as the same task-type.

The notion that the pragmatic requirements of task-type as opposed to variations in complexity are responsible for the lack of significant adjectival dependent clause variation across both Task 2 and Task 3 for both RQ1 and RQ2 might not be such a strong explanation given the significant increase in adjectival dependent clauses noted in Task 3-RQ2 (high complexity) when performed with pre-task planning time.

If it is considered that procedural skill level is the same for both RQ1 and RQ2 (see below), then the following explanation might be considered accurate. The findings for Task 2-RQ2 and Task 3-RQ1 are not statistically different. Task 2-RQ2 has pre-task planning time, but less cognitive task complexity than Task 3-RQ2; and Task 3-RQ1 has the same amount of applied cognitive task complexity as Task 3-RQ2, but no pre-task planning time. As a result, the major factors differentiating the results of Task 3-RQ2 seems to be the amount of

cognitive task complexity working in conjunction with pre-task planning time.

It appears that the ability to accurately measure cognitive task complexity is again an important factor, this time in conjunction with the potential processing benefits supplied by pre-task planning (discussed later in this section).

In the discussion of results for RQ1, procedural skill level was speculated as being at such a low level, the participants' output was immune to the possible effects of increased cognitive task complexity. With that in mind, it should be noted that the level of adjectival dependent clause produced between RQ1 and RQ2 mostly appear close.

Though the participants' level of proficiency was identified as being related to IELTS scores, which were supplied by each school (see methodology section), Task 1 was used as a means of double-checking performance ability. Task 1 (low complexity) was used as a test task in which low complexity and the same performance conditions were expected to provide some insight into procedural skill level between research questions (which had different groups of participants) by showing whether the difference in output was statistically different. As can be seen in Table 34 (section 4.2.13), the difference in adjectival dependent clause means between Task 1 for both RQ1 and RQ2 is slight.

Whether the small difference in procedural skill level, inferred by the minor variation between Task 1-RQ1 and Task 1-RQ 2, had any effect on the findings is debatable. In this case, either a very small difference in procedural skill levels has had marked effect on the results for RQ 2, when mixed with the correct amount of cognitive task complexity and planning time; or the differences in procedural skill levels for adjectival dependent clauses

between RQ1 and RQ2 are irrelevant, and it is only pre-task planning time in conjunction with the amount of cognitive task complexity applied in Task 3 that has affected the increase in adjectival dependent clauses.

If procedural skill level is removed as a central factor affecting the results between RQ1 and RQ2, and considering that Task 2 in RQ2 also had pre-task planning time but significantly less complexity than Task 3, it appears that the variation in adjectival dependent clauses is the result of planning time and the increased amount of complexity used in Task 3.

As mentioned earlier, exact measurements of cognitive task complexity are not possible at this time, so explaining in detail how the specific different amounts of cognitive task complexity affected the results is unlikely. However, past research has speculated on how pre-task planning time positively affects processing, and these insights may clarify how pre-task planning time facilitated the increased production of adjective dependent clauses.

Planning time is believed to be beneficial in a number of ways; for example, Kawauchi (2005) states that planning time allows learners performing difficult tasks to access maximal language knowledge. Adams, Armani, Newton, and Nik (2013) suggest that planning time facilitates the allocation of limited attentional resources to various language aspects and Ellis and Yuan (2004) suggest that it aids with the process of formulating what will be written. In this way, planning time is expected to effect the production of language.

Further detailing the proposed positive effects of pre-task planning time, Ellis and Yuan (2004) suggested that planning time might allow writers to focus more on the propositional

content (meaning) of the task, allowing for the formulation of specific types of complex language needed to address the appropriate meaning expression. Formulation, as mentioned here, refers the aspect of Kellog's (1996) writing process where prior to writing; participants formulate what they will write. Formulation entails planning, which involves establishing and organising goals and ideas; and translating, which involves selecting the syntactic and lexical items needed to meet the ideas and goals of the planning stage.

Taking into consideration the facilitative effects of pre-task planning mentioned above and the appearance that pre-task planning time has facilitated the best result when applied to Task 3, which has the highest level of task complexity used in this thesis, the following suggestions are made as possible processes leading to the positive findings for Task 3-RQ2.

The level of complexity in Task 3 is such that participants are induced to meet the complex pragmatic requirements of the task by producing more adjectival dependent clauses. During the pre-task planning phase, participants may have structured or prepared information with the effect of relieving the cognitive burden on the formulating process. This relieving of the cognitive burden could be understood as meaning that the participants do not have to hold so much task information in the memory while formulating.

As a result of fewer burdens on the memory and attentional resources during the formulation process, participants were able meet the propositional needs of the content more efficiently by using adjective dependent clauses to insert as much relevant detail as possible in the limit time and space.

Finally, it should be noted that the minor variation in procedural skill level between RQ1 and RQ2 may or may not have contributed to RQ2-Task 3 having a positive result for increases in cognitive task complexity and pre-task planning time. It should be regarded as a potential contributing factor because each small increase in individual dependent clauses noted in Task 1 (see Tables 34, 35, and 36) is followed by a subsequent increase in production across all tasks for that research question. What this means is that higher production on Task 1 is followed by comparatively higher production on Tasks 2 and 3, and lower production on Task 1 is followed by comparatively lower production on Tasks 2 and 3 for all dependent clauses.

In short, pre-task planning might have relieved online processing pressure during the formulating stage. With the students not needing to maintain so much information in the memory (as well as possibly having a slightly higher adjectival dependent clause procedural skill level) the formulation of more complex syntax (complex syntax meaning adjectival dependent clauses as an appropriate response to the mix of appropriate levels of cognitive task complexity, pragmatic task requirements, and current procedural skill level) was possible.

Finally, one needs to consider what it is about Task 3 that required adjectival dependent clauses, which are used to provide more information about noun phrases. Given that Task 3 required large amounts of information to be transmitted to a person who had no prior knowledge of the places or people being written about, adjectival dependent clauses could have been the means by which the participants could best clarify the large amount of information within a limited space.

Nominal dependent clauses: Nominal dependent clauses showed no significant change ($p=.28$) in the depth of nominal dependent clauses between Tasks 1, 2, or 3. Both Task 1 (low complexity) and Task 3 (high complexity) had the same mean scores. Task 2 (medium complexity) scored higher than both Tasks 1 and 3; however, the increase was statistically non-significant.

The within group results for RQ2 appear to show that pre-task planning time had no statistically significant effect on nominal dependent clauses; however, it might be helpful to look at the between groups results to view whether there are any significant differences for nominal dependent clauses between RQ1 and RQ2.

If compared against the results from RQ1, two things become apparent. Firstly, the results for RQ1 and RQ2 are similar with each task in each group producing mean length of nominal dependent clause scores that are not significantly different (see Table 36 in section 4.2.13). Secondly, the variation between the tasks in RQ2 follow a similar pattern as in RQ1, with a jump in variation from Task 1 (low complexity) to Task 2 (medium complexity) followed by a drop in variation from Task 2 (medium complexity) to Task 3 (high complexity).

Though the amount of variation between all three tasks is similar within both RQ1 and RQ2 (with patterns and amounts of variation between tasks appearing similar) the overall production of nominal dependent clauses is slightly smaller for Task 1, Task 2 and Task 3 in RQ2.

As mentioned above in the discussion of adjectival dependent clauses, it is difficult to predict how the small variations in nominal dependent clause length accurately reflect quantifiable variations in the participant's procedural skill level. Furthermore, it is also difficult to accurately predict what effect these small variations have (if any) on syntactic production when subjected to variations in cognitive task complexity and pre-task planning time. As a result, it is difficult to know if the slightly lower level of nominal dependent clause production viewed in Task1-RQ2 indicates that procedural skill level has affected the results. However, it is possible that the uniform nature of the variations in nominal dependent clause length between and within groups shows that nominal dependent clause production is being duplicated in RQ2 but at a slightly lower procedural skill level, and importantly, without effect for pre-task planning time.

The results for adjectival dependent clauses above suggest that pre-task planning, the appropriate combination of cognitive task complexity, pragmatic task requirements, and possibly very minor variations in procedural skill level of dependent clauses can come together to affect the results. Therefore, it is possible that in the findings for RQ2 and nominal dependent clauses, the combination of these factors has not been favourably balanced enough to facilitate statistically significant results. Following this line of thinking, it appears the pre-task planning time might have had no effect at all on the production of nominal dependent clauses in Task 2, and Task 3 in RQ2.

Adverbial dependent clauses: Adverbial dependent clauses showed no significant variation ($p=.69$) in the depth of adverbial dependent clauses between Tasks 1, 2,or 3. From Task 1 (low complexity), which had the highest mean score, the mean scores dropped

incrementally, with Task 2 (medium complexity) being lower than Task 1 and Task 3 (high complexity) having the lowest mean score of all three tasks.

The within group variation of adverbial dependent clauses for RQ2 does not show any significant scores; however, the between group comparisons reveals that the same pattern of degradation seen in RQ1 can be seen in RQ2, except in this case, the addition of pre-task planning time may have contributed to lesser degradations of adverbial dependent clauses than that seen in RQ1 (see Table 36 in section 4.2.13).

The small variation in mean scores between Task 1-RQ1 and Task 1-RQ2 are viewed as representing similar procedural skill levels, which have resulted in the participants of RQ1 and RQ2 producing similar amounts of adverbial dependent clauses under the same conditions. However, it is not possible in this thesis to say with certainty what the difference between the mean scores (1.16 in Task 1-RQ1, and 1.17 in Task 1-RQ2 in Table 36 in section 4.2.13) actually signify in terms of accurate indicators of variations in skill level between the participants of RQ1 and RQ2. What is interesting to note is that, across all the findings on dependent clauses in this thesis, whichever group has the lower score on Task 1 also has the lower scores on Tasks 2 and 3.

When considered in relation to the within groups results for adverbial dependent clauses in RQ2, cognitive task complexity appears to have had a minor negative effect, and pre-task planning seems to have had no positive effect. However, it is in comparing the results between groups that a different picture emerges where cognitive task complexity and pre-task planning appear to have had a noticeable positive effect.

252

Pre-task planning, cognitive task complexity, and a possible effect for a small increase in procedural skill level, appears to have noticeably increased the production of adverbial dependent clauses in Tasks 2 and 3 for RQ2 over Tasks 2 and 3 for RQ1.

However, the same potential combination of influences (pre-task planning, cognitive task complexity, and a possible effect for a small increase in procedural skill level) in RQ2 have resulted in decreases in adverbial dependent clauses as tasks have become more cognitively complex It is the position of this thesis that this result again says something about the inability to accurately measure cognitive task complexity between tasks and the delicate balance required between that measurement and procedural skill level and the pragmatic requirements of tasks.

For Tasks 2 and 3 in RQ2, pre-task planning time has allowed the participants to hold less information in the memory while formulating output. Having relieved online processing pressure (and working in conjunction with a possibly slightly higher adverbial dependent clause procedural skill level, and pragmatic task requirements) pre-task planning time appears to have allowed participants to formulate increased adverbial dependent clauses to meet the complexity requirements of Tasks 2 and 3. However, this is only noticeable compared to Task 2 and 3 RQ1, where the absence of pre-task planning time seemed to result in excessive cognitive pressure, leading to fewer adverbial dependent clauses (amongst others) and a trade-off in production as lexical complexity was utilised.

Conversely, the amount of cognitive task complexity applied in Tasks 2 and 3 appears too much for participants to process in relation to Task 1. Even with the mitigating effects of pre-task planning time (which can be seen working between Tasks 2 and 3 between both RQ1 and RQ2), the amount of cognitive task complexity and level of procedural skill do

not appear balanced enough to elicit positive outputs.

The effect of task-type on adverbial dependent clause production is more complicated when viewed in terms of between group results. In RQ1, the results could fairly be assumed to be the result of differing task-types, as there is a clear statistically significant difference between Task 1, which could be viewed as one type of task; and Tasks 2 and 3, which could be viewed as another task type. However, though the overall pattern of decreases in mean scores are similar between RQ1 and RQ2, there is no significant decrease in adverbial dependent clause production in RQ2, so the potential differential effects of task-type are no longer so apparent or strong.

With the addition of the proposed mitigating effects of pre-task planning, (and the potential effects of small amounts of procedural skill levels) the degradation in adverbial dependent clauses seems a little less like that of different task-types, and maybe a little more likely related to the amounts of applied cognitive task complexity and procedural skill levels.

To summarise, it appears that pre-task planning time has had a moderating effect on the production of adverbial dependent clauses that is only noticeable between RQ1 and RQ2. Tasks 2, and 3 in RQ2 have more adverbial dependent clauses than Tasks 2 and 3 in RQ1. However, regarding the relationship between tasks within RQ2, the amount of cognitive task complexity and level of procedural skill is such that not even the mediating effects of pre-task planning time seems able to prevent degradations in the production of adverbial dependent clauses between Task 1 and Tasks 2 and 3.

5.3.2.3 The holistic and discrete measures of mean length of T-units

As with RQ1, RQ2 has utilized two different approaches to measuring the mean length of T-units. RQ2 explores whether increases in cognitive task complexity and 10 minutes pre-task planning time had any effect on the mean length of T-units (a measure of subordination), using both the standard measure where all the dependent clauses are measured as one group, and using a non-standard measure where dependent clauses were analysed separately.

The non-standard measure of T-units in RQ2 was used because the results in a standard measure are the sum of the variations of each dependent clause, and viewing the dependent clauses as one group might obscure the combined effects of cognitive task complexity and 10 minutes pre-task planning time on each dependent clause.

In RQ2, the standard measure where all dependent clauses were measured as one group revealed a statistically non-significant result for the combined effects of 10 minutes pre-task planning time and cognitive task complexity on mean length of T-units. Ostensibly, this result appears to be consistent with the non-significant findings produced by past research (Kuiken & Vedder, 2007, 2008, 2011; Sercu et al., 2006) using similar independent and dependent variables, and letter writing tasks.

However, much like the findings for RQ1, considering the results only in terms of statistical significance might be misleading as the non-significant results may also provide something important about the effects of the independent variables. A between groups comparison showed that the results for RQ2, using the standard measure, displayed a trend towards increase in complexity whereas RQ1 showed a trend towards decreases. As such,

the standard measure of T-units provided some useful insights when compared between groups.

A subsequent analysis of the findings of RQ2 using the non-standard measure where dependent clauses were measured separately revealed one significant variation that was not observable when dependent clauses were clustered together in the analysis. Additionally, the other dependent clause results, though not significant, appear to have provided more perspective on the relationship between amounts of cognitive task complexity's effect on levels of procedural skill level and the pragmatic requirements of tasks both within RQ2 and between RQ1 and RQ2.

