The Comparative Effects of Traditional, Vocabulary, and Grammar Instruction on the Reading Comprehension and Fluency of High School English Language Learners

by

Kent Ka-Kin Lee

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Studies in Teaching and Learning English as a Second Language

Department of Educational Psychology

University of Alberta

Abstract

To meet the academic demands of high school content courses such as science, English language learners (ELL) require more exposure to academic language than they currently receive (Wong Fillmore, 2014). Although vocabulary is a key source of difficulty for students' comprehension of science texts (e.g., Coxhead, Stevens, Tinkle, 2010), academic discourse also contains complex structures such as long noun phrases (Biber & Gray, 2011) that are problematic for ELLs to understand (Abedi, 2006).

In Bernhardt's (2005, 2011) compensatory model of second language (L2) reading, grammatical knowledge plays an important role due to its strong relationship with L2 passage level reading comprehension (Jeon & Yamashita, 2014). Based on Anderson's (1993, 2007) skill acquisition theory and findings from pedagogical research grounded in Systemic Functional Linguistics (e.g., Fang & Schleppegrell, 2010), instruction and practice in unpacking meaning from complex noun phrases should help improve ELLs' academic reading comprehension and fluency; however, the effect of this type of instruction on high school ELLs remains to be investigated through a classroom-based experiment. To fill this gap, I conducted a randomized pre- post-test experiment with 81 high school ELLs. Three groups were provided with 15 hours of instruction. Each group read identical grade 10-level science passages followed by different activities: group one answered comprehension questions; group two studied Academic Word List (Coxhead, 2000) vocabulary that appeared in the passages; group three analyzed the meaning of complex noun phrases from the passages. MANOVA was used on pre-post reading comprehension and fluency gain scores to examine the comparative effects of the three interventions. Participants' comments from a post-intervention questionnaire and researcher field notes were thematically analyzed to augment the quantitative findings. Results provide evidence to support linguistic instruction beyond vocabulary.

Preface

This thesis is an original work by Kent K. Lee. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name: "Improving Reading Comprehension of Scientific Texts: A Comparison of Traditional, Academic Vocabulary, and Grammar-based Online Modules for K-12 English Language Learners", No. Pro00066067, July 24, 2017.

Acknowledgements

I would also like to express my sincerest gratitude to the participants, their teachers, and school board consultants and administrators who were paramount to this study. To the faculty members of the TESL Program, Dr. Tracey Derwing, Dr. Marian Rossiter, Dr. Leila Ranta, and my supervisor Dr. Marilyn Abbott, at the University of Alberta, thank you for your guidance and encouragement during the entire course of my graduate studies. Also, thank you to all of my other professors who have contributed to my development as a scholar. I am also grateful to Dr. William Dunn and Dr. Florence Glanfield in the Department of Secondary Education for the opportunity to grow as a teacher-trainer.

I am thankful to the members of my Dissertation Committee for providing me with invaluable advice. More importantly, my gratitude to Dr. Marilyn Abbott is beyond what words can describe – thank you for being a dedicated mentor.

I would like to thank my family for their unwavering support. Most importantly, I wish to thank my loving wife, Amie, for her patience, understanding, and believing in me throughout my studies.

Finally, the financial supported provided by the Social Sciences and Humanities Research Council of Canada has allowed me to devote my time and energy to the completion of this dissertation.

Table of Contents

Abstract .... **ii**
Preface .... **iv**
Acknowledgements .... **v**
Table of Contents .... **vi**
List of Tables .... **ix**
List of Figures .... **x**
Chapter One – Introduction **1**
Introduction 1
Study Overview 5
Research Questions 6
Organization of the Dissertation 7
Chapter Two – Literature Review **8**
Academic Performance of ELLs 8
Defining Reading and Theoretical Framework 14
Components of Reading – Lower Level Skills 15
Components of Reading – Higher Level Skills 19
Models of Reading 20
L2 Reading 21
L2 Reading Model 23
Components of L2 Reading 24
L2 Grammar and Reading Comprehension 27
L2 Vocabulary Knowledge and Reading Comprehension 33
Inseparability of Lexis and Grammar 39
Assessing Silent Reading 45
Role of Text in Science Education 47
Academic Language and the Language of Scientific Texts 50
SFL and Scientific Writing 55
Nominalization 57
Compressed Writing 58
Text Cohesion, Thematic Structure, and Nominal Groups 60
Lexical Density and Ambiguity 62

Academic Language Interventions 66
Methodological Issues and Computer Assisted Language Learning and Research 71
Review Summary 77
Study Overview 78
Chapter Three – Method 80
Introduction 80
Overall Research Design 80
Textbook Analysis – Tools and Procedures 82
Pilot Phase – Participants and Setting 84
Pilot Phase – Instruments 85
Pilot Phase – Procedures 91
Phase Two – Participants and Setting 93
Phase Two – Instruments 94
Phase Two – Procedures 94
Data Analysis 95
Chapter Summary 96
Chapter Four – Results 97
Introduction 97
Textbook Analysis 98
Comparative Effects of the Intervention Conditions 104
Phase Two Results 109
Post-Intervention Questionnaire Results 115
Results Summary 119
Chapter Five – Discussion, Summary and Conclusions 120
Overview 120
Discussion of Results 120
Research Question 1: 120
Research Question 2a: 125
Research Questions 2b & 3: 126
Research Question 4: 133
Research Questions 5 & 6: 136
Discussion of Contributions to the Research Community 138
L1 and L2 Reading 138
SLA 141

CALL as a Research Platform *144*
Pedagogical Implications 144
CALL as a Tutor *149*
Limitations and Future Directions 152
Conclusion 155
References 157
Appendix A 198
Appendix B 199
Appendix C 202
Appendix D 206
Appendix E 215
Appendix F 221

List of Tables

Table 1 Word Counts by Types and Tokens, and Calculated Lexical Density 99
Table 2 Lexical Coverage by the AWL 100
Table 3 Descriptive statistics of group performance on dependent variables from the pilot study
.... 107
Table 4 Pilot phase post hoc Tukey HSD multiple comparisons.. 109
Table 5 Descriptive statistics of group performance on dependent variables from the Phase Two
.... 110
Table 6 Phase Two post-hoc Tukey HSD multiple comparisons 115

List of Figures

Figure 1. Noun phrase constituent example including parts of speech and information type...... 58
Figure 2. Syntax tree of an example noun phrase containing multiple levels of embedding....... 60
Figure 3. Lexical frequency profile of both textbooks by cumulative and unique token coverage 102
Figure 4. Syntactic complexity measured as average of number of complex nominals per T-unit and clause.. 103
Figure 5. Coh-Metrix measures of text readability 104
Figure 6. Scatterplot of individual gain scores versus pre-test scores on the knowledge of nominal groups test.. 112
Figure 7. Scatterplot of individual gain scores versus pre-test scores on the silent reading fluency and comprehension test (both forms combined). 113

Chapter One – Introduction

Introduction

In recognition of the importance of immigration to the economic future of Canada, the Canadian government, through its multi-year Immigration Levels Plan, has called for annual increases in immigration targets from 2018-2020 totalling approximately one million newcomers (Immigration, Refugees and Citizenship Canada, 2017a, 2017b). Consequently, immigrants who do not speak English as a first language (L1) will continue to arrive and their children will enroll in kindergarten to grade 12 (K-12) schools. Both Statistics Canada and Immigration, Refugees and Citizenship Canada (IRCC) project that by 2036, immigrants or children of immigrants could constitute half of the Canadian population (Morency, Malenfant, & MacIsaac, 2017) and that more than a quarter of the population could speak a mother tongue other than English or French (Houle & Corbeil, 2017). For example, since 2015, 40,081 Arabic speaking Syrian refugees have been resettled in Canada (Government of Canada, 2017). Learning one of Canada's official languages is critical for successful resettlement, yet remains a challenge for many newcomers and their children (e.g., CBC News, 2016).

Young immigrants may face difficulties in social, educational, and health arenas during the resettlement process (Ngo & Schleifer, 2007); some even gain undesired first-hand experience with the justice system (Rossiter, Hatami, Ripley, & Rossiter, 2015; Rossiter & Rossiter, 2009). While many agencies serve the needs of immigrants, schools in the K-12 system arguably play the most influential role in the lives of children and youth partially because the larger part of their day is spent in schools. Teachers, school counselors, and other staff are charged with forming a scaffold of support for immigrant students in transition to becoming contributing members of society. In fact, the Province of Alberta mandated school authorities in the K-12 system to inspire students to "reach their full potential by developing the competencies of Engaged Thinkers and Ethical Citizens with an Entrepreneurial Spirit, who contribute to a strong and prosperous economy and society" (Alberta Education, 2013a); the corollary of this government ministerial order is that all students, including English language learners (ELLs), are expected to rise to the challenge.

Indeed, many young immigrants gladly accept the challenge, as immigrants typically view education as the key to a prosperous life in Canada (e.g., Krahn & Taylor, 2005; Shakya et al., 2010). Language is key. From a sociological perspective, learners are invested "in the target language…with the understanding that they will acquire a wider range of symbolic resources (language, education, friendship) and material resources (capital goods, real estate, money) which will increase the value of their cultural capital and social power" (Norton, 2013, p. 6). French-speaking immigrants arriving in Alberta have the option of attending Francophone schools; however, for non-English speaking immigrants, studying academic content in an additional language often presents a significant barrier to realizing their educational ambitions. School boards across Canada are already seeing their English language learner (ELL) demographics exceeding the aforementioned projections for Canada. In Alberta, the two major school districts in Edmonton reported having experienced a quintuple increase in English language learner (ELL) enrollment over the past decade, with ELLs representing approximately 25% of the student population in 2015 (Zabjek, 2015). Similarly, 46% of classrooms in British Columbia had one or more ELLs in the 2016-2017 school year (BC Ministry of Education, 2016a). In Ontario, where the majority of permanent residents initially land, the Toronto District School Board (Toronto District School Board, 2018) reported that over 50% of their student population speak a language other than English at home.

Provincial governments recognize that ELLs require additional supports to thrive in a relatively monolingual education system. Although subgroups exist in the ELL population who match or exceed native English speakers' (NSs) performance in schools, Garnett (2010) observed that grade 12 ELLs in British Columbia did not always "attain returns on their academic investments equal to those of [NSs]" (p. 704). Data from provincial ministries of education in Alberta in British Columbia support Garnett's findings. For example, Alberta Education (2009) found that students who were coded as English as a second language (ESL) in grades seven to nine, after three years of entering high school, had a lower high school completion rate (61.7%) compared with students without ESL coding (75.6%). British Columbia reported even more dismal results: only 53% of the ELL students in the 2015-2016 grade 12 cohort completed high school on their first attempt (BC Ministry of Education, 2016b). These figures do, however, rise with additional years of high school enrollment.

To receive a high school diploma, students must complete a number of mandatory courses. The common issue that hinders many students' learning is their weak reading skills (e.g., August & Shanahan, 2006). In Alberta, graduation requirements include credits in grade 12 level English language arts, social studies, and grade 11 level mathematics and science (Alberta Education, 2017). These core subjects require higher-level English reading comprehension skills that ELLs are in the process of developing. In many instances, these courses also serve as prerequisites for further post-secondary education and vocational training.

Room for improvement exists in the preparation of English language learners (ELLs) for academics beyond high school, especially in the area of academic language development (e.g., Valdés, 2004; Wong Fillmore, 2014). In classrooms where English is the language of instruction, ELLs' competence in oral communication during daily classroom routines may mask their challenges with reading comprehension (e.g., Cummins, 2008). Expository texts, such as those found in science textbooks, pose additional comprehension challenges for poor readers due to the use of discipline-specific academic language that is atypical of spoken English (e.g., Fang & Schleppegrell, 2010). Teachers of content-area subjects at the secondary level may struggle to meet the linguistic needs of ELLs (e.g., Bruna, Vann, & Escudero, 2007; Duff, 2001), hold different conceptualizations of academic English (DiCerbo, Anstrom, Baker, & Rivera, 2014), possess negative attitudes towards the inclusion of ELLs in their mainstream classrooms (Harrison & Lakin, 2018), or deem reading instruction as beyond their responsibility (Gunderson, 1986).

Nonetheless, reading plays a crucial role in instruction and assessment in the education system. Unfortunately, Toohey and Derwing (2008) found that many ELLs steer "clear of examinable subjects that had a heavy language component" (p.188), and that they "either fail textually based courses, or they do worse in secondary school examinations in those areas than in other subjects" (p. 191). Given that all four of the core subjects required for graduation contain some degree of reading, ELLs' avoidance of and failure in these courses potentially jeopardizes their futures.

A logical solution would see teachers devote more time to reading instruction, particularly as part of content courses. However, the task of helping ELLs is often bestowed upon ESL teachers or education assistants who typically do not have expertise in all subject content matter at the high school level – nor can they realistically be expected to possess such knowledge. These ESL teachers often work in isolation from their content-area peers (Bell & Baecher, 2012). Secondary school teachers generally hold a feeling of inadequacy in their ability to meet the linguistic and academic needs of ELLs (Rubinstein-Avila & Lee, 2014). Although all

Canadian K-12 teachers must hold a minimum of a bachelor's degree in education, Canadian universities have yet to universally adopt courses in ELL pedagogy as part of mandatory teacher training programs. Continuing professional development may assuage the oversight in training concerning second language education pedagogy, but professional learning sessions on teaching ELLs may amount to only a few hours over the course of a school year. Unfortunately, as Goldenberg (2013) remarked, "it is an inconvenient truth that we lack the knowledge base to fully prepare teachers to help…[ELLs] overcome the achievement gaps they face" (p. 10). Therefore, efforts must be made by educational researchers to identify effective interventions that could help ELLs improve their English reading skills so they can comprehend the complex content area subject materials found in courses such as high school science. Although the literature for ELL educators offers a number of suggestions for accommodated instruction and assessment in science (e.g., Munro, Abbott, & Rossiter, 2013), few experimental studies have been conducted to examine the effectiveness of linguistic interventions for improving ELLs' reading of scientific texts, particularly at the high school level.

Study Overview

The purpose of this study was to address the gap in research examining reading comprehension interventions designed for high-school aged ELLs. I investigated the impact of vocabulary and grammar instruction on ELLs' reading comprehension performance, as well as the students' experiences with the interventions. To begin, I analyzed two Alberta Education approved grade 10 Science textbooks to understand the type of text (vocabulary, grammar, cohesion) that ELLs may be exposed to in a content-area classroom. Using passages from the textbooks, I created three different Internet-based modules (30 sessions each): traditional reading comprehension (answering comprehension questions and learning Science 10 technical vocabulary), Academic Word List (Coxhead, 2000) vocabulary, and scientific grammar. Approval to conduct a classroom-based experiment with ELLs was granted by a local K-12 school board in Alberta; 81 high school ELLs from five ESL classes in two schools participated in the study. A pilot study was first conducted with 34 students. A pre-test, intervention, post-test design used in both the pilot study and the main study allowed me to examine the relative effects of the three types of online modules on ELLs' reading comprehension, silent reading fluency, and ability to interpret noun phrases. In addition, a survey with follow-up interviews was administered to explore the students' experiences using the modules, namely, the benefits and challenges they perceived. Ultimately, my intent was to assess the effects of my interventions on students' knowledge, skills, and attitudes.

Research Questions

1. What are the linguistic characteristics of the two Alberta Education-approved Science 10 textbooks in terms of lexicon, syntactic construction, and passage-level readability?
2. What are the comparative effects of three online modules (traditional reading comprehension, AWL vocabulary, scientific grammar) on ELLs' passage-level reading comprehension
    a. as measured by a commercialized standardized exam?
    b. as measured by maze passages generated from a Science 10 textbook?
3. What are the comparative effects of three online modules (traditional reading comprehension practice, academic vocabulary practice, or grammar practice) on ELLs' passage-level reading fluency, as measured by performance on maze passages constructed from a grade 10 science textbook?

4. What are the comparative effects of three online modules (traditional reading comprehension practice, academic vocabulary practice, or grammar practice) on ELLs' understanding of nominal structures as measured by a noun phrase interpretation activity?
5. What are the ELLs' attitudes (positive or negative) toward the use of the modules in their classes?
6. What are the ELLs' conceptions of the role of the online modules in their learning?

Organization of the Dissertation

In Chapter One, above, I situated my research in the Canadian context, outlined the purpose of the research, presented the research questions, and provided an overview of the research design. In Chapter Two, I review the relevant literature on the academic performance of ELLs in Canada as a justification for the need for reading instruction. Then I move to a review of the reading theory and research on the components of L2 reading. The linguistic features of academic language used in science are explored, followed by a review of pedagogical practices centred on those features. In Chapter Three, I describe the methodology of my study, which was designed to fill the gaps in research identified in Chapter Two. The results of the textbook analysis and intervention are presented in Chapter Four. A discussion of the results follows in Chapter Five. I conclude with pedagogical implications, limitations, and future directions.

Chapter Two – Literature Review

Addressing the academic needs of ELLs necessitates a review of interdisciplinary knowledge of education, linguistics, and second language acquisition. I begin with a review of research on the academic performance of ELLs in the Canadian context. Then I shift the focus to theoretical constructs related to reading and subskills of second language (L2) reading. Next, I consider the role of language in science education, the special linguistic features of scientific language, and suggested pedagogical practices for helping ELLs better comprehend scientific language. Methodological issues of quantitative, classroom-based L2 research are also reviewed.

Academic Performance of ELLs

Researchers sometimes use the immigrant/Canadian dichotomy when examining the educational outcomes of immigrant youth. The term 'immigrant' may encompass foreign-born youth, as well as children of first generation immigrants. Putting aside the question of "when does one stop being an immigrant?", the reality is that immigrants are confronted with a combination of challenges that could potentially jeopardize their futures. Ngo and Schleifer (2007) conceptualized an interactive model of issues that immigrant youth face in their daily lives, including "linguistic, acculturative, psychological and socioeconomic challenges" (p. 29). This model comprises four arenas: social services, health, justice, and education. Each arena spans three environments: family, school, and community. First, the social services arena includes issues such as cultural adjustment, social support and belonging, cultural identity, sexual identity, gender roles, internalized racism, and employment. Second, health refers to nutrition, communicable diseases, sexual and reproductive health, chronic health conditions, mental health problems, and migration-related trauma. Third, justice recognizes problems such as transitions, criminal gangs and violence, substance abuse, prostitution, and issues in the youth justice process. Finally, education involves unfamiliarity with schooling in Canada, appropriate assessments, ESL instruction, support in content classrooms, support for students with special needs, and support for heritage languages. Although this is not an exhaustive list of needs, Ngo and Schleifer's model illustrates the complexity of the interplay between numerous factors that influence the well-being and success of immigrant youth. Research in the area of pedagogical support for ELLs in content classrooms is one avenue to addressing needs in the education arena.