As a result of the current findings, it is the position of this thesis that for both RQ1 and RQ2, using the non-standard measure of mean length of T-units provides a clearer base upon which to build a picture of how cognitive task complexity and pre-task planning time affect the mean length of T-units.

5.3.3 The effects of pre-task planning time and cognitive task complexity on lexical complexity

The second complexity measure used in this thesis was a mean segmental type-token ratio, which is a measure of lexical complexity. Overall, the findings for lexical complexity showed statistically non-significant increases ($p=.089$). When compared against the findings for RQ1 (see Table 37 in section 4.2.13), which showed a similar but statistically significant increase in lexical complexity as tasks became more complex, the results for RQ2 seem to imply that the addition of planning time has had an adverse effect on lexical complexity.

The mean scores for Task 1 are similar between RQ1 and RQ2. With both RQ1 and RQ2 producing similar amounts of lexical complexity when operating under the same conditions and with the same low level of complexity, the inference is that both groups should have a similar proficiency level.

In section 5.2.3 the statically significant increases in lexical complexity noted in RQ1 (which increased as syntactic complexity decreased) were posited as resulting from a trade-off between lexical complexity (less demanding) and syntactic complexity (more demanding). Participants utilised the less cognitively demanding lexical complexity option by using the increased cache of lexical items supplied in the task rubrics to meet the increasingly complex pragmatic requirements of the tasks.

The findings for lexical complexity in RQ2 are posited to follow a similar process, with students using the increasingly varied cache of lexical items as a way to meet the needs of the increasingly complex tasks; however, the drop in lexical complexity from statistically significant to statistically non-significant may be the results of another trade-off. In RQ2, the combination of limited attentional resources, and the focus on increased syntactic complexity facilitated by pre-task planning time during the formulating process may have impeded the automatic retrieval of lexical units to be used in the output.

Ong and Zhang (2010) have suggested that pre-task planning engages participants in online planning in a way that impedes lexical complexity, whereas tasks with no planning time prompt the automatic retrieval of lexical items by participants. It could be argued that the difference between the results for RQ1 and RQ2 might reflect this position, as lexical complexity is slightly lower where planning time is utilised.

To summarize, though most of the findings for dependent clauses in RQ2 are not statistically significant, there is an overall trend towards increases in syntactic complexity as the overall trend for lexical complexity slips below statistical significance.

It is possible that the increasingly demanding pragmatic requirements of Task 2 and Task 3 (when freed up by the facilitating effects of pre-task planning time) elicit the more complex and detailed subordinate structures as a way to meet those demands. However, due to the proposed limited nature of attentional resources, a trade-off takes place where the automatic retrieval of lexical items is gradually impeded during the formulating process as attention is directed to meeting the propositional needs of the tasks with subordination.

5.3.4 The inclusion of a low complexity task

As with RQ1, RQ2 explored the potential benefit of using a patently low complexity task (Task 1) in conjunction with the more complex tasks (Task 1 and Task 2) when testing for the effects of cognitive task complexity and pre-task planning time on complex output.

The reason for adding a manifestly low complexity task in RQ2 is the same as in RQ1. The low complexity task was added because cognitive task complexity cannot be accurately measured when increasing and decreasing complexity. As a result, the findings might be obscured if the tasks were either too complex or not complex enough. Therefore, it is the position of this thesis that adding a task with manifestly lower complexity would highlight variations in complex output obscured by inaccurate measuring of more complex tasks.

Additionally, when analysing the findings for RQ1, RQ2, and RQ3 between groups, the low complexity task was used as another method of gauging procedural skill levels between

the different groups of students who participated in each research question. It was believed that a low complexity task, which allowed participants to use their own resources to answer the question under the same task performance conditions, would elicit output that might indicate how close the participants were in terms of dependent clause procedural skill level.

The inclusion of Task 1 as a means of gauging procedural skill level raised issues not considered prior to the designing of this thesis. Though the results for Task 1 were similar for RQ1 and RQ2, inferring there was some parity in procedural skill levels, there was a pattern noticed in the findings. Every time Task 1 had a minor variation in dependent clause length that was lower than the other Task 1, the Task 1 with the lowest dependent clause scores always had a Task 2 and Task 3 that had lower dependent clause scores than the Tasks 2 and 3 associated with the higher Task 1 score.

Though this might appear reasonable given that Task 1 was supposed to help gauge procedural skill levels, the variations between Task 1-RQ1 and Task 1-RQ2 were often very small. It implied that if these findings did represent procedural skill levels, then very small variations in procedural skill levels might be having significant effects on the findings for Task 2 and Task 3 when combined with cognitive task complexity and pre-task planning time. If very minor variations in procedural skill levels are interacting with other variables to create significant results, then the process of accurately manipulating the variables to get the desired results becomes more complicated.

Looking at the findings for both syntactic and lexical complexity, there appears to be evidence that the inclusion of a patently lower complexity task might be helpful in clarifying the effects of cognitive task complexity on written output.

In all the analyses, Task 1 provided a useful counterpoint for analysing the more complex tasks (Task 2 and Task 3). This is especially the case in RQ2 for adjectival dependent clause length where the statistically significant variation was found between Task 1 and Task 3.

Earlier qualifications that the patently lower complex task might be a different task-type with different pragmatic requirements might still be a legitimate concern; however, the results for RQ2 may lessen the possibility of different task requirements. The strongest case for the notion of Task 1 being a different task-type was adverbial dependent clauses in RQ1, in which the production of adverbial dependent clauses seemed to clearly show a demarcation line in production between Task 1 and Tasks 2 and 3. However, when compared against the results for adverbial dependent clauses in RQ2, the increase in production of adverbial dependent clauses for Tasks 2 and 3, bring them closer to Task 1, appeared to be related more to the effects of levels of cognitive task complexity and the addition of pre-task planning time.

In conclusion, the addition of a patently low complexity task (Task 1) appears to have generated findings for RQ2 that might not have been noticeable had only the two most complex tasks been used in the collection and analysis of the data. Furthermore, using Task 1 as a means to gauge procedural skill levels raised the issue that minor variations in procedural skill level might combine with the independent variables to create significant effects. Finally, though the matter of Task 1 being a different task-type remains a possibility, comparing the results between RQ1 and RQ2 appears to support the notion that variations in output between Task 1 and Tasks 2 and 3 are related to issues of complexity and procedural skill level.

5.3.5 The findings and predictions about resource-directing and resource-dispersing elements

In this section, Robinsons' predictions for the effects of resource-directing and resource-dispersing elements are discussed in relation to the findings of this thesis. The independent variables, number of elements and reasoning demands, used in this thesis have been termed resource-directing variables by Robinson (2001a, 2001b, 2005, 2007a). It is Robinsons' position that increases in cognitive task complexity facilitated by increasing resource-directing variables should lead to increases in complex language.

Furthermore, planning time has been described by Robinson as a resource-dispersing element. Resource-dispersing elements are believed to have a detrimental effect on complex language production when they are increased in complexity. For example, increasing resource-dispersing complexity by removing planning time is predicted to negatively effect the production of complex language. Alternatively, decreasing complexity along resource-dispersing lines would involve adding planning time, which is predicted to facilitate complex language production.

Robinson (2001b, p.31) states that

> ….making a pedagogic task simple along resource–depleting lines (e.g. by allowing planning time) and complex along resource-directing dimension (e.g. by requiring reasoning) allows optimum resource allocation to satisfy the linguistic demands of the task, compared to making the task more complex on both dimensions simultaneously.

This thesis has increased complexity along resource-directing lines by increasing the number of elements and reasoning demands between the task in RQ1, RQ2 and RQ3.

However, it is only in RQ2 that resource-dispersing complexity is decreased by adding 10 minutes pre-task planning time to the conditions under which Tasks 2 and 3 were performed.

The findings from this thesis, when viewed through the lens of Robinsons' theory, seem to show statistically non-significant support for Robinsons' predictions regarding the combination of resource-directing and resource-dispersing dimensions when the results of RQ2 were compared against those of RQ1. Support for Robinsons' predictions could be inferred from the one statistically significant result that emerged when the syntactic complexity findings for RQ2 were analysed using the non-standard method of separating dependent clauses during the analysis.

However, it should be noted that the predictions made by Robinson were not borne out for lexical complexity in RQ2 in terms of statistical significance. Whereas RQ1 showed significant increases in lexical complexity across Tasks 1, 2, and 3 when only the resource-directing dimension was used, the inclusion of the complexity facilitating element pre-task planning time (less resource-dispersing) in RQ2 resulted in decreases in lexical complexity to below statistical significance.

This thesis takes the position that the findings mentioned above represent partial support for Robinsons' theory. The increases in dependent clause production in RQ2 seem consistent with predictions for the combination of resource-directing and resource-dispersing elements on complex output. However, since neither fluency nor accuracy was tested concurrently, this thesis cannot speculate on Robinsons' assertion that more than one aspect of language production can be increased simultaneously. In addition, the decreases in lexical

complexity in the findings for RQ2 are speculated as possibly being the result of a trade-off in types of complex production resulting from a limited attentional capacity.

As in the results for RQ1, this thesis takes the position that both the Cognition Hypothesis and the Limited Attentional Capacity Model have a part to play in explaining the results. Robinson's predictions for the effects of resource-directing and resource-dispersing elements appear to work within the confines of the type of limited attentional capacity processing posited by Skehan, where trade-offs in production are required when limits are reached. In this case, the trade-off is between types of complexity as opposed to complexity and accuracy.

5.3.6 The effects of modality

Finally, the effect of modality is considered in relation to the findings for RQ2. The positive effects for cognitive task complexity and pre-task planning time on syntax, noted in the findings above, could be interpreted as giving support to the notion that the written modality by itself is not always sufficient to promote complex output under cognitive complex pressure. However, it is possible that the written modality works best with other facilitative elements (like planning time) to create the type of favourable conditions needed for increases in syntactic complexity. Previous studies in the written modality with positive outcomes for syntactic complexity (Ellis & Yuan, 2004; Ishikawa, 2006) also utilized pre-task planning time.

It is possible that the demands of formulating syntactic structures, which intrinsically place a cognitive burden on memory capacity (Cheung & Kemper, 1992; Lord, 2002), while processing the increasingly complex pragmatic requirements of complex tasks, requires the

alignment of a few key elements. Both the attention reliving effects of structuring work before the task provided by pre-task planning and the proposed inherent facilitative aspects of the written modality may well work synergistically with the procedural skill level of the subordinate clauses being utilised by the writer.

Kormos and Trebits (2012) stated that the written modality might provide more available attentional resources to monitor what is being produced during the encoding (formulation) process, where thoughts are transformed into words and syntactic structures. However, the extra resources provided by this process may not be sufficient when the added pressure of cognitive task complexity and low procedural skill levels are adding pressure to limited attentional resources. The reliving of attentional pressure created by the pre-structuring during pre-task planning time (meaning leaners do not have to hold information in the memory) may work in conjunction with relevant procedural skill levels to allow participants to address the propositional content of the task by using syntactic structures.

Finally, it is difficult to say with any certainty whether mode has significantly contributed to the positive results found in this thesis; however, when considering the positive effects for syntactical complexity found in RQ2 alongside the proposed benefits of the written modality and pre-task planning time, it is possible to make a cautious defence of the written modality as one constituent facilitative element in the development of complexity.

5.3.7 Summary of the discussion for RQ2

The discussion on RQ2 addressed the effects of cognitive task complexity and 10 minutes pre-task planning time on syntactic and lexical complexity. It also covered the additional relevant issues addressed in RQ1. In RQ2, mode was suggested as a potential contributor to

increased complex output, while once again the patently lower complexity task highlighted variations in complex output and procedural skill levels.

The effects of increased cognitive task complexity and pre-task planning time on the standard measure of mean length of T-units revealed non-significant increases as tasks became more complex. However, the non-standard measure of T-units revealed more detailed variations for each dependent clause, including one statistically significant result. Increases in the mean length of T-units were found both within RQ2 and in comparing results between RQ2 and RQ1. As with RQ1, the interplay of procedural skill levels, the inability to accurately measure cognitive task complexity, and pragmatic task requirements were viewed as contributing to the findings

These results appeared to partially support the Cognition hypothesis, with increases along resource-directing dimensions and decreases along resource-dispersing dimensions leading to some increases in syntactic complexity. Conversely, RQ2's lexical complexity, which showed non-statistically significant increases, was less than in RQ1. This was attributed to limited attentional resources being focused more towards syntactic production.

The results appeared to provide limited support for the Cognition hypothesis, which was posited as operating within a limited attentional capacity and subject to the delicate alignment of procedural skill levels, amounts of cognitive task complexity, and the appropriate pragmatic task requirements.

5.4 RQ3: What are the effects of post-task editing combined with cognitive task complexity on written complexity?

5.4.1 Overview

RQ3 investigated the relationship between increases in cognitive task complexity (reasoning demands and number of elements), and 10 minutes unguided post-task editing time on complex written output, measured as syntactic and lexical complexity.

Syntactic and lexical complexity were analysed in the same way as RQ1 and RQ2 with both sets of findings subjected to a repeated measures ANOVA as well post hoc tests whenever further in-depth analyses were required. Additionally, the results from RQ3 were subjected to a mixed between-within subjects ANOVA so that the overall results could be compared with the results from RQ1 and RQ2.

Had the results from RQ3 been significant, the supplementary issues discussed in RQ1 and RQ2 would also have been discussed in relation to the results; however, as clarified in the following section, the findings for RQ3 were not conducive to further discussion.

5.4.2 Mean length of T-units across all dependent clauses

The results for RQ3, where the mean length of T-units was measured using dependent clauses as one group, revealed that increases in cognitive task complexity in conjunction with 10 minutes post-task editing time appeared to result in no significant variation ($p= .33$) in the mean length of T-units.

Task 1 (low complexity), which had no post-task editing time, had marginally more T-unit depth than both Task 2 (medium complexity) and Task 3 (high complexity), which both had post-task editing time. The variation in mean scores between Task 2 and Task 3 was also slight.

These results indicate that there was no positive effect for post-task editing time on the mean length of T-units; however, these findings should be considered with caution when ascribing the low level of T-unit output between tasks to editing time because a low level of procedural skill may have influenced the findings. The post hoc test for the mixed between-within subjects ANOVA (see section 4.2.14) showed that the overall mean score for RQ3 was significantly lower than those of both RQ1 and RQ2.

When viewing the means scores for each task from each research question (see Table 33 in section 4.2.13), it is clear that the mean scores for Tasks 1, 2, and 3 for RQ3 were lower than Tasks 1, 2, and 3 for both RQ1 and RQ2.

Had the significantly lower means scores found in RQ3 only been located in Task 2 and Task 3, then the possibility that the additional independent variable (post-task editing time) had contributed to the low scores would have been considered. However, the significantly lower mean scores extended to Task 1, suggesting that some other mediating factor may have affected the results for RQ3.

Considering the minor variation in T-unit length within RQ3 between Tasks 1, 2, and 3, and RQ3's significantly lower production of T-unit length compared to both RQ1 and RQ2,

it is possible that the participants of RQ3 might have had a lower level of proficiency than the participants of RQ 1 and RQ2. As a result, these findings do not appear to contribute anything to the understanding of the effects of post-task editing time and cognitive task complexity on T-unit length.

5.4.3 The effects of post-task editing time and cognitive task complexity on lexical complexity

As with RQ1 and RQ2, the second complexity measure used for RQ3 was a mean segmental type-token ratio, which is a measure of lexical complexity. Similarly to the results for T-unit length, the findings for RQ3 showed no significant variation ($p=.67$) between Tasks 1, Task 2, and Task 3.

Task 1 (low complexity), which had no post-task editing time, had marginally higher mean segmental type-token ratio than Task 2 (medium complexity) and Task 3 (high complexity), which both had post-task editing time. The difference variation Task 2 and Task 3 was not significant.

As with the results for syntactic complexity above, these findings should be considered with some caution when ascribing the low level of lexical output between Tasks 1, 2, and 3 to post-task editing time, as these scores were potentially influenced by the same lower proficiency levels ascribed to the participants of RQ3 that also affected the syntactic results. Table 37 (in section 4.2.13) clearly shows that the mean scores for Tasks 1, 2, and 3 for RQ3 were lower than Tasks 1, 2, and 3 for both RQ1 and RQ2.

As with the results for syntactic complexity, the post hoc test for the mixed between-within subjects ANOVA (see section 4.2.14.2) showed that the mean segmental type-token ratio scores for RQ3 were significantly lower than those of both RQ1 and RQ2.

The low performance of RQ3 across all tasks compared to RQ1 and RQ2, and the small variation in lexical variety between Tasks 1, 2, and 3 within RQ3 appears consistent with the finding for syntactic complexity in RQ3. Comparatively low proficiency levels are likely a major contributing factor for both of these findings. As a result, these results do not appear to contribute to the understanding of the effects of cognitive task complexity and post-task planning time on complex output.

5.4.4 Summary of the discussion for RQ3

The results for RQ3 have mostly been dismissed as providing no real evidence of the effects of cognitive task complexity and post-task editing time on syntactic and lexical complexity due to problems with the proficiency levels of the participants.

Though the participants in RQ1, RQ2, and RQ3 all came from different schools, each school claimed that the students were at an intermediate proficiency level, with all students having IELTS levels ranging from 4.5 to 5.5. However, the performance of the participants in RQ3 appears to have been the result of comparatively lower proficiency as illustrated by the patently lower complexity task (Task 1), which was used as an additional means of gauging procedural skill levels.

5.5 RQ4: What is the relationship between the participants' attitudes and complex written output?

5.5.1 Overview

RQ1, RQ2, and RQ3 focused on the causal relationships between task input and performance conditions on written output, with subsequent discussions attmepting to provide credible explanations for the causes and effects. However, this thesis acknowledges that the proposed cause and effect relationships noted between task input and performance conditions on written output do not happen in isolation.

As such, this section endeavours to widen the scope of the investigation into potential influences mediating the effect of cognitive task complexity on complex written output by investigating the phenmoena of task motivation. Task motivation includes the motivational effects resulting from students' reactions to the characteristics of tasks, which are predicted to affect the amount of commitment a student might make to any given task (Dörnyei, 2002, 2005, 2009).

RQ4 investigates whether the phenomena of task motivation had any noticeable correlation with the production of cognitively demanding language, in this case the production of mean length of T-units. Using Likert scale questionnaires, correlations were sought between the attitudes of all 94 participants (from RQ1, RQ2, and RQ3) towards the letter-writing tasks and their T-unit production in Tasks 1, 2 and 3.

Of the 11 questions answered in the Likert scale questionnaire for RQ4, only three questions had statistically significant correlations; as a result, only those three questions are

discussed here. LQ4 (Likert question 4) addressed the task motivation dimension of task relevance (Crookes & Schimdt, 1991; Dörnyei, 1994; & Keller, 1983), which is the degree to which class instruction appears related to a students' goals, needs, and values. LQ6 (Likert question 6) and LQ8 (Likert question 8) addressed the task motivation dimension of task expectancy (Crookes & Schimdt, 1991; Dörnyei, 1994; & Keller, 1983), which is a students' perception of the amount of effort required for the successful completion of a task.

5.5.2 LQ4: I understood how these tasks were supposed to help me learn English

In LQ4, the association between mean length of T-units and Likert question 4 was explored using a Pearson product-moment correlation. The results showed a negative correlation between decreases in the mean length of T-units and increases in perceptions of task relevance across all three tasks. This suggests that the mean length of T-units became shorter as the participants became more aware of how the task was supposed to help with the learning of English.

As with all the correlational data analysed in this thesis, the correlation of the two variables does not represent a cause and effect relationship. The correlation only shows that this association exists between the two variables. As such, the following discussion infers causality between the two variables only as a means of discussing how the relationship between task construal and performance might work.

Additionally, when referring to a task as blueprint in this section, the reference is to the distinction between a task as a behavioural blueprint, which is the same for each

participant, as opposed to the activity elicited by the task, which may be different for each individual (Coughlan and Duff, 1994).

There are posited to be a number of factors that might influence a participants' reaction to a task as behavioural blueprint other than the cause and effect relationship suggested in RQ1, and RQ2. A participants' response to a task might vary based on individual learner factors like anxiety, working memory capacity, motivation, and aptitude (Robertson & Gilabert, 2007). Additionally, the motives driving task responses can be affected by a variety of social or biological influences (Lantolf & Appel, 1994; Leontiev, 1981).

Viewed together, it appears that the behavior/output generated by a task may be influenced by a number of external and internal factors that can either be common for each participant or they can vary between participants in any particular group that is performing the same task.

In terms of task motivation, it seems that any effect apportioned to a task (as behavioural blueprint) on levels of motivation could be partly dependent on how the task is construed and that construal could be informed by a combination of the factors mentioned above. The numerous contributing factors suggests that there may be noticeable variability within groups when studying task motivation, something that appears to be borne out by results of the correlations in which variations between task construal and output have only correlated with a small percentage of the overall group.

Focusing specifically on the small percentage of participants who correlated task construal and mean length of T-unit production, there initially appears to be an intuitive contradiction in the findings.

Perceiving a tasks' educational utility seems like the type of positive task construal that might lead to increases in motivation, which is posited to induce temporary increases in resources (Robinson, 2001b) and contribute to improved output. However, when viewing the result for LQ4, there appears to be a contradiction between the idea that understanding a task's utility might contribute to increased motivation and performance, and the findings, which correlate understanding of the tasks utility with decreases in production (theorized as resulting from decreases in motivation).

Regarding the perceived contradiction between a positive construal of the task and decreases in performance noted in the correlation of LQ4, it is possible that the small percentage of participants, who were captured by the correlation, may have believed that they understood the tasks educational utility. However, they may have viewed the perceived educational utility of the task as being irrelevant to their current goals, needs, or values.

The negative construal of utility could be based on any number of contributing factors; for example, it could be social in nature having been based on past exposure to letter-writing tasks in classroom environments, poor interaction with teachers in the form of inadequate teaching, or the general attitude towards this type of tasks amongst the group of participants.

With a prior negative construal of the educational utility of the task already set in the participants' mind, the bias may initially take effect at the execution phase of Dörnyeis', (2003) dynamic three-part cycle of motivation, as the execution phase is where the task is initially engaged. Once in effect, the negative construal of the task as behavioural blueprint,

in conjunction with the stresses placed on current procedural skill levels, might have a corrosive effect on the dynamic motivational system, affecting the subsequent appraisal and action control parts of the process.

The multiple drains on attentional resources, which are in effect due to cognitive task complexity and the requirements of subordination in the formulation process (which Kellog (1996) hypothesises places the most burden on working memory), may place extra pressure on the action control mechanism (the self-regulating mechanism needed to maintain engagement during the task execution stage). The addition of negative opinions of the writing tasks, initiated during the execution phase, could lower the available resources being used to maintain engagement with the task during the formulation part of the writing process.

At this point, it might be considered that the behaviour being generated by the task is operating under the combined influence of utility bias, stresses placed on the action control dimension of motivation, all of which happen during cognitively stressful formulating, which is already strained by the demands and procedural levels of the learners.

It is worth noting that the largest percentage of participants with a negative correlation correlated on Task 1, which was the lowest complexity task. Some of these participants, who had the negative construal, went on to increase their production on the more complex tasks. This is why the percentage of participants whose correlation between task construal and decreases in performance drops in Task 2 and Task 3.

This suggests that the proposed negative effects of task construal on Dörnyeis'(2003) three-part cycle of motivation, and processing pressure during formulation, may not have carried over to the more complex tasks, where the perception of utility did not have such a strong relationship with decreasing motivation.

Motivation can be viewed as being constantly reappraised as it is subjected to influence both internal and external (Ellis & Larsen-Freeman, 2006), with learners having variations in their levels of commitment even during the course of one class (Dörnyei & Skehan, 2003). As a result, it is possible that the effects of negative task construal on motivation varied between tasks as the participants' ability to sustain motivation fluctuated from negative to positive for the approximate 5% of participants who initially correlated on Task 1, but didn't correlate with decreases in performance on Task 2 and Task 3 (see Table 44 in section 4.3.6).

This fluctuation suggests that task construal may have a variable effect on individuals who can change during a task, and possibly between tasks, making predictions about the effects difficult. Additionally, there may have been participants in this group who had a negative task construal which never affected or correlated with decreases in performance, reinforcing the point made earlier about the variability of any proposed potential cause and effect for task motivation on output.

5.5.3 LQ6: When tasks became difficult, I lost interest in completing them

LQ6 explored the association between mean length of T-units and Likert question 6 using a Pearson product-moment correlation. The results showed a negative correlation between decreases in the mean length of T-units and increases in perceptions of task expectancy

across Task 2 and Task 3, suggesting that T-units became shorter as participants lost interest in completing them.

When conceptualising the relationship between decreases in T-unit production and the belief that increases in task difficulty affected interest in task completion, the apparent contradiction initially noted in LQ4 is not so apparent in LQ6. The correlation of decreases in motivation coinciding with decreases in task performance seen in LQ6 appears consistent with this thesis' previous assertion regarding the potential effects of decreases in motivation on syntactic output.

Table 44 (in section 4.3.6) illustrates that, for the small percentage of participants captured by the analysis, the correlation between perceived losses of interest and decreases in performance occurred on the Tasks 2 and 3, which were the most cognitively complex tasks.

Though the construal of the task could be based on similar past or present factors mentioned in LQ4, theoretically, the LQ6 correlation between participant attitude and the T-unit decreases fits into one of this thesis' general paradigmatic views regarding the affects of cognitive complexity. What is meant by paradigmatic view is the theoretical assumption that increases in cognitive task complexity can cause decreases in complex output (Skehan, 1998, 2003; Skehan & Foster, 1999, 2001). Subsequently, the negative construal of the task could be wholly, or partially formed as a result of the stresses placed on intermediate level participants while they are trying to formulate dependent clauses during a cognitively complex task.

Taking into account that the correlation between the two variables fits one of the theoretical assumptions underlying the results for RQ1, it is interesting to note the large amount of participants from RQ1 that were not part of the correlation. The participants from RQ1 alone constitute approximately 35% of the total sample group, which is far more than the approximate 5% (see Table 44 in section 4.3.6) who stated that the perceived difficulty of the task (as behavioural blueprint) affected their interest in completion.

The overall result for T-unit production in RQ1 was one of decreases in the mean-length of T-units as tasks became more complex; therefore, the findings for LQ6 suggest that, similarly to LQ4, decrements in performance might not always be connected to negative construal of the tasks, and task motivation may vary between participants that are producing similar results.

For those participants who did correlate decreases in interest (and potentially declining motivation) with decreases in performance, the processes involved may be much the same as those noted in LQ4, where negative task construal may manifest as decreases in motivation.

These decreases in motivation may reduce the levels of motivation available during action control, the self-regulating mechanism needed to maintain engagement during the task execution stage (Dörnyei, 2003), which will already be operating under cognitive stress due to the combined effects of increases in cognitive task complexity and procedural skill levels on subordination during the formulating process.

Finally, when trying to understand the relationship between the participants' answers for LQ6 and mean length of T-unit production, it is worth considering that the participants might have interpreted LQ6 differently from the expectations of this thesis. Describing the tasks as "difficult" may have had different interpretations for the participants answering the questionnaire than that intended by the researcher. The suggestions made above about a potential cause and effect relationship between the participant's perception of difficulty and syntactic output is partly based on the notion that participants will infer the meaning of "difficult" as somehow synonymous with increases in cognitive task complexity. However, there is no way to know for sure how the participants actually interpreted the meaning of difficulty in relation to the task as blueprint and their own individual learner factors.

5.5.4 LQ8: These tasks had too many things to concentrate on

In LQ8, the association between mean length of T-units and Likert question 8 was explored using the Pearson product-moment correlation. The results showed a positive correlation between increases in the mean length of T-units and increases in negative perceptions of task expectancy across all three tasks, suggesting that T-units became longer as the participants believed there were too many item to concentrate on.

The findings for LQ8 are the only results in which a positive outcome was found in the correlation. However, much like the findings for LQ4, this could be viewed as counter intuitive, or not fitting the paradigmatic view, given that the assumption of difficulty (with its potential drain on motivation) has not correlated with decreases in production.

Ostensibly, for LQ8 the negative construal of the task as blueprint does not seem to be tied to motivation depleting drives in the participants. There are a number of potential reasons

why a negative construal may not have had any noticeably corrosive effect on the motivational process.

It is possible that the participants, who are represented in the correlation, construed the task without any motivational attachment to their opinion because not every negative construal need necessarily produce a cognitive burden. If the proposed relationship between a negatively construed task and the resources depleting act of formulating subordination during the action control phase (as suggested in LQ4) is the same here, then the task construal suggested by LQ8 does not appear to have affected the sustaining of motivation during formulation.

Alternatively, the apparent lack of effect on motivation by the negative task construal might represent something about the causal nature of the relationship between task construal and performance. The inference is that task construal effects motivation from the moment the participant engages the task, or during the task performance.