Without adequate English language proficiency, ELLs may confront difficulty demonstrating their intellect in a largely monolingual educational environment. The acquisition of academic skills, particularly English reading comprehension skills, provides students with advantages leading to academic mobility and advancement, and contributes to overall quality of life. For example, Kanno and Kangas (2014) found that despite earning high marks in ELL sheltered content classes, ELLs at an American high school were unable to enroll in higher-stream courses because of their failure to demonstrate strong reading skills on standardized exams. Research in the Canadian K-12 education context also suggests that subgroups of ELLs have lower completion rates, higher dropout rates, and/or lower grades on standardized exams than their native English-speaking peers (e.g., Alberta Education, 2009; BC Ministry of Education, 2016b; Garnett, 2010; Gunderson, D'Silva, & Odo, 2012; Odo, 2012; Roessingh & Douglas, 2012a, 2012b; Toohey & Derwing, 2008).

The latest publicly accessible *High School Completion Longitudinal Study* available from Alberta Education (2009) revealed that only 61.7% of ELLs who were coded as English as second language (ESL) in junior high school (grades 7-9) completed high school within three years of beginning grade 10, compared with 75.6% of students with no ESL coding. Given another year, the statistics rose to 68.5% compared with 80.8% for non-ESL students. In contrast, students who were coded as ESL only during grades three to six showed parity in completion rates with non-ESL peers (75.4% and 75.6%, respectively). Although the proportion of ESL students in the analyzed data set was small ($n = 577$) compared with non-ESL students ($n = 30,044$), the lower ELL completion rates nevertheless suggest the need to provide additional ESL support to ELLs at the high school level.

Graduation rates from British Columbia (BC Ministry of Education, 2016b) show similar patterns to those in Alberta. In the 2015-2016 academic year, the graduation rate of all eligible ELLs including repeaters (87.1%) was higher than the general student population (83.6%). However, of those who were in grade 12 for the first time, only 53% of ELLs graduated, compared with 79% for the general student body. It appears that ELLs require more time to finish high school than native speakers (NSs). A closer examination of ELL performance in grade 10 yields additional insight; for example, from 2011 to 2016, a consistent percentage of approximately 48% to 50% of ELLs in BC received a grade of C or lower on their Science 10 exams (BC Ministry of Education, 2016c). Therefore, the additional time required to finish high school may be a cumulative effect of unmet needs that were present at the beginning of an ELL's high school career.

Toohey and Derwing (2008) investigated data from the Vancouver school district that included 1,554 ESL students between 1997 and 2002. They found that students avoided language-demanding classes and "were far more likely to write examinations in subjects that have a strong visual or hands-on component, such as Math and Chemistry" (p. 13). Of the 933 ELLs in grade 12 who graduated, over 75% of them wrote the English and Mathematics provincial exams, 61% wrote Chemistry, but the percentages dropped to 35% and to 26% for Physics and Biology, respectively. Despite English being a linguistically demanding subject, the authors believed that graduates wrote the grade 12 English exam because it was a graduation requirement. Far fewer non-graduates wrote English (9%) than math (15%) exams. Some ethnic group differences were also observed, with the Japanese and Chinese (Mandarin and Cantonese) first language (L1) student groups being the only ones to achieve graduation percentages over 70%. The authors attributed these groups' success to their higher socio-economic status (as indicated by their independent class immigrant status and the location of their schools).

Odo (2012) examined the disparity of failure rates between ELLs and mainstream students by scrutinizing the validity of British Columbia's provincial high school exit exam questions and results. Although Odo echoed Garnett's (2010) concern that the ELL label potentially shrouded other factors affecting performance (e.g., SES), the results from provincial exams between 2006 and 2011 show unequivocally that ESL students were at a disadvantage. Even mathematics, arguably the least linguistically demanding core subject, contains word problems: ELL failure rates were double those of other students. Closer examination of the exams revealed that questions favoured students with cultural background knowledge and English language proficiency, both of which ELLs are likely to lack. Similar to researchers in the United States (e.g., Abedi, 2006; Menken, 2008), Odo suggested that linguistic complexity, coupled with the lack of appropriate exam accommodations (e.g., dictionary use), seriously undermines the validity of standardized government exam results, which in turn jeopardizes ELLs' academic success. However, the fact remains that standardized exams serve as a check and balance of the education system and, as such, hold both students and their schools accountable. Even with exam accommodations such as allowing dictionaries or extra time, bolstering ELLs' linguistic knowledge remains a necessity.

ELLs' graduation from Canadian high schools signifies an academic milestone; for many, it marks a transition to post-secondary education (Thiessen, 2009). Earning a high school diploma, however, does not necessarily predict preparedness for or success in post-secondary study. Roessingh and Douglas (2012a) found that high school graduate ELLs may be linguistically ill-prepared for studies at university. They examined grade 12 Alberta provincial diploma exam scores in English 30-1 and Math 30-1 from 128 ELLs who entered university. Their analysis revealed significant differences between ELLs' and native speakers' scores on both exams, with ELLs outperforming their native-speaking peers in Math but lagging behind in English. Roessingh and Douglas argued that ELLs' strength in Math indicates intellectual competence; thus, their academic vulnerability lies in language proficiency: ELLs failed the diploma exam in English at significantly higher rates than NSs. These results support the argument that ELLs' linguistic proficiency may be insufficiently developed to meet post-secondary literacy demands.

For further analysis, from the same pool of 128 ELLs in Roessingh and Douglas (2012a), Roessingh and Douglas (2012b) randomly selected 30 Alberta high school graduates who completed their degrees at the University of Calgary. These 30 ELLs were subdivided into two equal groups based on their age of arrival to Canada, using 14 years of age as the boundary. Graduation rates, grade point averages (GPAs), and total semesters taken were compared with a NS group ($n = 15$). Common degrees attained by the ELLs ($n = 11$) and NSs ($n = 9$) included engineering, computer science, nursing, and psychology. The ELLs also graduated with degrees in commerce, economics, international relations, biology, chemistry, and actuarial science; NSs graduated from programs in linguistics, political science, Canadian studies, sociology, and zoology. Encouragingly, ELLs proportionally outnumbered NSs in graduation rates; however, an examination of their GPAs revealed a different picture. Only 13% of the ELLs who arrived in Canada under 14 years of age graduated with a GPA of over 3.0 (of 4.0 maximum), and none who arrived over 14 attained a GPA of higher than 3.0. In contrast, 33% of their NS peers earned a GPA of greater than 3.0. While one may be satisfied with a bachelor's degree as an end goal, Roessingh and Douglas argued that a GPA of less than 3.0 jeopardizes one's admissibility for graduate studies. Furthermore, ELLs on average required more semesters to graduate from university and were more likely to be placed on academic probation during their program due to subpar academic performance. Unfortunately, while diverse faculties were represented, the random selection of participants did not yield entirely comparable programs of study. Regardless, the authors advocated for support for K-12 ELLs prior to university to prepare them to meet the linguistic demands of university.

Clearly, if we desire an equitable education for all Canadians, then linguistic support must be provided to ELLs. If this support begins in K-12 schools, Canada may better cultivate the intellectual capital found in the ELL population. Low SES may limit ELLs' and their families' abilities to independently access extra support (e.g., tutoring); therefore, the moral obligation of providing students with necessary supports (e.g., academic language support, exam accommodations, more time to complete high school) rests with the education system. More specifically, schools must help students develop competent reading abilities. Mackay and Tavares (2005) surveyed teachers in Manitoba and found that an overwhelming majority (88%) identified limited literacy skills, defined as reading and writing, as a barrier to success for refugee students. The role of reading in school and academic success is obvious: by high school, students must be able to read to learn (Chall, 1996). In addition, reading provides one with "the ability to understand and employ printed information in daily activities at home, at work and in the community – to achieve one's goals, and to develop one's knowledge and potential" (OECD & Statistics Canada, 2011, p. 316). Canada is a literate, print-based society in which the ability to read and interact with textual information, paper or digital, not only enhances one's quality of life but is necessary for becoming a fully functioning member of society.

Defining Reading and Theoretical Framework

Defining reading is far from a simple task. Parents may boast that their young child's ability to sound out the letters and words on a page of a book is being able to read, even if the child does not yet possess full comprehension of the text. Reading involves more than the extraction of phonological information from printed words. While knowledge of the correspondence between the written letters and words and sounds is essential, according to Koda (2010), "the ultimate goal of reading is to construct text meaning based on visually encoded linguistic information" (p. 5). Doing so requires readers to utilize a complex set of cognitive processes and subskills, and reading is the composite of these multiple components in operation (e.g., Bernhardt, 2011; Grabe, 2009; Koda, 2005, 2007; Stanovich, 2000). In other words, reading is more than making accurate letter-sound (grapheme-phoneme) correspondences; reading occurs when a reader also comprehends a text.

Skill acquisition theory or more specifically, Anderson's (1993, 2007) Adaptive Control of Though-Rational theory (ACT-R) provides a theoretical framework for understanding the general reading process. The ACT-R model helps researchers in their attempt to understand the organization of knowledge and production of intelligent behaviour by the human mind. According to this theory, skill acquisition follows three phases. In the first cognitive phase, learning begins with explicit acquisition of relevant knowledge (declarative knowledge) which requires conscious effort. In the second associative phase, explicit knowledge is converted to implicit knowledge through practice (Ellis, 1994). Finally, automaticity is achieved as the access of knowledge becomes effortless due to the strengthening of associated pathways in long-term memory from practice. Reading is a learned skill; learning to read in a second language (L2) proceeds through the process as outlined by ACT-R. The ACT-R coincides with the strong interface position in second language acquisition (SLA) theory: that is, explicit (declarative) knowledge that L2 learners gain (i.e., from instruction) can become implicit (procedural) knowledge through meaningful practice (DeKeyser, 2007). The implication of ACT-R for the present study is that reading and its subcomponents are skills that can be taught explicitly and that their execution may become streamlined through repeated practice.