Potentially, any construal of the task might be partially or fully formed after the task was completed. The task construal may be formed by the participant reflecting back on how difficult it was to complete once it was finished, but the construal may not be representative of the state of mind held during the performance. Motivation has been posited as having two causal relationships to learning. In one sense it is viewed as being a strong support for learning, in another sense it is viewed as being the result of learning (Ellis, 2008). In this case, the task construal may be a past reflection, and any subsequent motivational affects might not manifest until future performances of similar tasks now that the negative construal has been made.

Finally, the relationship shown in the correlation for LQ8 might be viewed as consistent with Robinsons' Cognition Hypothesis where increases in complexity are met with increased output. The apparent counter intuitive relationship between the construal of the task blueprint and the increase in complex output may just be the considered reflection of the types of cognitively stressful processing required to perform well on increasingly complex tasks. Like a runner who exhausts himself to win a race and then claims the race was hard. The claims represent the effort required to win, not a negative motivational burden.

As mentioned in LQ6, it is also interesting to consider the participants who did not figure in the correlation. RQ2, in which increases in subordination were noted under cognitive task complexity and pre-task planning time, contains about 35% of the participants, while the percentage in the correlation for LQ8 never reaches above 6.8% (see Table 44 in section 4.3.6). This suggests that a large group of participants who produced increases in mean length of T-units recorded a variety of types of task construal across similar performances. This would support the view that the relationship between task construal (as an element of task motivation) and performance may be widely variable within groups, making strong statements about the relationship from these results difficult.

5.5.5 Summary of the Likert scale questionnaire discussion

The items in the Likert scale questionnaires are theorized to represent different elements of the participants' task construal, which is viewed in this thesis as being a dimension of task motivation.

Task construal is thought to be variable in cause and effect; for example, proposed causes include social influences (past and present) and current procedural skill levels. Additionally, the disparate ways the Likert items can be interpreted by individuals might also affect any causal analysis of task construal an output. The proposed effects of task construal on performance included decreases in motivation resulting in decreases in the mean length of T-units, or no effects at all if the task construal had no cognitive burden attached to it. Finally, adding to the variability issues (which make defining causes and effects between task construal and performance difficult) was the possibility that the relationship between task construal and production can change during the course of one session across tasks of different cognitive task complexity.

While acknowledging the potential effects of the variability issues, this study interpreted the significant findings for the three Likert questions as follows. For LQ4, a negative construal of the educational utility of the tasks was posited to have a negative impact on motivation and output. Furthermore, for LQ6, a negative construal of maintenance of interest due to increasing amounts of items in the tasks was posited to have a negative impact on motivation and production. However, for LQ8, a negative construal of the amount of items in the task was theorized to have either a positive effect on the motivation and language production of some participants, or it had no effect on their motivation or their ability to produce complex language.

Where a relationship exists between task construal and variations in motivation, the effects might be more noticeable during the attention demanding formulation of dependent clauses under increasing cognitive task complexity. Negative task construal may create a corrosive effect, reducing the maintenance of attentional resources across the motivational task

281

processing system (Dörnyei, 2002, 2003, 2009, 2010), with a flow on effect negatively impacting the task execution, and task appraisal dimensions.

5.6 Conclusion of the discussion chapter

Four research questions were the focus of this thesis; however, only three were answered in detail. RQ3 was not discussed, as the findings seemed to show a lower competence level compared to RQ1 and RQ2, thus rendering the findings unusable. As a result, any significant analysis for post-task editing time has not been addressed.

Overall, the non-standard measure of T-units in conjunction with the use of a patently non-complex task revealed detailed findings on the effects of increasing cognitive task complexity on syntactic complexity. It appears that planning time in RQ2 had a facilitating effect on increasing mean length of T-units, whereas RQ1, which had no planning time, mostly showed decreases. Without pre-task planning time, the attentional pressure appears too high to formulate increasing amounts of subordinate in order to meet the increasing pragmatic task requirements. However, lexical complexity, which significantly increased in RQ1 but increased without significance in RQ2, appeared to decrease as limited attentional resources were applied to formulating subordinate clauses.

Mode did not appear to be a noticeable contributing factor in the production of complex language in RQ1; however, it may have been one contributing factor when applied in conjunction with pre-task planning time in RQ2.

The syntactic findings for RQ2 appeared to partially support the Cognition hypothesis, with increases along resource-directing dimensions and decreases along resource-dispersing

dimensions leading to increases in the mean length of T-units, in both within group and between group analyses. However, caution is advised when attributing the positive findings for lexical complexity in RQ1 to predictions by Robinson because the significant findings were attributed to increases in lexical items in the task instructions, as opposed to any extra mental effort induced by cognitive task complexity. Despite the disparate results for lexical and syntactic complexity, this thesis suggests that the findings give limited support for the Cognition hypothesis, which is posited as operating within a limited attentional capacity and subject to the delicate alignment of procedural skill levels, amounts of cognitive task complexity, and the appropriate pragmatic task requirements.

Finally, the results for RQ4 widened the scope of potential elements effecting the production of mean length of T-units in RQ1 and RQ2. Task construal (as a dimension of task motivation) is viewed as a variable phenomenon both in cause and effect. Where a potential relationship exists between task construal and variations in motivation, the effects were theorized to be potentially more noticeable during attention demanding processes like the formulation of dependent clauses under increasing cognitive task complexity pressure. Furthermore, this might have a corrosive effect, reducing the maintenance of attentional resources across the motivational task processing system (Dörnyei, 2002, 2003, 2009, 2010).

CHAPTER SIX

CONCLUSION

6.1 Overview

In this chapter, the findings from chapter 5 are discussed in relation to their significance to contributions to theory, research, and pedagogy. Section 6.2 is a reminder of the central aims of this thesis and provides a detailed summary of the findings from each of the research questions. Section 6.3 presents the contributions of this thesis, starting first with contributions to theory (6.3.1). This is followed by contributions to research (6.3.2) and pedagogy (6.3.3). Section 6.4 discusses the limitations of this thesis and section 6.5 recommends areas of future research. This chapter concludes with some final remarks in section 6.6.

6.2 Summary of findings

The central aim of this study was to explore the relationship between increases in cognitive task complexity on complex written output. Further to this central objective was the addition of two extra variables, pre-task planning time and post-task editing, which were added in order to gauge their additional effects on the production on complex written output. In addition to the investigation of cognitive task complexity on complex writing was the exploration of the potential effects of task motivation (as task construal) and any potential consequences for the production of complex written output.

A number of supplementary issues were considered in relation to the main objectives of this thesis. These subsidiary elements were considered factors with potential implications for the relationship between the central variables. The subsidiary elements, considered in conjunction with the research questions, were the effect of mode, a non-standard mean length of T-units measurement, the application of a patently lower complexity task, and the findings relationship to Robinsons' predictions for resource-directing and recourse-dispersing elements.

The data was collected in four language schools in Auckland city. The population from which the samples were obtained consisted of a variety of non-native speakers of English who were studying English in New Zealand. All the participants were classified as intermediate level language learners by the schools in which they were enrolled with IELTS levels ranging from 4.5 to 5.5.

This thesis utilised a positivist approach, utilising two types of quantitative research: experimental and associational. The experimental approach, which looked for cause and effect between variables, was utilized to answer RQ1, RQ2, and RQ3. The results were analysed within groups, using a repeated measures ANOVA; and between groups, using a mixed between-within subjects ANOVA. The associational approach, which looks for relationships between variables without specifying cause and effect, was used for RQ4. The results from RQ4 were analysed using a Pearson product-moment correlation.

RQ1: What are the effects of cognitive task complexity on written complexity?

A number of conclusions were extrapolated from RQ1's findings. Firstly, mode did not seem to be a contributing factor in the production of complex language. Secondly, this

study found a patently lower complexity task useful for highlighting differences in complex output. However, questions were raised regarding whether the low complexity task used in this research was a different task type.

Mean length of T-units were analysed in two ways. The use of a standard measure of mean length of T-unit, where all dependent clauses were measured as one group, indicated that increases in cognitive task complexity generated non-significant degradations in dependent clauses. However, viewing the results at the individual dependent clause level (using the non-standard measure where dependent clauses were analysed separately) revealed a multiplicity of effects and potential causes, operating in conjunction with cognitive task complexity.

The level of procedural skill of each dependent clause, the current inability to accurately and incrementally manipulate cognitive task complexity, and the dependent-clause-specific pragmatic requirements of a task were all suggested as central components affecting complex syntactic output. It was noted that these elements would also be affected by other factors not addressed specifically in RQ1, such as planning time, the social aspects of language learning, and other individual difference factors.

Contrary to the results for syntactic complexity, the effects of cognitive task complexity did appear to have a positive effect on the mean segmental type-token ratio, though this thesis surmises that this was not due to any extra mental effort induced in the participants by cognitive task complexity. It has been posited that the increases in lexical complexity may have been the result of participants choosing an easier means of expression by taking lexical items from the increasingly varied items available in the task instructions. This was

viewed as a less demanding means of expressing complexity than the more demanding syntactic production, which mostly deteriorated as the tasks became more complex.

Finally, although the results for syntactic complexity mostly showed overall degradations in dependent clauses, there was one incident where a non-significant increase was detected. This thesis speculates that had the multiple elements posited to affect T-units length, in conjunction with cognitive task complexity, been favourably aligned, predictions for the Cognition Hypothesis may have been borne out. Eliciting favourable results for T-unit length is viewed as a complicated endeavour requiring, in part, the alignment of the participant's procedural skill levels, the correct dependent-clause-specific pragmatic requirements of a task, and a more precise way to manipulate cognitive task complexity between tasks.

RQ2: What are the effects of pre-task planning time combined with cognitive task complexity on written complexity?

A number of conclusions were extrapolated from the findings for RQ2, both within RQ2 and between RQ2 and RQ1. Firstly, mode may have been a contributing factor in the production of complex language, with the findings for RQ2 having some similarities to other work in the written modality that also utilized pre-task planning time. It is the position of this thesis that the written modality may function as one constituent part of a group of elements, like pre-task planning time and appropriate procedural skill levels. These elements may combine to facilitate the production on complex language under cognitively complex conditions.

Regarding the use of Task 1 (low complexity), RQ2 showed that there was utility in using a manifestly lower complexity task to analyse complexity in this thesis. This thesis found Task 1 useful for highlighting differences in complex output that might not have been noticed between the two more complex tasks. Task 1 also brought to light questions about how very small difference in procedural skill levels might have significant impact on complex output produced under cognitively complex conditions.

The use of two different measures of T-unit length also proved worthwhile. The standard measure of mean length of T-unit, where all dependent clauses were measured as one group, indicated that increases in cognitive task complexity and pre-task planning time mostly led to non-significant increases in dependent clauses. However, this result did not give a clear picture of the various effects of cognitive task complexity and pre-task planning time on the mean length of T-units.

The findings for RQ2 on syntactic complexity, when viewed at the individual dependent clause level, showed various effects for cognitive task complexity and pre-task planning time on dependent clauses, with the exposing of one statistically significant result not apparent when using the standard measure. The non-standard measure of T-unit length revealed more detailed findings that appeared to show the how the combination of varying levels of cognitive task complexity, pre-task planning time, and the procedural skill level of the individual dependent clauses might be some of the key elements contributing to overall production of mean length of T-units.

Contrary to the results for syntactic complexity, the effects of cognitive task complexity and pre-task planning time did not appear to have a positive effect on lexical variation

when measured as a mean segmental type-token ratio. Though the trend in RQ2 was still positive, with lexical variation increasing from Task 1 (low complexity) through to Task 3 (high complexity), the mean scores fell below statistical significance, something that this thesis attributes to a trade-off between complex outputs. The trade-off between complex outputs was speculated as being the results of limited attentional resources being allocated to the formulation of syntactic structures, which were facilitated by the introduction of pre-task planning time.

To summarize, the results for RQ2 suggest that the development of complex language through the increasing of cognitive task complexity and pre-task planning time is possible but complicated, requiring the combination and balance of a number of key elements. In this respect, the conclusion is similar to the one reached in RQ1.

RQ3: What are the effects of post-task editing combined with cognitive task complexity on written complexity?

The results for RQ3 have mostly been dismissed as providing no real evidence of the effects of the independent variables on the dependent variables due to an apparent problem with the proficiency levels of the participants. Though the participants in RQ1, RQ2, and RQ3 all came from different schools, each school claimed that the students were at an intermediate proficiency level, with all students having IELTS levels ranging from 4.5 to 5.5.

However, the performance of the participants of RQ3 appears to have been the result of lower procedural skill levels than the other two groups as illustrated by the low

performance on Task 1 (low complexity), which was used as an additional means of gauging proficiency (procedural skill level).

It has been suggested in this thesis that low procedural skill levels can leave participants susceptible to cognitive overload when cognitive task complexity is increased; and as a result, the participants may fall back upon using the default level of production associated with using their own resources. However, even Task 1, which was designed to elicit the default level of production associated with the use of individual resources, was unable to extract a level of lexical and syntactic complexity from the students that was mostly consistent with RQ1 and RQ2.

It is assumed that there was some difference in the way the proficiency of the students in RQ3 was gauged by the school from which they were selected compared to the schools that supplied the participants for RQ1 and RQ2.

RQ4: What is the relationship between the participant's attitudes and complex written output?

The correlation between the attitudes in the Likert scale questionnaire and the performance of T-units does not signify any causal relationship; however, the results were discussed in terms of causality as a way to present potential reasons for the effects of task motivation.

The proposed effects of task motivation have been discussed in this thesis as partially resulting from the effects task construal has on output. A wide range of factors likely affects the relationship between the construal of a task and task output; for example, task construal might result from various underlying factors, such as social influences (past and

present), current procedural skill levels, or a combination of both. As well as having variation in causes, task construal may also have various effects on performance, such as decreasing complex output by lowering motivation levels during attention demanding activities, or having no effects at all if the task construal has no cognitive burden attached to it.

Other issues arose during the discussion of RQ4 that suggest caution must be used when interpreting the relationship between task construal and output. For example, there can be a difference between what the Likert scale questionnaires items are designed to express and the distinctive ways in which the individuals interpret the items. As a result, understanding what is being correlated between the Likert scale questionnaire items and the output is not always going to be clear. Additionally, there is no assurance that a causal relationship always exists between the task construal and the output, with the construal of a task potentially manifesting before, during, or on completion of a task.