Reading skills can be broadly divided into two categories: lower-level and higher-level. These two sets of processes work in tandem to allow the reader to comprehend written text (e.g., Perfetti, 1999). Based on research from first language (L1) reading, a number of different models of reading exist that explain the two systems' contribution to reading comprehension (e.g., Hoover & Gough's [1990] Simple View of Reading; Kintsch's [1988, 1998] Construction-Integration Model; Stanovich's [1980, 2000] Interactive-Compensatory Model; van den Broek, Risden, Fletcher, & Thurlow's [1996] Landscape Model). Researchers generally agree that different component skills are involved in reading, regardless of the model, with the implication that deficiencies in any of the component subskills may compromise reading for the reader. To develop my reading interventions, I needed to first identify targeted component skills as instructional goals.

Components of Reading – Lower Level Skills

Lower-level skills refer to a set of skills that contribute to fluent reading once automatized (Koda, 2005). These skills include word recognition, syntactic parsing, and semantic-proposition encoding, and they are thought to occur in working memory (Grabe, 2009). Efficient word recognition is a significant contributor to reading comprehension (e.g., Perfetti, 2007; Perfetti, Landi, & Oakhill, 2005). According to Perfetti (2007), lexical knowledge includes properties of phonology, orthography, grammar, and meaning (along with a fifth element of how well the four are bound in a lexical item); limited mental representations of these properties hamper reading.

In spoken language, phonology involves the combination of individual sounds to form meaningful units; in signed languages, phonology refers to handshapes and movements that together encapsulate meaning (MacSweeney, Waters, Brammer, Woll, & Goswami, 2008). Phonological processing contributes to word reading ability and has three aspects: phonological awareness, phonological recoding, and phonological memory (Lesaux & Geva, 2006). Phonological awareness is the conscious awareness of sounds or handshapes in a language. Phonological recoding is the conversion of a visual stimulus into its corresponding sound(s)/sign(s). Phonological memory is the coding of auditory or visual information in working memory for temporary storage. Phonological processing has been shown to be a predictor of reading success in studies with monolingual children (Adams, 1990, 1999).

Orthographic processing refers to visual recognition of writing systems and may lead to word recognition via an alternate route that relies only partially on or completely bypasses phonological processing (i.e., sight word recognition) (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Cunningham, Perry, & Stanovich, 2001). Where reading in English is concerned, longer words (i.e., more letters) require more word-recognition time (Rayner & Pollatsek, 1989). Semantic and syntactic processing possibly contribute to word recognition via spreading activation mechanisms (Coltheart et al., 2001): when readers access a word in their repertoire of known vocabulary (lexicon), the recognized word spreads activation energy to surrounding words, thus stimulating recognition of those associated words (Grabe, 2009).

Through effective practice, readers become more efficient at phonological, orthographic, semantic, and syntactic processing. As a result, retrieval of words from one's lexicon becomes rapid and requires few resources (i.e., automatic) (Perfetti, 2007). Rapid processing of words well represented in the reader's lexicon allows for "processing resources to be devoted to higher level comprehension" (Perfetti & Hart, 2001, p. 76). Ineffective or non-automatic word recognition could, therefore, compromise reading (Perfetti, 1985).

While word recognition has importance, the other two lower-level processes also play vital roles in reading. Syntactic parsing involves deriving meaning from text through the use of grammatical knowledge to access information from words and sentence structures (Fender, 2001; Grabe, 2009; Perfetti et al., 2005). In L1 reading, syntactic awareness appears to predict reading comprehension (Nagy, 2007). For example, in a four-year study with 23 L1 English-speaking children, Demont and Gombert (1996) found that syntactic awareness (as measured by performances in correction of morphemic anomaly and correction of asemantic and agrammatical sentences) in kindergarten and grade one accounted for 3.3% to 24.4% of the variance in reading comprehension measured at the end of grades two and three. Likewise, Nation and Snowling (2000) found syntactic awareness separated good readers from poor ones. They compared the reading comprehension scores of 15 children who were poor comprehenders (reading comprehension was at least one year below grade level) with the scores of 15 normal readers, matched for age, nonverbal ability, and decoding skills. Poor comprehenders performed significantly worse than normal readers on word order correction tasks. Therefore, the development of struggling readers' grammatical knowledge as part of reading interventions appears to be warranted.

The process of syntactic parsing occurs almost as simultaneously as word recognition; grammatical information and syntactic processing constrain ambiguity, help to track referents, construct phrasal and clausal units, and set up processing expectations (e.g., "the verb *put* …will [likely] be followed by an upcoming noun phrase and locative prepositional phrase") (Grabe, 2009, p. 200). The meaning of clauses and sentences are constructed using information from word recognition and syntactic processing, leading to the creation of semantic proposition units (Fender, 2001; Kintsch 1998; Perfetti & Britt, 1995). These units of information are linked in that they "establish key referents and predicates, track overlapping referents and index the relative importance of referents, events, and descriptors" (Grabe, 2009, p. 201). In short, grammatical information helps readers to interpret a text as a cohesive whole and to arrive at the author's intended message. Difficulty in reading, as measured by processing time, can increase with more complex or ambiguous sentence structures (Fender, 2001). Students unfamiliar with certain sentence patterns (e.g., phrasal elaboration and complex noun phrases found in academic writing used in textbooks) could experience further reading difficulty.

As stated above, lower-level processing occurs in working memory. Baddeley and Hitch (1974) conceptualized a model of working memory consisting of a central executive, a phonological loop, and a visuo-spatial sketch pad. Generally considered as limited in capacity, working memory both stores and processes information. The central executive dictates the allocation of attentional resources; the visuo-spatial sketch pad stores and processes incoming visual information; and new words are phonologically rehearsed and stored in the phonological loop of working memory (Baddeley, 2006, 2015). Although an extensive examination of research on working memory is beyond the scope of the present study, Daneman and Merilke's (1996) meta-analysis of 77 studies on working memory and language comprehension (6,179 participants) indicated the existence of strong correlations between the storage and processing capacity of working memory and L1 reading comprehension. In a more recent meta-analysis of studies that focused on readers between 8 and 30 years of age, Carretti, Borella, Cornoldi, and De Beni (2009) found results consistent with Daneman and Merilke's meta-analysis. Grabe (2009) explained working memory as necessary for active storage and processing at the word and clausal levels, the suppression of distracting information, and retention of ideas from sentences to be later assembled with other ideas into a coherent whole (i.e., higher-level processing). In simpler terms, readers must remember and incorporate what they just read while concurrently decoding and processing new information. Faster L1 readers at the secondary grade level (grade eight) have been observed to be better comprehenders as a result of decreased cognitive processing load (van Silfhout, Evers-Vermeul, Mak, & Sanders, 2014).

More research, however, is still required to ascertain the role of working memory in L2 reading comprehension. In a recent study, Indrarathne and Kormos (2018) showed that better working memory abilities provide undergraduate L2 English learners with advantages in learning new, targeted English grammatical forms. To address the differences in working memory abilities between L2 learners, the researchers advocated for "extensive exposure to target syntactic constructions and additional instructional support" (p. 371). However, no further elaborations regarding additional instructional support were provided.

Components of Reading – Higher Level Skills

Although lower-level reading skills mainly concern vocabulary and grammar, knowledge of those components alone does not equate to the ability to read. Comprehension requires a reader to connect the lower-level pieces of what is being read to form a coherent whole and construct meaning from the text. Higher-level processing draws upon a reader's lexical, semantic, and syntactic knowledge, as well as background knowledge about the text to aid in comprehending the text as a whole, assuming that "the reader can direct attentional resources to [the] component skills" (Grabe, 2009, p. 39).

The contribution of higher-level skills to reading are illustrated in a text model and in a situational model of reading comprehension. According to the text model, as a reader begins to read, input from the text base activates meaning elements and is stored in working memory, and new information from additional words is added to the existing information network. As the network grows, some information overlaps, less relevant information becomes suppressed, and summary restriction occurs (Gernsbacher, 1990, 1997). This model represents the writer's intended message. Conversely, the situational model reflects the reader's approach to the text. As people read, they also amalgamate their own erudition with the text; success in comprehension relies on the integration of the reader's background knowledge and textual information (Kintsch, 1998). In fact, prior misconceptions of a topic have been demonstrated to adversely affect reading comprehension of scientific texts (van den Broek & Kendeou, 2008). Clearly, strong abilities in text-level processing, combined with an extensive and accurate background knowledge are conducive to reading comprehension.

Models of Reading

As previously mentioned, several theoretical models have been conceptualized to explain the reading process, at least for L1 reading. Traditionally, two paradigms exist in the literature: a bottom-up and a top-down approach to reading (Hudson, 2007). Bottom-up approaches hold that readers interact with text and build meaning in a linear fashion. Readers transform visual input from textual units such as letters, words, and sentences into phonemic units. Those units carry lexical meaning and readers continue to build meaning as they read; through this process, readers decode the writer's intended message. Alternately, top-down approaches assume that readers have preconceived notions about the text (i.e., background knowledge), continually predict upcoming information, and evaluate new incoming information against their background knowledge.

Currently accepted models of reading are neither purely bottom-up nor top-down, but rather interactive, in which both processes support one another (Grabe, 2009; Hudson, 2007). In the interactive paradigm, the dominant view since the early 1980s, various models attempt to account for the contributions of word recognition, syntactic parsing, and semantic-proposition encoding, working memory, and background knowledge. Of note is Stanovich's (1980, 2000) interactive-compensatory hypothesis; it has influenced the L2 reading model that informs the current study. In this model, Stanovich claimed that neither bottom-up nor top-down models adequately explain individual differences in reading ability. The strong version of bottom-up processing is serial in nature: that is, lower-level processes, such as word recognition, must be completed before higher-level processing can occur. Conversely, in a strictly top-down model, reading is conceptually driven, which Stanovich criticized as too time-consuming for even word recognition to occur. Stanovich explained that when readers lack ability in one level of processing, the other level can compensate for weaknesses. For example, a poor reader deficient in word recognition skills may exhibit greater reliance on contextual clues or other knowledge sources to arrive at text comprehension. Stanovich's model has informed several models of L2 reading, as described in the next section.

L2 Reading

A cluster of interrelated variables (L2 proficiency, L1 reading ability, age and educational experience at onset of L2 learning, cross-linguistic factors, dual language processing) differentiates L2 from L1 reading (e.g., Koda, 2007). According to the threshold hypothesis (Cummins, 1976, 1979), ELLs in the process of learning English need sufficient language skills to effectively process a text (e.g., Bernhardt & Kamil, 1995). Oral skills may also support L2 reading comprehension. For example, in the Simple View of Reading model (Hoover & Gough, 1990), reading comprehension is viewed as the product of word decoding and listening comprehension, where both interact to contribute to understanding text; empirical evidence from studies with elementary school ELLs (Gottardo & Mueller, 2009; Verhoeven & van Leeuwe, 2012) and adult undergraduate foreign language learners (Yamashita & Shiotsu, 2017) support the validity of this view in L2 reading. Beyond young ages, however, L2 reading is not so simple; thus, a different model of L2 reading is needed for older learners. Even highly fluent L2 adult bilinguals lack the automaticity in L2 word decoding skills possessed by L1 readers (Segalowitz, Poulsen, & Komoda, 1991). Some adult L2 learners have reading comprehension skills that far exceed their oral proficiency. Unlike Canadian-born NSs who learn to read as children, ELLs may begin learning to read in English well after childhood. As a result, older ELLs' "prior literacy experience …potentially provides substantial facilitation" (Koda, 2005, p. 7). On the other hand, those who lack educational experience in their native language or whose concepts of literacy conventions differ from those in English may face challenges in accessing textual information (e.g., Bell & Burnaby, 1984; Bigelow & Vinogradov, 2011; Kurvers, 2015). Cross-linguistic factors such as L1 reading proficiency certainly affect L2 reading (e.g., Cummins, 1991; Genesee, Geva, Dressler, & Kamil, 2006; Koda, 2005, 2007; Sparks, Patton, Ganschow, & Humbach, 2009, 2012; van Gelderen, Schoonen, Stoel, & de Glopper, 2007). L1 orthography and writing systems may affect L2 reading strategy use (e.g., Abbott, 2006; Bialystok, Luk, & Kwan, 2005). Issues with decoding orthographic information could hamper fluent reading, thereby overloading readers' short-term memory, ultimately hindering acquisition of new knowledge from the text (e.g., National Reading Panel, 2000). Limited mastery of vocabulary and syntactic knowledge also pose problems for L2 readers (e.g., Bernhardt, 2005; Schmitt, 2010; Verhoeven, 2000; Vermeer, 2001). Finally, monolingual information processing occurs during L1 reading, but L2 reading may involve dual languages (Koda, 2005). Clearly, L2 reading is an even more complex process than reading in L1.

Alderson (1984) posed an important question: are problems with reading in the L2 a language problem (e.g., vocabulary, grammar) or a reading problem (e.g., working memory)? The answer to Alderson's question – L2 reading interventions should focus on the problem areas – has pedagogical implications. Bernhardt (2005) maintains that language and reading are not mutually exclusive; that is, the question should be how much one accounts for the other. L1 reading research has shown that word recognition (Adams, 1990, 1999; Perfetti, 1999; Perfetti et al., 2005), vocabulary (e.g., Alderson, 2000; Laufer, 1992; Laufer & Ravenhorst-Kalovski, 2010), syntax (e.g., Gaux & Gombert, 1999; Nation & Snowling, 2000), and prior knowledge (e.g., Kintsch, 1998; van den Broek & Kendeou, 2008) affect reading comprehension performance; however, the existing models of L1 reading cannot be directly applied to L2 reading, as they do not account for the contribution of L1 skills to L2 reading.

L2 Reading Model

Only one general descriptive model of L2 reading exists to date. Informed by interactive models of reading such as Stanovich's (1980), Bernhardt (2005, 2011) attempted to elucidate the relationship between L1 and L2 reading skills and to create a model of L2 reading with substantial explanatory power. The result is a compensatory model of second language reading. It encapsulates variables that include L1 literacy (i.e., phonemics, text structure, purposes for reading, sentence configuration), L2 language knowledge (e.g., grammatical form, vocabulary knowledge, cognates), and a concept called unexplained variance (e.g., motivation, comprehension strategies) that accounts for how readers interact with the text. In this model, one source of knowledge compensates for deficiencies in others. Bernhardt (2005) noted that previous studies in L2 reading (Bernhardt & Kamil, 1995; Bossers, 1991; Brisbois, 1995; Carrell, 1991) estimated that L1 reading ability and L2 knowledge predicted 14-21% and 30% of L2 reading performance, respectively. Accordingly, in her model, Bernhardt postulated that L1 literacy accounts for 20% of the performance variance in L2 reading; L2 language knowledge, 30%; and unexplained, 50%. The three knowledge sources "operate synchronically, interactively, and synergistically" (Bernhardt, 2005, p. 140).

Components of L2 Reading

Similar to the approach taken in L1 reading research, a components-based interactive model of L2 reading allows for investigations of individual sub-skills' contributions to reading. Common constructs have been found to underlie L1 and L2 reading comprehension, as observed in research comparing ELL and NS students' reading skills (Lipka & Siegel, 2012). Components-based investigations have produced insights into the nature and degree of contributions of key predictors of L2 reading comprehension. For example, Lesaux, Koda, Siegel, and Shanahan (2006) analyzed research investigating factors that influence the development of L2 reading comprehension in L2 children in the K-12 context. They identified the following variables: word reading skills, understanding of literacy, L1 literacy, metalinguistic awareness, reader strategy, background/cultural knowledge, motivation, acculturation, text factors (e.g., genre, discourse style), and demographic factors (e.g., SES). In turn, the identification of key components of L2 reading helps to inform the development of evidence-based pedagogical interventions.

In support of a components view of L2 reading, Nassaji (2014) emphasized the importance of lower-level skills in L2 reading in a detailed synthesis that focused predominantly on L2 word recognition and its subprocesses (orthographic and phonological processing, and lexical access). Following the view of reading as a complex information processing task (e.g., Perfetti & Hart, 2001; Stanovich, 1980), Nassaji argued that reading is hierarchical in nature, whereby "successful operation of a higher-level process…depends on the efficient and accurate execution of those at the lower-level" (Nassaji, 2014, p. 4); higher-level processing (in reading) cannot occur in the absence of text, and automaticity of lower-level processing frees up attentional resources in working memory, which then allows for comprehension processing. The importance Nassaji places on lower-level processes is attributable to evidence from L2 reading research that showed lower-level processes as principal contributors to reading comprehension. For instance, efficient and precise word recognition has been shown to be a predictor of reading comprehension in beginning (e.g., Geva & Wang, 2001; Gottardo, Yan, Siegel, & Wade-Woolley, 2001; Grant, Gottardo, & Geva, 2011) and advanced L2 readers (e.g., Akamatsu, 2003; Favreau & Segalowitz, 1983; Nassaji, 2003; Nassaji & Geva, 1999; Shiotsu, 2009). These findings support other studies that demonstrated causal relationships between word recognition and reading comprehension for elementary aged L2 learners (Droop & Verhoeven, 2003; Verhoeven, 2000). Increasing the size of an ELLs' vocabulary would thus appear to facilitate reading comprehension.

In a longitudinal study, Geva and Farnia (2012) examined the dynamic nature of predictors of reading comprehension in 390 ELLs from multiple L1 backgrounds and 149 NSs in grade two and again in grade five. They assessed the students' cognitive abilities (pattern identification, working memory); phonological processing (naming speed, phonological awareness); oral English language proficiency (vocabulary, syntactic knowledge, listening comprehension); reading abilities (pseudo-word decoding, word identification, word and text reading fluency); and reading comprehension skills. Reading comprehension was assessed using a standardized multiple-choice exam. ELLs performed on par with their NS peers on cognitive abilities. As expected, phonological awareness and vocabulary measured in grade two predicted reading comprehension in grade five for both groups. However, only for the grade five ELL group did syntax, listening comprehension, and text fluency contribute unique variance to reading comprehension. Geva and Farnia argued that as L2 proficiency develops, ELLs are better able to utilize other language skills to help them process text, and that the associations with reading comprehension are manifested only when these skills are sufficiently developed. Furthermore, readings in higher grades possibly require language skills beyond vocabulary (e.g., syntactic knowledge), as opposed to those in lower grades where word recognition may contribute to adequate comprehension.