What is clear from trying to interpret the findings for RQ4 is that variability seems to be a key element of task motivation, with variations in the causes and effects of task construal needing to be considered along with variations of when construal towards a task might actually manifest. These issues also need to be considered alongside potential variations in how the Likert scale questionnaire is interpreted by participants; and finally, there is the potential for the relationship between task construal and production to change during the course of one session across tasks of different cognitive task complexity.

However, despite the issues mentioned above, some cautious claims are made regarding potential reasons for the correlations noted in the findings. For LQ4, a negative construal of

the educational utility of the tasks was posited to have a negative impact on the motivation of some participants. For LQ6, a negative construal of an individuals' ability to maintain interest given the amount of items in the task was also posited to have a negative impact on the motivation of some participants. Further to these claims, suggestions were made as to how these negative impacts may have affected output.

It has been suggested in this thesis that where a relationship exists between a task construal and variations in motivation, it is possible that this relationship might have a noticeable effect on the type of production that intrinsically places a burden on limited intentional resources, such as the formulation of dependent clauses under cognitive task complexity. A negative task construal may affect the action control process, which may already be under pressure to maintain motivation during formulation where participants are working close to their level of ability. Negative task construal may create a corrosive effect that reduces the maintenance of attentional resources across the motivational task processing system (Dörnyei, 2002, 2003, 2009, 2010), with a flow-on effect negatively impacting the task execution, and task appraisal dimensions.

For LQ8, a negative construal of the amount of items in the task was posited to have had either a positive effect, or no effect, on the motivation of some participants. For some participants, the negative construal may have had no negative effect, if the construal was a retrospective reflection on the difficulty involved in completing the task.

6.3 Contributions of this study

This section discusses the proposed contributions of this thesis. Section 6.3.1 discusses the contributions to theory. Following that, section 6.3.2 illustrates the contributions to research, and finally, 6.3.3 suggests the contributions to pedagogy.

6.3.1 Theory

It is the position of this thesis that the findings can be interpreted as giving cautious support to both the Cognition Hypothesis and the Limited Attentional Capacity Model. However, in the case of the Cognition Hypothesis, this support is based on excluding the notion of extended resource limits, and including the possibility that there are performance trade-offs between performance dimensions of complexity.

During the discussion of the findings, the results were considered in relation to the Cognition Hypothesis and to a lesser extent the Limited Attentional Capacity theory. Both of theses theories are often presented as competing positions in terms of complexity and language production, in part due to the differing approaches to attentional limits underpinning each theory. However, the findings from this thesis suggest that they do not necessarily have to be viewed in terms of a competing dialectic but may, in fact, both be viable processes operating under the limitations placed on language production by limited attentional resources, which can lead to trade-offs in performance.

Skehan and Foster (2012) stipulate that even though trade-offs in the performance of accuracy and complexity will not always be apparent because other factors may influence the outcome, the need for trade-offs in performance can be viewed as the default position. Though this thesis has not been able to make any claims regarding trade-offs between

production dimensions like accuracy and complexity, this thesis suggest that trade-offs in performance may also a defining feature within the single dimension of complexity.

This thesis proposes that, in terms of complex output, resource-directing and resource-dispersing variables (as described by the Cognition Hypothesis) are localised processes operating within the wider process suggested by the Limited Attentional Capacity theory. Furthermore, the success of the Cognition Hypothesis may be in part due to the level of attentional resource availability created by variable alignment, which is the appropriate alignment of all the variables need to promote positive complex outcomes.

For example, if there is a misalignment of complexity levels and procedural skill levels, the subsequent drain on limited attentional levels may lead to trade-offs in production. These trade-offs could manifest in two ways. The first is a trade-off between the higher or lower levels of complex output in one measure (for example mean length of dependent clauses). More or less dependent clauses may be used to meet the pragmatic requirements of a task, in part depending on cognitive task complexity and appropriate variable alignment. The second trade-off will manifest as trade-offs between types of complex output that are more or less cognitively demanding (for example between lexical complexity and syntactic complexity). The use of either lexical or syntactic means to meet the pragmatic requirements of a task may also depend to some extent on cognitive task complexity and appropriate variable alignment.

The conditions needed to facilitate the positive outputs predicted by the Cognition Hypothesis may be so detailed that it is like trying to hit a very small target, whereas the default position mentioned by Skehan and Foster is the very wide area that can easily be hit

when the target is missed. The default position suggested by Skehan and Foster does not necessarily dismiss the possibility of the Cognition Hypothesis as a viable process, only that the conditions under which it may happen are extremely difficult to arrange.

As a result, this thesis suggests that the notion of resource-directing and resource-dispersing elements of the Cognition Hypothesis may work when used in conjunction with attentional limits and the appropriate variable alignment. This means that increases in complexity along resource-directing and resource-dispersing dimensions can produce positive complex outputs when the appropriate affective variables are aligned; however, when aligning other affective variables, there are potentially very small margins of error due to limited attentional resources. When cognitive task complexity is manipulated along resource-directing and resource-dispersing dimensions, the resulting trade-offs between higher or lower levels of complex output, or trade-offs between types of complex output may be influenced by how appropriately aligned the contributing variables are.

6.3.2 Contributions to research

The discussion of the results in chapter five has raised some potentially important points about research in the field of cognitive task complexity and writing. The issues relate to modality, writing-to-learn, the application of patently lower complexity tasks, issues of proficiency, the importance of non-significant results, the trade-off process between individual complexity target measures, the significance of pre-task planning time in complexity studies, and the use of a non-standard measure of mean length of T-units.

During this thesis, the findings for cognitive task complexity on complex output, as they relate to effects of modality, were considered in relation to past findings in both the oral

and written modality. The results for modality alone did not present any clear evidence of the predicted benefits of the written modality (Grabe & Kaplan, 1996; Grabowski, 2007; Kormos & Trebits, 2012; Ellis & Yuan, 2004) on the complex output of individual writers.

When considering only the results for RQ1, which can be viewed as relating to mode alone without the added benefits of planning time, the findings from this thesis (which show positive increases in lexical complexity but none for syntax) are not too different from past studies. Oral modality studies have provided one result with positive effects for lexical complexity as diverse speech (Michel et al., 2007) but nothing for syntax. Written modality, using the same variables as this study, provide some support for lexical complexity as lexical sophistication (Kuiken & Vedder, 2007) and lexical variety (Sercu et al., 2006), but they provide nothing for syntactic complexity (Kuiken & Vedder, 2007, 2008, 2011, 2012; Sercu et al., 2006). Similarly, written modality studies using various complexity variables also provided some support for lexical complexity (Kormos, 2011; Ong & Zhang (2010) but no significant results were found for increases in complexity on syntactic complexity for second language learners.

Despite the lack of strong evidence for mode shown in the findings from RQ1, this thesis does make some suggestions regarding the potential benefits of mode when considered in conjunction with pre-task planning time. The results for RQ2, which had the additional facilitating effects of pre-task planning time with the written modality, produced positive effects for syntactic complexity. This is consistent with studies by Ishikawa (2006) and Ellis and Yuan (2004), who both used unguided pre-task planning time in conjunction with different types of cognitive task complexity to produce results with favourable outcomes for syntactic complexity.

The results from this thesis, in conjunction with Ishikawa (2006) and Ellis and Yuan (2004), could be viewed as support for the potential positive aspects of mode (as beneficial to promoting complex language under cognitive duress) when the written modality is considered a contributing facilitating dimension, and not a singular entity that succeeds or fails on its own.

The findings from this thesis, when viewed both between RQ1 and RQ2, and within RQ2 alone, imply that the addition of pre-task planning time to cognitive task complexity was the major influencing factor increasing the production of mean length of T-units; however, without a similar study in the oral modality running concurrently, using exactly the same elements, it is difficult to be sure if modality did not contribute. In fact, it is possible that the decreases in dependent clauses noted in RQ1 could have been much worse if not for the output being subject to the potential processing benefits associated with writing.

If mode alone is not sufficient to facilitate complex output, but contributes to a sufficient condition when combined with other variables like planning time, then it might be advisable to design future studies on cognitive task complexity and writing in a way that defines this issue with more clarity. In future research, it might be advisable to run concurrent oral modality studies, with exactly the same variables and performance conditions, in order to clarify this issue.

Further to the issue of mode, assertions have been made that the written modality promotes certain positive elements of linguistic processing (Manchon, 2011). Though this thesis is unable to make any definite claims about the proposed benefits of writing on linguistic processing, some potential benefits of writing in conjunction with forced complex output

297

are speculated upon below within the framework of writing-to-learn, which views writing as a process that promotes learning (Harklau, 2002, Manchon, 2011).

It is possible to theorise about the potential benefits to language learning that might have come from the act of writing in conjunction with the variables from RQ2, which produced positive results for increases in syntactic complexity, and RQ1, which produced positive results for lexical complexity. These findings show that the participants of this thesis have been able to write higher levels of complex language to meet the increasingly complex pragmatic requirements of the tasks, and potentially, this may have had some beneficial effect on their language abilities.

Two points should be considered when viewing the findings in terms of the potential effects of writing-to-learn. The first is Cummings (1990, p.483) statement about the potential effects of writing as modality.

> Composing might function broadly as a psycholinguistic output condition wherein learners analyse and consolidate second language knowledge that they have previously (but not fully) acquired…..Composition writing elicits an attention to form-meaning relations that may prompt learners to refine their linguistic expression-hence their control over their linguistic knowledge-so that it is more accurately representative of their thoughts and standard usage.

The second point regards the potential positive effects of pushed complex language output. Complex language output can be viewed as pushing learners to create more elaboration and structure in their developing language ability. Additionally, complex language output may

bring learner language closer to target language use, enable the efficient and precise expression of complex ideas, and might be a sign that acquisition is taking place (Cheng, 1985; Cook 1994; McLaughlin, 1990; Skehan, 1996; Swain, 1985).

Taken together, both points suggest that pushed complex language in the written modality could have combined to result in numerous benefits for the participants in this thesis. In terms of RQ 1, the combination of writing and cognitive task complexity has resulted in higher lexical diversity, which can be viewed as representing a more advanced native speaker form of expression. Though levels of attentional duress were considered to be smaller on the lexical complexity than on the syntactic complexity, the output still represents a more native like form of expression. The issue here is one of whether benefits to a student accrue when complex output occurs without high levels of cognitive duress.

For syntactic complexity, the combination of planning time, the potential language learning benefits of the writing, and increases in cognitive task complexity could have acted synergistically to direct learner attention to (and the production of) a more native speaker like control over dependent clauses (as seen in RQ2). This control may have resulted in a more native speaker like form of expression by pushing the participants to more efficiently and precisely formulate the complex ideas needed to meet the increasingly complex pragmatic requirements of the tasks.

The benefits noted above for syntactic complexity are based on the notion that increases in dependent clause usage is a more efficient and precise way to formulate ideas. Dependent clause usage has certainly been characterised as a more difficult and mature type of expression than the ability to use coordination. Increasing dependent clause usage can be

seen as increases in the means to express relationships, complicated propositions, and form coherent organization between statements that are related (Mendhalson, 1983). Any increases in the ability to do this would seem to be bringing participants closer to native like targets.

Carroll (1999) and Harklau (2002) have proposed that writing may have some beneficial effects on acquisition. Carroll (1999) states that all the different theories about acquisition suggest that meaningful exposure to language's grammatical properties allows learners to know the language they are exposed to. Additionally, Harklau (2002) has stated that there might be acquisition benefits in writing; for example, mnemonic strategies such as listing vocabulary and often-used phrases. Harklau also suggests that writing might aid acquisition by promoting the analysis of grammatical rules or sentences.

Exposure to the types of increasingly complex language, both syntactic (RQ2) and lexical (RQ1) in the meaningful context created by the necessity to meet increasingly complex pragmatic requirements may promote the "knowing" that Caroll refers to. Furthermore, the types of acquisition benefits derived from the analysis of grammatical rules or sentences connected to writing may be enhanced by the extra attentional focus required by formulating dependent clauses under cognitive complexity.

Finally, one has to wonder if being able to process larger amounts of complex material into complex written outputs affords opportunities to practice and successfully use language similar to the types required of more advanced communication like academic writing skills. The types of reasoning used in these tasks can be seen as basic practice for the types of reasoning used in academic writing where students will have to form opinions based on

supplied information. The results from RQ1 and RQ2 may have had some benefit in preparing students for the type of processes (both cognitive and in the practice of certain written forms) used in academic compositions.

It is the position of this thesis that future research in the field of cognitive task complexity and writing should also focus on the potential beneficial effects on linguistic processing (Manchon, 2011) and any potential effects on acquisition. Given the potential dual benefits of writing and pushed complex language output, this may be a rich area of research.

Another important issue for future research is the addition of patently lower complexity tasks during the data collecting process. Though this thesis had lingering questions over whether or not the patently low complexity task (Task 1) had different pragmatic requirements than Tasks 2 and 3, the variations in findings seemed to show that cognitive task complexity and planning time were the factors affecting changes at dependent clause level between the tasks. With that in mind, the patently low complexity tasks in future research need to remove any possibility of variation being affected by different pragmatic requirements. Other than this issue, the need to use a task that can demonstrably be shown to have lower levels of complexity seems like an essential element in tracking the effects of task complexity considering the current inability to accurately measure cognitively complex between complex tasks.

It is the position of this thesis that the application of the patently low complexity task has provided clearer tracking of variations in complex output than those of past studies (Kuiken & Vedder, 2007, 2008), which have relied on the current practice of imprecise

manipulation of cognitive task complexity, a practice necessitated by the current lack of accurate measuring tools for cognitive task complexity. Importantly, more work needs to be done on finding ways to more accurately and incrementally increase complexity between complex tasks.

As well as tracking complexity between tasks, the patently low complexity task (Task 1) was also used as an additional method of checking the participant's procedural skill levels on the target measures. An issue of concern arose during the analyses when Task 1 showed small variations in the procedural skill level of individual dependent clauses that seemed to be linked to a noticeable and sometimes statistically significant effect on dependent clauses when cognitive task complexity was increased between tasks.

If the slight differences in mean scores noted in the patently low complexity tasks (Task 1) actually represent slim variations in procedural skill levels, and if those slim differences in ability translate to significant effects on production when cognitive task complexity is applied, it is not hard to imagine how delicate the balance might be in pushing formulation to either positive or negative outputs when under cognitive duress. Further studies need to be conducted that investigate the effects that slim margins of procedural skill levels might have on the production of complex output during written cognitive task complexity studies.

The problems associated with the potential effects of small amounts of variation in procedural skill level highlights an important problem mentioned earlier; that is the current inability to accurately manipulate the degrees of measurement associated with complexity. One of the most the pressing issues with cognitive task complexity studies is the lack of detail that currently exists in executing complexity studies versus the potential level of

detail that might actually be required when increasing cognitive complexity on the limited attentional capacity formulation process (in this case on intermediate students).