To identify key correlates of L2 reading comprehension, Jeon and Yamashita (2014) performed a meta-analysis of studies that examined L2 reading comprehension with various reading components. The analysis comprised 59 studies with 9,461 participants of all ages, L2 proficiency levels, and diverse L1 backgrounds. Inclusion criteria consisted of studies that reported a correlation coefficient (regardless of statistical significance) between passage-level L2 reading comprehension and at least one component of reading, and that examined participants with no language-related disabilities. The authors categorized the variables into high-evidence correlates (L2 decoding, L2 vocabulary knowledge, L2 grammar knowledge, L1 reading comprehension) and low-evidence correlates (L2 phonological awareness, L2 orthographic knowledge, L2 morphological knowledge, L2 listening comprehension, working memory, metacognition). In addition, moderator variables (age, L1-L2 orthographic distance, L1-L2 language distance [cognates, grammar], vocabulary measurement methods [productive/receptive, in-context/discrete], grammar measurement methods, and L2 proficiency) that may affect the strength of the relationship between L2 reading comprehension and the key correlates were taken into account. The four high-evidence correlates were identified as most frequently investigated in the literature with respect to L2 reading comprehension, while the six low-evidence correlates were less frequently so. Their results revealed the strongest correlates to be L2 grammar knowledge (r = .85), L2 vocabulary knowledge (r = .79), and L2 decoding (r = .56). Contrary to the researchers' expectations, they did not find L2 proficiency to be a significant moderator between L1 and L2 reading comprehension; however, due to a lack of standardized L2 proficiency assessments across the sample of studies, Jeon and Yamashita resorted to coding L2 proficiency into two levels, a decision that they regarded as possibly "too crude a method" (p. 195). In light of these rather large correlations, in the next two subsections, I focus more closely on the top two correlates: L2 grammar and L2 vocabulary knowledge.

L2 Grammar and Reading Comprehension

Grammar plays a prominent role in Bernhardt's (2005, 2011) compensatory model of L2 reading. Commonly-used related terms for grammar in the L2 reading literature include syntactic knowledge, syntactic awareness, and syntactic/sentence parsing/processing. The relationship between grammar knowledge and reading comprehension has been well established in L1 and L2 reading research; some researchers even consider grammar knowledge to be the dominant predictor of L2 reading (e.g., Jeon & Yamashita, 2014; Shiotsu, 2010). Recently, van den Bosch, Segers, and Verhoeven (2018) found young (between eight and ten years old) beginning L2 readers with low syntactic knowledge (as measured by a grammaticality judgement test) to show longer sentence processing times, and argued for the development of syntactic knowledge (in addition to decoding and vocabulary) as part of reading instruction for beginning L2 readers.

According to Gernsbacher's (1990, 1997) Structure Building Model of discourse comprehension, grammar contributes to reading comprehension in that it signals for readers to maintain or shift their mental representation of text structure. After words enter working memory, syntactic parsing and proposition forming occur instantaneously, thus initiating the building of text structure. As the first clause is read, a foundation for comprehension is laid, and succeeding information is mapped onto this foundation and a network structure is built. If a new clause being read overlaps with information in the existing network, it is readily incorporated into the foundation structure. On the other hand, new information involving new referents signals the building of a new foundation structure. As new information is read, it may be suppressed or added to the text structure by a suppression-elaboration mechanism, depending on contextual relevance, with the eventual goal of building a coherent discourse. An efficient suppression-elaboration mechanism contributes to reading comprehension. Inappropriate information has been shown to remain activated in poor readers, suggesting that they possess less efficient suppression mechanisms, leading to a greater tendency to shift to new foundation structures (Gernsbacher, 1991; Gernsbacher, Varner, & Faust, 1990). As Grabe (2009) summarized,

> weak readers have difficulty maintaining coherence … leaving many fragments that are hard to assemble … and retain much superfluous and irrelevant information that leads to overall processing inefficiencies. (p. 93)

Since grammatical information helps readers identify overlapping or new material, instruction to increase ELLs' grammatical knowledge could potentially improve their reading comprehension at the passage level.

As stated above, Jeon and Yamashita (2014) identified L2 grammar knowledge as the top contributor to reading comprehension. In their meta-analysis, they examined 16 studies that included participants from the sixth grade to post-graduate studies. For the purpose of the current study, I review only those studies that focused on reading in English and were conducted in the past 15 years.

In a study conducted in Israel, Abu-Rabia and Sanistsky (2010) compared the acquisition of English by trilinguals and bilinguals. They divided 80 sixth-grade students into two groups: an erroneously labeled experimental group (no experiment was conducted) that consisted of Russian students who were trilinguals (Russian L1, Hebrew L2, and English L3) and a control group with native-Hebrew speakers who were bilinguals (English L2). Students' cognitive and language abilities were measured, including their syntactic knowledge and reading comprehension. Syntactic knowledge was measured by a grammaticality judgment test (GJT). The results showed significant correlations between English syntactic knowledge and English reading comprehension ($r = .46$, $p < .01$). In addition, the trilinguals outperformed the bilinguals on English syntactic awareness and reading comprehension. Participants' L1 syntactic awareness also appeared to enhance their knowledge of grammar in their second and third languages. These results support Bernhardt's (2005, 2011) model of L2 reading that takes L1 proficiency into account.

August (2010) examined the relationship between the L2 grammar knowledge and L2 reading comprehension of Spanish-speaking adults between the ages of 18 and 55 who were living in the United States. Reading comprehension was measured by the Stanford Diagnostic Reading Test designated for grades 4.6-6.5. English grammar knowledge was assessed by the Michigan ESL Placement Test, which contained 30 multiple-choice items designed to test various forms (e.g., verb tense, articles, sentence connectors). August a found significant correlation between English grammar and English reading comprehension scores ($r = .31$, $p < .05$). As well, L1 grammar knowledge positively correlated with L1 reading, which in turn correlated with L2 reading. Learners with weaker L1 reading skills need to develop their L2 knowledge (e.g., academic vocabulary, complex linguistic structures) "more completely to compensate for the lack of L1 skills" (August, 2010, p. 258).

In a study with Mandarin L1 university students, Guo and Roehrig (2011) found that a combination of English syntactic awareness and vocabulary knowledge explained 87% of the variance in L2 reading comprehension. Guo and Roehrig measured syntactic awareness with two instruments: a test that required participants to combine sentences and an 11-item questionnaire that asked for explicit explanations of grammatical functions. The TOEFL Reading Comprehension Subtest and the Gray Silent Reading Tests served as the reading measures. Syntactic knowledge from both measures correlated positively and significantly with results from both reading tests.

Shiotsu and Weir (2007) reported results of three studies that investigated the contribution of L2 syntactic knowledge to reading comprehension. The first study involved 107 undergraduate ELLs at a university in the United Kingdom. Students' knowledge of grammar was measured by a test containing 30 multiple-choice items that presented a sentence stem and required students to choose from four possible responses that were semantically similar. Only one response, however, fit within the syntactic constraints. Reading comprehension was measured by an end-of-course test with 20 items: test item formats included short answer, true-false, and table/flowchart/sentence completion tasks. The correlation between grammar and reading was moderately high ($r = .62$, $p < .05$). In the second study, the participants were 181 Japanese L1 undergraduate ELLs in Japan. A twenty-question multiple-choice test containing four passages from two different tests served as the reading comprehension instrument. Syntactic knowledge was measured with a similar test to the one used in the first study. Results showed a higher correlation between syntax and reading ($r = .89$, $p < .05$). Participants in the third study were also Japanese L1 ELLs ($n = 591$) enrolled in Japanese universities. Data collection instruments were identical to the ones employed in the second study. L2 syntax was found to explain 72% of the variance in reading ($r = .85$, $p < .05$). Further subgroup analysis of higher and lower achievers also showed significant correlations between syntax and reading for both groups ($r = .62$, $p < .05$ and $r = .78$, $p < .05$, respectively). In all three studies, syntax also correlated significantly with vocabulary knowledge, and vocabulary correlated with reading; however, syntax appeared to predict reading performance better than vocabulary did in all three cases.

A longitudinal study was conducted by van Gelderen et al. (2004) with Dutch L1 ELLs in grades eight, nine, and ten. The researchers examined L2 reading comprehension and the contribution of various components (word recognition and sentence comprehension speeds, vocabulary and grammar knowledge, L1 and L2 metacognitive linguistic knowledge, L1 reading comprehension). Reading comprehension was measured using texts from previous research and national secondary school exams; the reading test consisted of six passages and 35 multiple-choice questions. Unfortunately, the test for syntactic knowledge was not well described, other than that it contained "difficult" items. The correlation between L2 grammar knowledge and reading comprehension was high ($r = .80$, $p < .05$). In this study, grammar knowledge showed a higher correlation with L2 reading than did vocabulary knowledge or word recognition.

Shiotsu (2010) provided a detailed account of a study with Japanese L1 university-level ELLs in Japan. Syntactic knowledge was evaluated using a multiple-choice test containing 35 items selected from past TOEFL tests. Questions consisted of fill-in-the-blank prompts designed to test "knowledge of sentence structures and that of acceptable sequences and forms of words in terms of syntax" (Shiotsu, 2010, p. 61). Test construct validity was ensured by having English-L1 and Japanese-L1 linguists evaluate the appropriateness of test items prior to administration. Grammatical knowledge was a moderate predictor of reading comprehension ($r = .62$, $p < .05$), more so than vocabulary knowledge, for both stronger and weaker readers.

The general picture, as Jeon and Yamashita (2014) described, shows a substantial relationship between L2 grammar knowledge and reading comprehension. However, the validity of the assessments employed to measure grammatical knowledge and reading comprehension has been challenged. For example, Alderson and Kremmel (2013) criticized Shiotsu's (2010) use of expert judges to establish construct validity. They maintained that experts' perception of whether a test item measures vocabulary or grammar knowledge is unreliable, and that the two constructs are too difficult, if not impossible, to distinguish. Concerning the relative importance of grammar and vocabulary, Zhang's (2012) study with graduate EFL students yielded contrary results to Jeon and Yamashita's (2014) findings: structural equation models "suggested vocabulary knowledge was more strongly predictive of learners' reading comprehension abilities than was their grammar knowledge" (p. 569). Zhang surmised that the divergent findings from previous studies resulted from differences in the sensitivity of the instruments used to distinguish between the two constructs, the language proficiency of the participants, or the influence of inconsistent reading comprehension subskills (e.g., locating the answer versus inferencing) measured in different studies. To address these concerns, Jeon and Yamashita (2014) analyzed task type (sentence completion tasks vs. GJT) as a moderator of L2 grammar and reading. They found different degrees of correlations for the two task types: completion task ($r = .91$) and GJT ($r = .57$); however, the effect of task type only "approached significance" (Jeon & Yamashita, p. 193). Given the relatively large number of studies reviewed by Jeon and Yamashita, one can reasonably conclude that L2 grammatical knowledge indeed plays a large role in L2 reading.

Reading involves a complex interplay between various components, as evidenced in the literature by the presence of several significant correlates along with moderator variables on each of those variables. While one component can play a greater role in reading comprehension, the whole would be incomplete in the absence of the other pieces. Nevertheless, the strength of the relationship between L2 grammar knowledge and reading comprehension warrants further attention: more complex designs beyond mere correlations (i.e., experimental designs) should be used to examine the differential impact of remedial instruction based on grammar and vocabulary.

L2 Vocabulary Knowledge and Reading Comprehension

The second highest correlate found in Jeon and Yamashita's (2014) meta-analysis was L2 vocabulary, which has long been regarded as a necessary component of L2 reading comprehension (e.g., Bernhardt, 2005, 2011; Grabe, 2009; Koda, 2005). In the K-12 context, after a comprehensive review of studies examining English vocabulary and reading with school-aged L2 learners, August and Shanahan (2006) concluded that L2 vocabulary knowledge contributes to L2 reading; the consensus from experimental studies appears to be that explicit vocabulary instruction in meaningful contexts, especially those that incorporate an ELL's L1, lead to improvements in L2 reading (Shanahan & Beck, 2006).

More recently, Lesaux, Kieffer, Kelley, and Harris (2014) conducted a quasi-experimental study on the effects of an academic vocabulary intervention in 14 Californian middle schools. Participants included 50 teachers and 2,082 students in grade six, 1,469 of whom were language minority (LM) students (defined as speaking a language other than English predominantly at home). The researchers randomly assigned the 50 teachers and their students to a treatment (n = 971 with 700 LMs) or control group (n = 1,111 with 768 LMs). The treatment group received 20 weeks of daily 45-minute lessons that focused on a short expository text, along with target words selected from the texts and Coxhead's (2000) academic word list (AWL). Teachers were provided with lesson plans to present students with words that aligned with the content of the day's lesson. Vocabulary instruction began with an introduction to the word in context; was followed by exposure to alternative meanings, morphology, and uses in other contexts that would appear in later lessons; and ended with writing practice. Results from immediate post-tests showed significant intervention effects on the learning of the target words. The treatment led to improvements in expository text comprehension, as measured only by the researcher-designed assessments, not by the standardized reading test. The effects of the intervention were also larger for the LM students.

To my knowledge, L2 reading intervention studies in the K-12 context have focused only on elementary school ELLs. Although Lesaux et al. (2014) reported that their study was conducted in "middle school" and with "adolescents," the participants were in grade six, a grade that belongs to elementary school in most urban Canadian school districts. Currently, educators working to improve high school ELLs' English reading abilities have little empirical evidence to help guide their practices. More intervention studies, particularly with high school-aged ELLs, are necessary to identify or verify effective reading instruction and programs.

For adult ELLs, vocabulary knowledge has been shown to predict performance on reading comprehension in high stakes standardized English proficiency exams such as the TOEFL (Qian & Schedl, 2004). Even in the studies reviewed above, which concluded that grammar contributes more than vocabulary to reading, vocabulary still played a significant role (e.g., Shiotsu & Weir, 2007).

Richards' (1976) seminal ideas on the multidimensional nature of vocabulary knowledge arguably sparked interest in it as a complex construct. Anderson and Freebody (1979, 1981) introduced the notion of breadth and depth of vocabulary knowledge: breadth refers to the number of words for which one knows at least a partial meaning, whereas depth is the dimension that reflects one's knowledge of the word in all situations. Later, Nation (1990) distinguished between receptive and productive word knowledge: these two distinctions essentially correspond to the receptive skills of reading and listening, and the productive skills of writing and speaking. In addition, "knowing a word involves form, meaning and use" (Nation, 2001, p. 26). Therefore, according to Nation, to know a word when it is encountered during reading involves a number of aspects: recognition of its written form, recognition of word parts and their relationship to the word's meaning, recollection of the word's meaning in general, recollection of the word's meaning in the context used, knowledge of the word's concept which allows for it to be understood in different contexts, knowledge of related words, recognition of correct/incorrect usage in a sentence, recognition of typical collocations, knowledge of the word's frequency (i.e., common/uncommon) and its connotation (e.g., complimentary/pejorative). In L2 reading research, the focus has been on the contribution of breadth of vocabulary knowledge, but limited evidence suggests that depth also predicts ELLs' reading comprehension scores on standardized assessments (Qian & Schedl, 2004).

Research into L2 vocabulary learning, specifically long-term retention, has revealed a number of influential factors including (but not limited to) frequency (of appearance in a corpus), word length, complexity of inflections and derivations, polysemy, and cognateness (see Willis & Ohashi, 2012). Evidence exists showing cognateness and frequency to significantly affect learning of vocabulary over time (Willis & Ohashi, 2012); however, the majority of vocabulary studies that I have located focused on frequency.

A reciprocal relationship is thought to exist between vocabulary and reading; that is, growth in vocabulary can be gained through reading, which in turn facilitates further reading (Nation, 2001; Stanovich, 2000). Vocabulary knowledge gained from incidental exposure alone, while possible, appears to be an inferior avenue to acquisition for L2 learners in contrast to reading plus explicit instruction in context (e.g., Laufer, 2003; Min, 2008; Sonbul & Schmitt, 2010). When readers encounter a previously unknown word for the first time, they have to guess word meaning from context, which typically results in less than 15% of these new words being learned to some extent (e.g., Kuhn & Stahl, 1998; Nagy, Herman, & Anderson, 1985; Nation, 2001). Schmitt (2010) suggested that on the first exposure to an unknown written word, a "person may only remember the first few letters of the word, or its broad structural outline…it is only possible to gain the single meaning sense which was used in that context" (p. 20). Studies have indicated that multiple encounters, as high as 20, are necessary for learners to acquire the multiple dimensions of word knowledge (e.g., Pellicer-Sánchez & Schmitt, 2010; Pigada & Schmitt, 2006; Waring & Takaki, 2003).

For readers to be able to guess the meaning of new vocabulary from context, 95% of the text being read must comprise already known words (also called text coverage), which translates to one unknown word in 25 (e.g., Liu & Nation, 1985). Where reading comprehension is concerned, L1 reading studies have suggested that breadth of vocabulary knowledge has a stronger relationship with reading comprehension than depth (e.g., Tannenbaum, Torgesen, & Wagner, 2006). In L2 research conducted with 66 adult ELLs, Hu and Nation (2000) found that the density of unknown words in a text affects comprehension. Participants were randomly assigned to a text with 100%, 95%, 90%, and 80% text coverage. All words were within the 2,000 most frequent words in English; unknown words were artificially created by replacing words with nonsense ones. Results showed a linear relationship between text coverage and comprehension test scores. None of the ELLs reading the 80% coverage text showed adequate comprehension, and only a few reading the 90% and 95% texts did. Adequate comprehension was acknowledged as scoring 12 out of 14 on a multiple-choice test, and 70 out of 124 on a written response test. Hu and Nation (2000) posited that 98% coverage is necessary "for most learners to gain adequate comprehension" without assistance (p. 419).

Researchers interested in vocabulary have further examined the relationship between receptive vocabulary size, lexical text coverage, and reading comprehension scores (e.g., Nation, 2006; Schmitt, Jiang, & Grabe, 2011). Text coverage is usually expressed by the percentage of words present in a text, categorized by their relative occurrence frequencies (frequency bands) in the English language. Schmitt (2014) noted that most studies involving receptive vocabulary knowledge measured only the form-meaning connection, and he cautioned that estimates of vocabulary size could vary depending on the assessment method used (e.g., guessing on multiple-choice formats). Nonetheless, breadth of vocabulary is already well-established as a key facilitator of reading, and a limited number of studies have found that breadth and depth correlate with each other and that both constructs are useful in predicting reading comprehension performance (Qian, 2002; Zhang, 2012).

Using corpus analysis, Nation (2006) determined that knowledge of the 2,000 most frequent word families (including the root, inflections, and derivations) covers only about 88% of novels and 83% of newspapers; the 4,000 most common word families plus proper nouns raise the coverage to 95%. He estimated learners would require knowledge of up to at least the 8,000th frequency band in English to reach the 98% coverage necessary to fully comprehend novels or newspapers. Based on the British National Corpus (BNC), Nation calculated that 8,000 word families translate to approximately 34,660 individual words. These findings coincide with those from Hu and Nation (2000), showing that knowledge of 98% of vocabulary in a text is the threshold for reading comprehension during independent reading.