The relatively undifferentiated means by which researchers manipulate levels of complexity between tasks belies the potentially delicate balance required to align all the elements involved in the process of formulating complex language under cognitive duress. As a result, more work must be done to create a controlled incremental application of complexity that reflects the potentially thin margins for error being worked with under cognitive stress.

Another issue that came to light during the analyses of the findings was the potential importance of including statistically non-significant results. When a broader view of the results was taken, it appeared that all the non-significant results and all the significant findings presented a more detailed picture of the effects of cognitive task complexity on mean length of T-units and mean segmental type-token ratio.

For example, in this thesis the difference between some of the significant and non-significant results were viewed as representing the different ways in which levels of procedural skill interacted with variation in levels of cognitive task complexity. If only the significant scores had been considered, then insights into the variable nature of these interactions would not have been as apparent. This thesis suggest that in situations where multiple target measures are being measured over the course of incremental increases in complexity across multiple tasks, the various results above and below significance paint a clearer picture of the processes at work than just interpreting the significant findings.

Another important issue is the potential trade-offs between types of complexity measures.

303

When examining the effects of cognitive task complexity on written output, this thesis followed the suggestion of Ellis and Barkhuizen (2005) who stated that it is important to use more than one type of complexity measure because language learners may use different means to express complexity in their writing. Having examined the findings from RQ1 and RQ2, this thesis has come to the conclusion that different types of complexity measures may be partially influenced by the variables noted in this thesis that come to bear on complex tasks.

Trade-offs in terms of production may not only occur between different types of production such as complexity and accuracy; trade-offs in production may also be noticeable between different measures of complexity. Different complexity measures like syntactic or lexical measures may be affected by different demands on attentional resources. Additionally, the use of either syntactic and lexical complexity measures, may also be influenced by other variables like the pragmatic task requirements, procedural skill levels, amounts of cognitive task complexity applied on limited intentional resources, and in the case of this thesis, the specific design of the instructions, which seemed to increases lexical complexity due to the increased items supplied in the instructions.

What this means for future research is that the variations in success of the various types of complexity measures are going to be subject to a multiplicity of influencing factors which need to be considered during analyses, one of which may be trade-offs between the measures themselves.

When analysing whether the independent variables have affected complex measures, it needs to be considered whether or not the results are affected by the differentiated cognitive

burden associated with each measure and how those measures may relate to each other. For example, it is important to consider whether a dialectic relationship forms from the competition for limited resources between different measures of complexity. In future research, it may be advisable to consider what type of cognitive burden is associated with different types of complex target measures when attempting to ascertain how all the myriad of potential factors have affected complex task output.

The findings from this study also appear to have strengthened the case for the addition of pre-task planning time as a central element in research focusing on the effects of cognitive task complexity on written syntactic development. As mentioned earlier, pre-task planning time was considered an important facilitating factor in the production of syntactic complexity in writing under cognitive duress. This thesis has contributed to the small amount of studies (Ishikawa, 2006; Ellis & Yuan, 2004) providing positive effects for the combinations of these variables. It should be noted that Ishikawa (2006) is not usually classified as a pre-task planning time study; however, pre-task planning time was part of the conditions of the research.

The final contribution to research is the use of the non-standard measure of mean length of T-unit. Previous research on cognitive task complexity and writing (Kuiken & Vedder, 2007, 2008, 2012), which used essentially the same dependent and independent variables as this thesis, mostly showed non-significant effects for cognitive task complexity on measures of subordination. It has been suggested that a contributing factor to this may have been the insensitive nature of the standard measure of T-units where all dependent clauses are measured as one group.

The findings from this thesis, using the non-standard measure where dependent clauses measured separately, have found significant variations that were masked by the standard T-unit measure. Additionally, analysing the dependent clauses separately (even when the results were non-significant) provided a richer view of the effects of cognitive task complexity on the production of dependent clauses. It is the suggestion of this thesis that previous work by Kuiken and Vedder (2007, 2008, 2012) may have benefited from using this approach, and that it is an approach that could be utilized in future research.

6.3.3 Pedagogy

This thesis' findings have suggested a number of issues that could be relevant to the teaching of complex language skills. For example, the development of complexity lessons should be reflected in both task design and assessment. Additionally, the nature of complexity development may also require a focus on the best ways to facilitate optimal complex formulation.

As mentioned previously, the use of a non-standard measure of mean length of T-units revealed findings that seemed to show that the production of individual dependent clauses responded differently depending on the combinations of a number of factors. Central factors were theorized to be the varied amounts of applied cognitive task complexity, the participant's procedural skill levels of individual dependent clauses, and the pragmatic requirements of the task. What this suggests for the development of pedagogic tasks in language learning classes is that the design of tasks may need to take all of these issues into account if optimum results are to be expected considering the delicate alignment of factors required to promote complex language under cognitive duress.

When tasks are being designed to focus specifically on the development of dependent clauses, the participant's ability to produce dependent clauses needs to be identified, as does the specific pragmatic dependent clause requirements of the tasks. Specific types of tasks need to be designed and tested for this purpose because it cannot be assumed that any task will automatically be disposed towards generating dependent clauses. Under optimal conditions, these tasks should elicit the requisite dependent clauses in the least complex task even before complexity is increased between tasks.

Once these tasks have been compiled, the participants need to be tested for the ability to produce the actual dependent clauses being targeted in the tasks. When the student's ability to produce dependent clauses has been married with the types of tasks that elicit them, then the chances of deriving beneficial effects from increasing resource-directing and resource-dispersing dimensions may be improved.

In addition to the alignment of the dependent clause procedural skill levels and the dependent clause related pragmatic requirements of a task is the central concern of applying cognitive task complexity between tasks in a measured and systematic way. This is a central and complicated issue, which has already been mentioned previously so does not need to be mentioned further.

The development of complexity, as an attention demanding process, requires considerations outside of just the marriage of procedural skill levels and pragmatic task requirements. Considerations also need to be made regarding the best way to assure students attentional levels are utilised to the best of their ability. For example, complexity focused tasks might need to be factored into syllabi at times when participant attention levels are optimised,

possibly in the morning or after breaks. It might be helpful to survey students regarding their own perceived optimal times. Furthermore, there might be a place in complexity focused lessons for maximizing the ability of participants to operate under complexity pressure by finding which tasks are best construed by participants, or by first creating an environment where a positive construal can be shaped. Considering that the margins for positive or negative effects of cognitive task complexity on complex output are unknown, and that the students will most likely be working at the upper end of their ability, these factors could have important impacts on the viability of complexity focused classes.

Finally, the evaluation of complex tasks needs to focus on complex elements. Teachers need to familiarize themselves with assessment techniques other than those that focus on accuracy, which is likely the element most focused on by teachers. The extensive preparations for complexity mentioned above need to be followed up with complexity focused assessment.

6.4 Limitations

As is the case with most studies, this research also has its limitations. Firstly, the exclusive use of a quantitative methodological approach utilised in this thesis has limitations. Quantitative data tends to lack the rich contextual nature of qualitative data; in essence, quantitative data can lack the ability to probe for details like the potential effects of students' individual differences, and the effects of environment on task outcomes.

Another potential problem for this study was the assumption that increasing the levels of complexity in the writing tasks would, under the best conditions, lead to positive language outcomes. Another way this research could have been made better was to have tested the

complexity tasks on native speakers. By doing this, the effects of these types of tasks on learners that already have fluent procedural knowledge of dependent clauses would have provided a clearer indication of the effects of these tasks under optimal conditions for limited attentional resources.

One of the central current debates around cognitive task complexity and the limits of attentional resources is the potential for both complexity and accuracy in student output to be simultaneously improved by the application of cognitive complexity. Unfortunately, this thesis was unable to address the output dimension of accuracy and as such is unable to contribute to the continuing dialogue surrounding these two dimensions.

Another potential limitation of this research was a lack of in depth research into the writing-to-learn aspects of the results. The investigation of the written mode in this thesis only included comparisons to past results in the written and oral modality. Research on the effects of mode might have been better served had the effects of cognitive task complexity and writing included the additional perspective of writing-to-learn, with the combination of independent variables and the written modality being analysed in relation to the effects on different aspects of linguistic processing like acquisition.

Another potential issue was the design of the patently low complexity task. During the data collecting process, a patently low complexity task was used that had lingering questioned about different pragmatic requirement compared to the other tasks. One concern was that this potential difference might have influenced the findings. In future research, this issue needs to be resolved in order to assure that, as much as possible, the variables being tested are responsible for the various effects being analysed.

During the data collection process, the participating schools supplied participants that were all of a similar skill level; however, during the analysis, the low complexity task, which was also used to provide information on participant skill levels, revealed that one group appeared significantly lower than the other two. Had time allowed, it would have been better if the participants had been tested before the data collecting procedure rather than relying only on the school's testing processes.

Further issues to be considered were the development of the complex tasks and the Likert scale questionnaire. In hindsight, testing the increasingly complex tasks on native speakers, prior to data collection, may have provided a better understanding of the effects of these tasks on the type of optimal processing expected from native speakers, thus providing a clearer picture of the utility of manipulating these complexity variables. When the Likert scale questionnaire was initially designed, there was, to the best of the researchers' knowledge, no available questionnaires focusing on task motivation and complex written output to model the items on. Furthermore, the items used in the questionnaire might have benefited from prior testing on the participants to ensure that the meaning implied by the researcher was the same meaning understood by the participants. Unfortunately, limited access to the participants did not allow for this at the time.

Finally, issues arose regarding the participant's understanding of key words in the Likert scale questionnaire that might have affected the findings. There may have been potential problems related to the effectiveness of the Likert scale questionnaires if there was a disparity between the researcher's and the participant's understanding of key words. For example, in the Likert scale questionnaires, the word "difficulty" was used in the Likert items to express a state similar to what was understood to be pressure exerted by increases

310

in cognitive task complexity; however, there was no guarantee that this was how it was understood by the participants.

The potential problems regarding discrepancies between what students perceived the Likert scale questionnaires to say and the intended meaning of the researcher need to be considered if the data collected from the Likert scales is to be considered relevant. As a result, key words need to be clarified before or during the data collecting process to ensure the alignment of participant attitudes and the key foci of the questionnaire.

6.5 Recommendations for future research

As a result of the work in this thesis, a number of suggestions for future research have emerged. These suggestions are listed, in no particular order, below.

1. Written studies using cognitive task complexity need to isolate the contribution of modality by running equivalent studies in which modality is a manipulated variable.
2. More research needs to be performed on writing and cognitive task complexity, with a focus on more exhaustive research on one group of variables (like reasoning demands and number of elements) in order for stronger generalisation to be made from the results.
3. Cognitive task complexity studies in the written modality need to focus on writing-to-learn elements, to better understand the effects of cognitive task complexity on elements like acquisition.
4. Future studies utilizing the mean length of T-units to measure dependent clause production need to consider using the non-standard measure where independent clauses are analysed separately.

5. More studies need to be done on the effects of cognitive task complexity, writing, and pre-task planning time. There is a paucity of research using these variables, which have produced positive outcomes for syntactic production. These variables might also be considered in relation to the writing-to-learn model and potential effects on linguistic processing.

6. Synergistic studies combining cognitive task complexity, writing, and the development of academic compositions need to be considered as an area of research.

7. Further studies need to be conducted that investigate the effects that slim margins of procedural skill levels might have on the production of complex output during written cognitive task complexity studies.

8. Studies need to be conducted on the effectiveness of the types of complexity focused pedagogical suggestions mentioned in section 6.3.3, where complexity focused task design, complexity focused assessment, and the facilitating of optimal processing conditions are combined and utilised in language learning syllabi.

9. Studies need to be conducted on the potential effects of trade-offs between competing measures of complex production on complex output.

10. Future studies utilising the same variables as this research could incorporate the dimension of accuracy in order to gain a better understanding of the effects of cognitive task complexity on potential trade-offs between different output dimensions.

11. Any future research that investigates similar variables and data collecting instruments may need to investigate the utility of the complexity tasks by testing the resource-directing variables on native speakers first as well as testing the

questionnaires on the participants to ensure that there is agreement on the meaning of the items in the Likert scale.

6.6 Final remarks

The study of cognitive task complexity and writing is a relatively under researched area, and the results from this thesis have shown that there is a rich vein of research to be mined, with potential contributions to acquisition, pedagogy, and a potential rethinking of the dialectic relationship between competing theories on complexity and output.

The positive aspects form RQ1 and RQ2 have shown that under the correct conditions, students can benefit from increases in cognitive task complexity. However, the results from this thesis have shown that the focus on complexity in writing development is potentially very detailed and difficult, with much consideration needing to be placed into the alignment of contributing variables if students are going to produce complex language under cognitive duress using limited attentional resources. For teachers and researchers to be able to control positive outcomes, much more work will be required to be able to remove some of the random nature of the current work.

REFERENCES

Adams, R., Armani, S., Newton, J., & Nik, A. (in press). Planning and production in CMC writing. In H. Byrnes & R. M. Manchon (Eds.), *Task-based language learning: Insights from and for L2 writing*. Amsterdam: John Benjamins.

Brown, G., Anderson, A., Shilcock, R., & Yule, G. (1984). *Teaching talk: Strategies for production and assessment*. Cambridge: Cambridge University Press.

Brown, J, D. (2003). The coefficient of determination. *JALT Testing & Evaluation SIG Newsletter*, *7* (1), 14-16.

Brown, J. D., & Rodger, T. S. (2002). *Doing second language research*. Oxford University Press.

Bulte, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-46). Amsterdam: John Benjamins.

Bygate, M. (1999). Task as context for the framing, reframing and unframing of language. *System, 27*, 33-48.

Candlin, C. (1987). Towards task-based language learning. In C. Candlin & D. Murphy (Eds.), *Language learning tasks* (pp. 5-22). London: Prentice Hall.

Cheng, P .W. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review, 92,* 414-23.

Cheung, H., & Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. *Applied Psycholinguistics, 13*, 53-76.

Cohen, J. W. (1988). *Statisitical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cook, V. J. (1994). *Linguistics and second language acquisition.* London: Macmillan.

Corder, S, P. (1967). The significance of learners' errors. *IRAL, 5*, 161-169.

Coughlan, P., & Duff . P. (1994). Same task, different activities: Analysis of SLA task from an Activity theory perspective. In J. P. Lantolf & G. Appel (Eds*.), Vygotskian approaches to second language research* (pp. 173-193). Norwood, N.J: Ablex.

Cromer, R. (1991). *Language and thought in normal and handicapped children*. Oxford: Blackwell.

Crookes, G., & Schmidt, R . (1991). Motivation: Reopening the research agenda. *Language Learning, 41*, 469-512.

Cooper, T, C. (1976). Measuring written syntactic patterns of second language learners of german. *Journal of Educational Research, 69* (5), 176-183.

Crowhurst, M. (1980). Syntactic complexity and teachers' quality ratings of narrations and arguments. *Researching the Teaching of English, 14*, 223-231.

Crowhurst, M. (1983). Syntactic complexity and writing quality. *Canadian Journal of Education, 8* (1), 1-16.

Cumming, A .(1990). Metalinguisitc and ideational thinking in second language composing. *Written Communication, 7* (4), 482-511.

Dornyei, Z. (1994). Understanding motivation: On with the challenge! *The Modern Language Journal, 78* (4), 515-523.

Dornyei, Z. (1999). Motivation. In B. Spolsky (Ed.), *Concise encyclopaedia of educational linguistics* (pp. 525-532). Oxford: Elsevier.

Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences in second language acquisition* (pp. 137-158). Amsterdam: John Benjamins.

Dornyei, Z. (2003a). *Questionnaires in second language research*. Lawrence Erlbaum Associates, Inc.

Dörnyei, Z. (2003b). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53* (1), 3-32.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dornyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning, 59,* 230-248.

Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247-267). London: Continuum.

Dörnyei, Z., & Csizér, K. (1998). Ten commandments for motivating language learners: Results of an empirical study. *Language Teaching Research, 2,* 203-229.

Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research, 4* (3), 275–300.

Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics, 4*, 43-69.

Dörnyei, Z., & Skehan. P. (2005). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589-630). Oxford: Blackwell.

Elder, C., & Iwashita, N. (2005). Planning for test performance: Does it make a difference? In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 219-238). Amsterdam: John Benjamins.

Ellis, N, C., & Larsen- Freeman, D. (2006). Language emergence: Implications for applied linguistics—introduction to the special issue. *Applied Linguistics, 27* (4), 558–589.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4* (3), 193-220.

Ellis, R. (2003). *Task- based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2005). *Planning and task performance in a second language.* Amsterdam: John Benjamins.

Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Ellis , R., & Yuan, F .(2004). The effects of planning time on fluency, complexity, and accuracy in second language narrative writing. *Studied in Second Language Acquisition, 26* (1), 59-84.

Ferris, D. R., Liu, H., Sinha, A., & Senna , M. (2012) Written corrective feedback for individual L2 writers. *Journal of Second Language Writing, 22*, 307–329.

Foster, P. (1996). Doing the task better: How planning time influences students' performance. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp.126-135). London: Heinemann.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18* (3), 299-324.

Foster, P., & Skehan, P. (1999). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition, 18* (3), 299-324.

Gaies, S, J. (1980). T-Unit analysis in second language research: Applications, problems and limitations. *TESOL Quarterly, 14* (1), 53-60.

Gilabert, R. (2007a). Effects of manipulating task complexity on self-repairs during L2 oral production *IRAL, 45* (3), 215-240.

Gilabert, R. (2007b). The simultaneous manipulation of task complexity along planning time and here and now: Effects on L2 oral production. In M, García Mayo (Ed.), *Investigating tasks in formal language learning* (pp.44-68). Clevedon, Avon: Multilingual Matters.

Gilabert, R., Baron, J., & Llanes, A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *IRAL, 47*, 367-395.

Givon, T. (1985). Function, structure, and language acquisition. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition, Vol. 1* (pp. 1008–1025). Hillsdale, NJ: Lawrence Erlbaum.

Givon, T. (2009). *The genesis of syntactic complexity.* Amsterdam: Benjamins.

Grabe, W., & Kaplan, W. (1996). *Theory and practice of writing: An applied linguistic perspective*. London: Longman.

Grabowski, J. (2007). The writing superiority effect in the verbal recall of knowledge: Source and determinants. In M. Torrace, L. van Waes & D. Galbraith (Eds.), *Writing and cognition research and applications* (pp. 165-179). Amsterdam: Elsevier.

Granfeldt, J. (2008). Speaking and writing in L2 French: Exploring effects on fluency, accuracy and complexity. In A. Housen, S. van Daele, & M. Pierrard (Eds.), *Proceedings of the conference on complexity, accuracy and fluency in second language use, learning and teaching* (pp. 87–98). Brussels, Belgium: Koninklijke Vlaamse Academi.

Guiraud, P. (1954). *Les caracteres statistiques du vocabulaire.* Paris: Presses Universitaires de France.

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing, 11*, 329-350.

Hake, R., & Williams, J. M. (1979). Sentence expanding: Not can, or how, but when. In D. Daiker, A. Kerek & M. Morenberg (Eds.), *Sentence combining and the teaching of writing* (pp.134-146). Akron, Ohio: L & S Books.

Heckhausen, H. (1991). *Motivation and action.* Berlin: Springer

Heckhausen, H., & Kuhl, J. (1985). From wishes to action: The dead ends and short cuts on the long way to action. In M. Frese & J. Sabini (Eds.), *Goal-directed behaviour: Psychological theory and research on action* (pp.134-160). Hillsdale, NJ: Lawrence Erlbaum.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30* (4), 461-473.

Hudson, R. (2009). Measuring maturity. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), *The Sage handbook of writing development* (pp. 349-362). London: Sage Publications.

Hunt, K, W. (1965). *Grammatical structures written at three grade levels*. Urbana, IL: The National Council of Teachers of English.

Ishikawa, S. (1995). Objective measurement of low-proficiency EFL narrative writing. *Journal of Second Language Writing*, *4*, 51-70.

Ishikawa, T. (2006). The effect of manipulating task complexity along the [here-and-now] dimension on L2 written narrative discourse. In M. del Pilar García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136-156). Clevedon: Multilingual Matters.

Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design. *Language Learning, 21* (3), 401-436.

Johnson, M, D., Mercado L., & Acevedo, A. (2012) The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing, 21*, 264–282.

Julkunen, K. (1989). *Situation- and task-specific motivation in foreign language learning and teaching*. Finland: University of Joensuu.

Julkunen, K. (2001). Situation- and task-specific motivation in foreign language learning. In Z. Dornyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 29-42). Honolulu, HI: University of Hawaii 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). Amsterdam: John Benjamins.

Keller, J, M. (1983). Motivational design of instruction. In C. M. Reigelruth (Ed.), *Instructional design theory and models: An overview of their current status* (pp. 383-424). Hillsdale, NJ: Lawrence Erlbaum.

Kellog, R, T. (1996). A model of working memory in writing. In M. C. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-71). Hillsdale, NJ, England: Lawrence Erlbaum.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20,* 148–161.

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62 (2), 439-472.

Kuhl, J., & Beckmann, J. (1994). *Volition and personality: Action versus state orientation.* Seattle, WA: Hogrefe & Huber

Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster- Cohen & M. P. Garcıa-Mayo (Eds.), *Eurosla yearbook. Vol*. 5 (pp.195–222). Amsterdam: John Benjamins.

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *IRAL, 45*, 261-284.

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*, 48-60.

Kuiken, F., & Vedder, I. (2011). Task performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 91–104). Amsterdam: John Benjamins.

Kuiken , F.,& Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in l2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 199-120). Philadelphia/Amsterdam: John Benjamins.

Lantolf, J., & Appel, G. (Eds) (1994). *Vygotskian approaches to second language research*. Norwood NJ: Ablex.

Larsen- Freeman, D. (2002). Language acquisition and language use from a chaos/complexity theory perspective. In C. Kramsch (Ed.), *Language acquisition and language socialization* (pp. 33-46). London: Continuum.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27* (4), 590-619.

Leontiev, A, A . (1981). *Psychology and the language learning process*. London: Pergamon.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 177–197). Philadelphia/Amsterdam: John Benjamins.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modeling and assessing second language acquisition* (pp. 77-99). Clevedon, Avon: Multilingual Matters.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspectives* (pp.39-52). Amsterdam: John Benjamins.

Long, M. H., & Crookes, G. (1987). Intervention points in second language classroom processes. In B. Das (Ed.), *Patterns of classroom interaction*. Singapore: SEAMEO RELC.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26*, 27-56.

Lord, C. (2002). Are subordinate clauses more difficult? In J. L . Bybee & M. Noonan (Eds.), *Complex sentences in grammar and discourse: Essays in honour of Sandra A Thompson* (pp. 223-233). Philadelphia, PA: John Benjamins.

Mackey, A., & Gass, S. M. (2005). *Second language research.* Mahwah: NJ: Lawrence Erlbaum.

Manchon, R. M. (2011) . Writing to learn the language: Issues in theory and research. In R. M. Manchon (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp.61-82). Amsterdam: John Benjamins.

Manolopoulou-Sergi, E. (2004). Motivation within the information-processing model of foreign language learning. *System, 32*, 427-441

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11,* 113-28.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition 20*, 52–83.

Mendhalson, D. J. (1983). The case for considering syntactic maturity in E.S.L. and E.F.L. *IRAL, 21* (4), 299-311.

Michel. M. C., Kuiken, F., & I, Vedder .(2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *IRAL, 45,* 241–259.

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*, 11–37.

Monroe, J. H. (1975). Measuring and enhancing syntactic fluency in French. *The French Review*, *48*, 1023-1031.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30* (4), 555-578.

Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System*, *34*, 566–585.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21,* 109–48.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24,* 492–518.

Ong, J., & Zhang, J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL student's argumentative writing. *Journal of Second Language Writing*, 19, 218-233.

Pallant, J. (2002). *SPSS survival manual*. New South Wales, Australia: Allen & Unwin.

Pallotti, G. (2009). CAF: defining, refining and differentiating constructs. *Applied Linguistics, 30* (4), 590-601.

Prabhu, N. S. (1987). *Second Language pedagogy*. Oxford, Oxford University Press.

Rahimpour, M., & Safarie, M. (2011). The effects of on-line planning and pre-task planning on descriptive writing of Iranian learners. *International Journal of English Linguistics, 2* (1), 274-280.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45* (1), 99-140.

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge: Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22,* 27-57.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL, 43* (1), 1–32.

Robinson, P. (2007a). Criteria for grading and sequencing pedagogic tasks. In M. del Pilar Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp.7–27). Clevedon, Avon: Multilingual Matters.

Robinson, P. (2007b). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL, 45*, 193-213.

Robinson, P., & Gilabert, R. (2007). Task complexity, and the cognition hypothesis and second language learning and performance. *IRAL, 45*, 161-176.

Rutherford, K. (2001*). An investigation of the effects of planning on oral production in a second language*. Unpublished MA Thesis: University of Auckland, New Zealand.

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.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), Cognition *and second language Instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Selinker, L. (1972). Interlanguage. *IRAL, 10*, 209-231.

Sercu, L., De Wachter, L., Peters, E., Kuiken, F., & Vedder, I. (2006). The effect of task complexity and task conditions on foreign language development and performance: Three empirical studies. *ITL—International Journal of Applied Linguistics, 151*, 55–84.

Sharma, A. (1980). Syntactic maturity: Assessing writing proficiency in a second language. In R. Silverstein (Ed.), *Occasional papers in linguistics, No 6* (pp. 318-325). Carbondale: Southern Illinois University.

Skehan, P. (1996). A framework for the implementation of task based instruction. *Applied Linguistics, 17* (6), 190-204.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*, 1-14.

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30* (4), 510-532.

Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research, 1* (3), 185–211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49* (1), 93-120.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge: Cambridge University.

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 193-218). Philadelphia/Amsterdam: John Benjamins.

Skehan, P., & Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance: A synthesis of the Ealing research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 199-120). Philadelphia/Amsterdam: John Benjamins.

Slobin, D. (1993). Adult language acquisition: A view from child language study. In C. Perdue (Ed.), *Adult language acquisition: Cross-linguistic perspectives: Vol. 2. The results* (pp. 239–252). Cambridge: Cambridge University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.) *Input in second language acquisition* (pp. 235-256). New York: Newbury House.

Tavakoli, P., & Skehan, P. (2005). Planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-277). Amsterdam: John Benjamins.

Tremblay, P. F., Goldberg, M. P., & Gardner, R. C. (1995). Trait and state motivation and the acquisition of Hebrew vocabulary. *Canadian Journal of Behavioural Science, 27*, 356-370.

VanPatten, B. (1989). Can learners attend to form and content while processing input? *Hispania, 72* (2), 409-417.

VanPatten, B. (1990). Attending to form and content in the input, an experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287-301.

VanPatten, B. (1996). *Input processing and grammatical instruction in second language acquisition*. Norwood NJ: Ablex.

Weaver, C. (1996). *Teaching grammar in context*. Portsmouth, NH: Heinemann.

Wendel, J. (1997). *Planning and second language narrative production*. Unpublished Doctoral Dissertation, Temple University, Japan.

Wickens, C. D. (1980). The structure of attentional resources, In R. S .Nickerson (Ed.), *Attention and performance VIII* (pp. 239-257). Hillsdale NJ: Lawrence Erlbaum.

Wickens, C. D. (1989). Attention and skilled performance. In D. Holding (Ed.), *Human skill* (pp.71-105). New York: John Wiley.

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing, 14* (1), 21–44.

Wilkins, D. (1976). *Notional syllabuses*. Oxford: Oxford University Press.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu, HI: University of Hawaii Press

Yuan, F., & Ellis, R. (2003). The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24* (1), 1–27.

_____

Appendix A: Participant survey

Name   _____   email   _____

Gender   Male   _____   Female   _____

Age   _____

First language   _____

How many years have you been studying English?   _____

How old were you when you started studying English?   _____

How long have you been studying English in New Zealand?   _____

How long have you been studying English writing?   _____

Do you want to study at University in English?   _____

How long have you been studying academic writing?   _____

How often do you do writing homework?   _____

Do you write English outside of class-time (not homework)?   _____

What type of English writing do you do outside of class
work e.g. writing emails, writing letters etc?   _____

How often do you write in English outside of class
time and homework?   _____

Are you currently studying at more than one language school?   _____

If the answer to the last question was yes, how many schools,
including this school, are you studying at?   _____

How often do you read in English (not including school work)? _____

What type of things do you read in English outside of school
work e.g. English novels, English magazines etc?   _____

What level do you think your English writing is?

Very poor   _____   Good   _____

Poor   _____   Very good   _____

Average   _____   Really good _____

Above average   _____   Excellent   _____

Appendix B: Task 1 (low complexity)

LETTER ONE.   **Name.** ———————————————

This activity is about writing a letter to a friend. Read the information below then write a letter based on the situation and following the instructions.

Situation:
1. You have a close English-speaking friend called Peter.
2. Peter is thinking about coming to New Zealand.

Instructions:
1. You have to write a letter to your friend of about 200 to 250 words.
2. In this letter, you should tell Peter about New Zealand.
3. Start the letter below.

Dear Peter,

LETTER TWO. NAME ————————————

This activity is about writing a letter to a friend. Read the information below then write a letter based on the situation and following the instructions.

Situation:
1. Your friend John is coming to New Zealand for one weekend, and there are two restaurants he really wants to try.
2. There is only time to go to one restaurant. As a result, John wants you to choose one restaurant.
3. Neither of the restaurants you have checked are perfect for John and your requirements.

Instructions:
1. Look at John's requirements in list A.
2. Look at the restaurant information in list B.
3. Consider your own personal preferences.
4. Using the information from lists A and B and your own preferences, write John a letter of between 200 and 250 words telling him which restaurant you have chosen and why you choose it.
5. Start the letter below.

Dear John,

Appendix D: Task 2 (supplementary information)

List A. Information regarding you and John.

John's information:
1. He is arriving on Saturday morning and leaving on the following Monday afternoon.
2. Seafood and Pork are his favourite food.
3. He generally eats a lot.
4. He doesn't particularly love sweet food, but enjoys dessert some times.
5. He likes to drink a glass of wine with diner.
6. He only speaks English.
7. He will be staying with you during his time here, so transportation will be your responsibility.

Your information:
1.When you are considering the restaurant, consider your actual personal preferences.

List B: Restaurant information.

Restaurant 1:
Opening times: 4:00 pm to 8:00 pm Monday to Saturday.
Prices: Main courses (main meal) cost around $20.
Availability: Usually the restaurant is very busy and bookings (reserve a table) are necessary to get a table.
Critic's review of food quality:
The seafood selection (what is available) is good, and is considered very high quality.
The beef is average quality.
The pork is average quality.
There are no desserts (ice cream etc) at this restaurant.
The portions (size of meal) are average size.
Drink: The beer and wine is expensive. There is no BYO (bring your own drinks).
Staff: Some staff speak English, some staff only speak Japanese.
Service: The service is quick, but the staff do not appear friendly.
Entertainment: Karaoke after 7:00pm.
Location: In Auckland's central city.
Parking: Restaurant supplies no parking.

Restaurant 2:
Opening times: 4:00pm to 9:00 pm Tuesday to Sunday.
Prices: Main courses cost around $28.
Availability: Quiet during the week, sometimes busy on the weekend. No booking is necessary.
Critic's review of food quality:
The seafood selection is good and the quality is good.
The pork is good quality.
The beef is high quality.
The desserts are very high quality.
The chicken is average quality.
Portions are larger than average size.
Drink: Beer and wine are very expensive, however, you can BYO.
Staff: Some staff members speak English, and some staff members speak Chinese.
Service: The service is efficient, and the members of staff are helpful.
Entertainment: There is no entertainment supplied by the restaurant.
Location: In Auckland's central city.
Parking: Restaurant supplies a small amount of parking for customers, though much less than the restaurant requires.

Appendix E: Task 3 (high complexity)

LETTER THREE.   **NAME** ——————————————

This activity is about writing a letter to a friend. Read the information below then write a letter based on the situation and following the instructions.

Situation:
1.You have a friend called Kate who is coming to New Zealand for one weekend and there are three restaurants she really wants to try.
2. There is only time to go to one restaurant. As a result, Kate is asking you to choose one restaurant.
3.You are also bringing your friends Jerry and Marvin to the restaurant.
4. None of the restaurants you have checked are completely perfect for Kate, Jerry, Marvin and your requirements.

Instructions:
1. Look at Kate, Jerry, and Marvin's requirements in list A.
2. Look at the restaurant information in list B.
3. Consider your own personal preferences.
4. Using the information from lists A and B, write Kate a letter of between 200 and 250 words telling her which restaurant you have chosen and why.
5. Start the letter below.

Dear Kate,

Appendix F: Task 3 (supplementary information)

List A. Information regarding Kate, Jerry, Marvin, and you.

Kate's information:
1. She is arriving on Saturday morning and leaving on the following Monday afternoon.
2. Beef is her favourite food.
3. She generally eats small meals.
4. She enjoys desserts.
5. She drinks a little alcohol, but it is not essential for her.
6. She speaks English and Japanese.
7. She likes to have fun, e.g. listen to music or sing Karaoke.
8. She will be staying with you during her time here, so transportation will be your responsibility.

Jerry's information:
1. He lives in the city centre in an apartment.
2. He eats anything, but really like seafood.
3. He eats a lot.
4. He doesn't eat dessert.
5. He likes to drink beer.
6. He speaks Japanese and Korean, and his English is ok.
7. He has a car.
8. He enjoys Karaoke.

Marvin's information:
1. He lives on the North Shore.
2. He loves Chicken and beef, but doesn't eat seafood.
3. He eats a lot.
4. He eats dessert, but doesn't really care about it.
5. He drinks, but it is not important to him if the restaurant has beer or wine.
6. His English is average, his Japanese is average, his Chinese is good.
7. He has a car.
8. He doesn't care about entertainment at the restaurant.

Your information:
1. When you are considering the restaurant, consider your own personal preference

List B: Restaurant information.

Restaurant 1:
Opening times: 4:00 pm to 9:00 pm Monday to Saturday.
Prices: Main courses (main meal) cost around $24.
Availability: Usually the restaurant is very busy and bookings (reserve a table) are necessary to get a table.
Critic's review of food quality:
The seafood is average, but it is considered to be very high quality.
The beef is good quality.
The restaurant does not serve pork.
The chicken is very high quality.
There are no desserts (ice cream etc.) at this restaurant.
The portions (size of meal) are very large.
Drink: The beer is cheap and the wine high quality and cheap. There is no BYO (bring your own drinks).
Staff: Some staff members speak English, some staff members only speak Japanese, and some staff members only speak Italian.
Service: The service is quick, but the members of staff do not appear friendly.
Entertainment: There is no entertainment.
Location: In Auckland's central city.
Parking: Restaurant supplies no parking.

Restaurant 2:
Opening times: 4:00pm to 12:00 pm Tuesday to Sunday.
Prices: Main courses cost around $24.
Availability: This restaurant is usually quite. No booking is necessary.
Critic's review of food quality:
There is no seafood served at this restaurant.
The pork is good quality.
The beef is high quality.
The desserts are very high quality.
The chicken is very poor quality.
Portions are larger than average size.
Drink: Beer and wine are very expensive, however, you can BYO.
Staff: Some staff members speak English, and some staff members speak Chinese.
Service: The service is efficient, and the members of staff are helpful.
Entertainment: There is no entertainment supplied by the restaurant, however, the restaurant is next door to a very popular bar with good entertainment.
Location: In Auckland's central city.
Parking: Restaurant supplies a small amount of parking for customers, though much less than the restaurant requires.

Restaurant 3:
Opening times: 4:00pm to 10:00pm Tuesday to Sunday.
Prices: Main courses cost around $27.
Availability: Generally not very busy. No booking is required.
Critic's review of food quality:
The seafood selection is poor, and the quality is average.
The beef is very good.
The pork is average.
The chicken is very high quality.
The desserts are very high quality.
The portions are below average.
Drink: There is a bar at this restaurant that is always open, but no BYO.
Staff: The members of staff are very friendly, and speak English, Chinese, and Japanese.
Service: The service is very quick, and sometimes the staff gets the orders wrong.
Entertainment: Karaoke after at 8:00 pm on the weekends.
Location: In Pakuranga, which is about a 30-minute drive from Auckland city centre.
Parking: Free parking is available to restaurant customers only after 7:00 pm.

Appendix G: Likert Scale Questionnaire.

Read each statement then place an **X** in the square below that best describes your level of agreement or disagreement with the statement.
Do not take too much time to think about your answers. Use the first answer that comes to mind when marking the boxes.

1. These tasks were easy.

☐   ☐   ☐   ☐   ☐   ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

2. These tasks are the same as other tasks I have done in writing classes.

☐   ☐   ☐   ☐   ☐   ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

3. The English I wrote for these tasks is different to the English I use outside of class.

☐   ☐   ☐   ☐   ☐   ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

4. I understand how these tasks are supposed to help me learn English.

☐   ☐   ☐   ☐   ☐   ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

5. When the tasks became difficult, I really focused on individual words, not sentences.

☐   ☐   ☐   ☐   ☐   ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

6. When the tasks became difficult, I lost interest in completing them.

☐ ☐ ☐ ☐ ☐ ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

7. When the tasks became difficult, I concentrated harder to complete them.

☐ ☐ ☐ ☐ ☐ ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

8. These tasks had too many things to concentrate on.

☐ ☐ ☐ ☐ ☐ ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

9. I re-read my work repeatedly while I was writing it.

☐ ☐ ☐ ☐ ☐ ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

10. The instructions for these tasks were very clear.

☐ ☐ ☐ ☐ ☐ ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

11. I felt tired before starting these tasks.

☐ ☐ ☐ ☐ ☐ ☐
Strongly   Disagree   Slightly   Partly   Agree   Strongly
disagree   disagree   agree   agree

	Participant Information Sheet	AUT U N I V E R S I T Y TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project Title: Cognitive Complexity and Written Output.

1. An Invitation: My name is Mark Frear and I am inviting you to participate in a study on writing. This study is *voluntary*, if you join, you can drop out at any time.

2. What is the purpose of this research? The data being collected will be used as part of a PhD on writing in a second language.

3. How was I chosen for this invitation? You have been chosen because you are an adult, intermediate level writer of English, who is studying at AUT International House.

4. What will happen in this research? There are two stages: **A**. We meet on Friday the 28th at 12.30 pm in room WW 904. You will fill out a form and do a very short piece of writing. This takes 40 minutes. **B**. We meet once more, at a time that is convenient for you, between Monday the 7th of June and Friday the 11th of June. You will fill out two questionnaires and do two very short pieces of writing. This should take about 80 minutes.

5. How will any discomforts and risks be alleviated? 1. The researcher is not employed at AUT International House. 2. Your teachers are not involved. 3.You are free to leave anytime. 4.You can discuss your concerns with the researcher or the supervisor. 6. All information is confidential.

6. What are the benefits? 1. Participating will help with the fluency of your writing. 2. A koha in the form of a free movie ticket will be given to you on *completion* of the data collection. 3. You will have been part of a study that is intended to help develop writing skills. 4. If you wish, you may receive a report on the findings of this study.

7. How will my privacy be protected? All information is confidential and will be kept in a locked filing cabinet in Professor Bitcheners' office. After six years, all information will be destroyed. In addition, no names will be mentioned in the final study.

8. What are the costs of participating in this research? The cost to you will be in time, see 4.

9. What opportunity do I have to consider this invitation? You can withdraw at anytime. If you withdraw, please contact the researcher to avoid problems with the data collection process.

10. How do I agree to participate in this research? If you choose to participate, you can fill in the consent form and bring it with you on Friday 28th.

11. Will I receive feedback on the results of this research? If you wish to have feedback or discuss your results, please tick the box on the consent form.

12. What do I do if I have concerns about this research? Contact project supervisor: Professor John Bitchener, email address, *john.bitchener@aut.ac.nz* Tel. 09 921 9999 ext 7830

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Madeline Banda, *madeline.banda@aut.ac.nz*, 921 9999 ext 8044.

Whom do I contact for further information about this research?
Researcher Contact Details: Mark Frear email address, *mwfrear@hotmail.com*

Project Supervisor Contact Details: Professor John Bitchener, email address,
john.bitchener@aut.ac.nz Tel. 09 921 9999 ext 7830

Appendix I: Participant Information Sheet

	Consent Form	AUT U N I V E R S I T Y TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title: Cognitive Task Complexity and Written Output.

Project Supervisor: Professor John Bitchener.

Researcher: Mark Frear.

O   I have read and understood the information provided about this research project.

O   I have had an opportunity to ask questions and to have them answered.

O   I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

O   I am not suffering from any health problems that will be affected by participating in this study.

O   I agree to provide a survey, three written tasks, and two questionnaires.

O   I agree to take part in this research.

O   I wish to receive a copy of the report from the research (please tick one):

        Yes O

        No   O

Participant's signature:
        ....……………………………………………………………
……
Participant's name:
        ....……………………………………………………………
……
Participant's Contact Details (if appropriate):
…………………………………………………………………………………………..
…………………………………………………………………………………………..
…………………………………………………………………………………………..
…………………………………………………………………………………………..
Date:
Approved by the Auckland University of Technology Ethics Committee on 19/8/2009
AUTEC Reference number 09/142
Note: The Participant should retain a copy of this form.

M E M O R A N D U M
Auckland University of Technology Ethics Committee (AUTEC)

To:	John Bitchener
From:	**Madeline Banda** Executive Secretary, AUTEC
Date:	19 August 2009
Subject:	Ethics Application Number 09/142 **Cognitive task complexity and written output.**

Dear John

Thank you for providing written evidence as requested. I am pleased to advise that it satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC) at their meeting on 13 July 2009 and that I have approved your ethics application. This delegated approval is made in accordance with section 5.3.2.3 of AUTEC's *Applying for Ethics Approval: Guidelines and Procedures* and is subject to endorsement at AUTEC's meeting on 14 September 2009.

Your ethics application is approved for a period of three years until 19 August 2012.

I advise that as part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through *http://www.aut.ac.nz/research/research-ethics*. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 19 August 2012;

- A brief report on the status of the project using form EA3, which is available online through *http://www.aut.ac.nz/research/research-ethics*. This report is to be submitted either when the approval expires on 19 August 2012 or on completion of the project, whichever comes sooner;

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are reminded that, as applicant, you are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

Please note that AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to make the arrangements necessary to obtain this. Also, if your research is undertaken within a

jurisdiction outside New Zealand, you will need to make the arrangements necessary to meet the legal and ethical requirements that apply within that jurisdiction.

When communicating with us about this application, we ask that you use the application number and study title to enable us to provide you with prompt service. Should you have any further enquiries regarding this matter, you are welcome to contact Charles Grinter, Ethics Coordinator, by email at ethics@aut.ac.nz or by telephone on 921 9999 at extension 8860.

On behalf of the AUTEC and myself, I wish you success with your research and look forward to reading about it in your reports.

Yours sincerely

Madeline Banda
Executive Secretary
Auckland University of Technology Ethics Committee

Cc: Mark Wain Frear mwfrear@hotmail.com