In contrast, Schmitt et al. (2011) demonstrated the absence of a vocabulary threshold that, were ELLs to possess knowledge of a minimum number of words, would dramatically increase reading comprehension. They found a “straightforward linear relationship” (p. 39) between vocabulary knowledge and comprehension in their study of 661 ELLs aged 16 to 33. The researchers presented participants with two expository texts, one on climate change and the other on an unfamiliar scientific study. Both texts exhibited similar Flesch-Kincaid reading grade levels (9.8 and 9.7, respectively), a unidimensional measure of passage difficulty derived from word and sentence length without regard for other factors that affect reading comprehension (e.g., background knowledge). Vocabulary knowledge (form-meaning connection) was selfassessed using a yes/no checklist, which included all of the words from the two texts. However, no corrections for overestimation due to guessing were employed other than deletion of participants who "chose too many nonwords" that were inserted into the test. Reading comprehension was assessed with a multiple-choice test and a graphic organizer completion task. Results showed a linear relationship similar to Hu and Nation's (2000) findings: the higher the degree of vocabulary knowledge, the higher the degree of comprehension. For example, an estimated 90% text coverage was (based on vocabulary size estimations) required for 50-60% comprehension. Schmitt et al. deemed this relationship to be important, as it indicated that "substantial comprehension" is achievable, albeit not easily, when ELLs do not know one in ten words (p. 35). This is a rough estimate, as vocabulary difficulty (in turn, learners' vocabulary knowledge) depends on more than frequency (see Willis & Ohashi, 2012). More importantly, however, the group with 100% vocabulary knowledge on average obtained only 75.3% on the comprehension test. The authors concluded that knowledge of 98% text coverage, or 8,000-9,000 word families, would be necessary for unaided independent reading of novels and newspaper because other factors beyond vocabulary knowledge also contribute to reading comprehension. Therefore, focusing solely on vocabulary as a way to remediate ELLs' reading problems in content classes will likely be insufficient to help them achieve academic parity with their native-speaking peers.

Inseparability of Lexis and Grammar

As observed in the studies reviewed in the previous section, both grammar and vocabulary contribute significantly to reading comprehension. Some researchers (e.g., Alderson & Kremmel, 2013; Römer, 2009; Shiotsu & Weir, 2007) question the separability of the two components. For example, in Eskildsen's (2012) study of L2 negation construction, the adult learners appeared to acquire lexis and grammar together. Alderson and Kremmel (2013) highlighted the difficulty and futility in trying to isolate syntactic knowledge from lexical knowledge on assessments. In fact, L2 vocabulary researchers (e.g., Martinez & Schmitt, 2012) have expanded their attention beyond the word boundary to include formulaic phrases; applied to L2 foreign language reading, this line of research seemed to support the existence of a lexciogrammar continuum, in that knowledge of multiword units (phraseology knowledge), lying between lexis and syntax, is also predictive of L2 reading comprehension (Kremmel, Brunfaut, & Alderson, 2017). Nonetheless, the dichotomy between vocabulary and grammar has served a useful pedagogical function. It has conveniently allowed teachers to compartmentalize and prioritize their instruction; however, as Römer (2009) suggested, it would benefit ELLs at all levels of academic study to be introduced to lexical-grammatical patterns that will help "them become accepted members of the specific community of practice they aim to belong to" (p. 160). I will return to this point in the section on academic language.

Reading Fluency

Researchers' interest in L2 reading fluency (i.e., reading rate) as it relates to reading comprehension has been gaining momentum (e.g., Fujita & Yamashita, 2014; Gorsuch & Taguchi, 2008; Grabe, 2009, 2010). As alluded to earlier in the section on lower-level skills in L2 reading, fluent reading is an important contributor to reading comprehension. The National Reading Panel (2000) concluded that dysfluent reading could hamper students' acquisition of new knowledge. The panel's meta-analysis of 92 articles on L1 reading fluency practice (e.g., repeated reading, assisted reading, oral recitation) and comprehension with K-12 students found an overall moderate effect size of .41, suggesting "a clear impact on the reading ability of non-impaired readers through at least grade 4, as well as on students with various kinds of reading problems throughout high school" (Chapter 3, p. 3). Kuhn and Stahl (2003) also stated that fluency instruction "is generally effective" in improving the fluency of struggling readers, although it may be an effect of extensive reading (p. 3). Again, lower-level processing occurs in working memory, and inefficient processing may overtax this limited capacity system, preventing the necessary higher-level processing from operating (e.g., Koda, 2005). In light of an interactive model of L2 reading, reading comprehension cannot be solely data-driven by lower-level processing, but requires higher-level processing as well. However, for ELLs without sufficient background knowledge in content-area classes who are trying to learn from reading, instruction that promotes the efficient execution of lower-level processing would likely be a benefit.

A definition of reading fluency includes the elements of automatic processing of lower-level skills, rapid and accurate decoding, automatic word recognition, ease in reading text with speed, and appropriate prosody (Grabe, 2009; Kuhn & Stahl, 2003). This definition involves reading at the word-level and text-level. It also implies a complementary relationship between fluency and accuracy; however, this relationship has been under-researched in the L2 reading literature. Unfortunately, due to the limited number of studies on L2 reading fluency and reading comprehension at the passage-level, fluency as an independent variable was absent in the large scale comparison study by Jeon and Yamashita (2014). In their definition of fluency, Kuhn, Schwanenflugel, and Meisinger (2010) stated that reading fluency also facilitates or limits comprehension in oral and silent reading:

> Fluency combines accuracy, automaticity, and oral reading prosody, which, taken together, facilitate the reader's construction of meaning. It is demonstrated during oral reading through ease of word recognition, appropriate pacing, phrasing, and intonation. It

is a factor in both oral and silent reading that can limit or support comprehension. (p. 240)

The inability to read assigned academic material in allotted times during the school day could jeopardize students' academic performance, leading to a widening academic gap between proficient and inefficient readers (e.g., Grabe, 2009; Stanovich, 1986). For example, ELLs may read well enough to comprehend the text, but they may also require more reading time than fluent L1 readers to arrive at comparable levels of comprehension. Unfortunately, time is often limited in school testing situations. In university, where students are typically faced with copious amounts of required course readings, slower reading rates seriously tax students' time, limiting the amount of reading that they can complete. To illustrate, the seven ELLs in Hirano's (2015) investigation of reading in a post-secondary setting "were unanimous in saying that the amount of reading they were expected to do for some of their classes was beyond that which they could handle comfortably" (p. 181). Grabe (2009) also reminded readers that when viewing reading as a learned skill, the end goal should be autonomous control of the reading process, as outlined in Anderson's (2007) ACT-R theory. Practice, therefore, should increase fluency, which would only benefit students.

Of course, the interest in fluency extends beyond students' ability to rapidly read words aloud without comprehension. For example, in Carver's (1997) rauding theory based on evidence from L1 college students, the author proposed that efficient reading needs to account for both speed and accuracy, and that different types of reading are associated with different rates. Rauding, a portemanteau of reading and auding (listening), refers to "language comprehension situations where most of the thoughts being presented in the form of sentences are being comprehended as they are being presented" (Carver, 1977, p. 13). In this theory, reading is divided into five different processes called gears (scanning, skimming, rauding, learning, and memorizing) (Carver, 1990) that operate at different rates. Scanning requires only lexical accessing and therefore operates at a much higher rate (approximately 600 words per minute) than rauding (general comprehension – 300 words per minute). The slower rate results from additional processes that occur during rauding, namely lexical accessing, semantic encoding, and sentence integrating. Reading for learning (e.g., reading expository material such as academic texts) operates at an even slower rate of approximately 200 words per minute. Carver (1990) further argued that L1 readers generally read at a constant rate (i.e., rauding rate of 300 words per minute) unless the text is difficult or they are reading for different purposes. Efficient reading, then, is the fastest silent reading rate at which an individual can read an easy text with an accuracy of 64% or greater (Carver, 2000).

Research in reading fluency training at the word level has yielded promising results (Akamatsu, 2008; Fukkink, Hulstijn, & Sim, 2005; Segalowitz & Segalowitz, 1993). Over the course of seven weeks, Akamatsu (2008) trained 49 Japanese L1 university students in word recognition tasks. Students improved in both speed and accuracy. Contrary to criticisms of learning vocabulary in a decontextualized fashion, Akamatsu argued that word recognition practice with vocabulary items already in the learner's lexicon could improve lexical access speed. Fukkink et al. (2005) provided 41 Dutch L1 eighth grade students, who were intermediate English L2 learners, with computerized single-word recognition training in a classroom-based experiment. After engaging in two 40-minute long training sessions, the students in the treatment group showed improved speed in lexical access, but the control group did not. Although the training led to speedier word recognition, it improved neither reading speed nor reading comprehension. Before dismissing the effects, efficient word reading could contribute to reading comprehension as a moderator variable. For example, rapid and accurate word recognition could also facilitate comprehension indirectly through fluency or word learning (Hamada & Koda, 2008; Shiotsu, 2010; Yamashita, 2013).

Again, only a few studies have investigated the relationship between L2 reading fluency and comprehension. The aforementioned word recognition fluency training by Akamatsu (2008) and Fukkink et al. (2005) showed that although reading rates improved, no direct effect on reading-comprehension was found. Gorsuch and Taguchi (2008) examined repeated reading practice in a quasi-experimental study with 30 Vietnamese-L1 ELLs in college. At the end of an 11-week-long treatment, the experimental group read significantly faster and comprehended previously unread, novel passages better than the control group. These results are consistent with previous studies on repeated reading, also by Taguchi and colleagues (Taguchi & Gorsuch, 2002; Taguchi, Takayasu-Maas, & Gorsuch, 2004).

Taguchi and Gorsuch (2002) divided 18 Japanese-L1 university ELLs into two experimental and control groups (of nine participants each). Over 10 weeks, the experimental group received 28 reading sessions, each of 30 to 40 minutes. During each session, participants reread an identical passage, distinct from the readings in the control group, seven times, whereas the control group read extensively from a selection of educational books designed for language learning. Although comparison of pre- and post-tests showed both groups improved on their reading comprehension performance and silent reading rate, the experimental group outperformed the control group in reading rate. However, the small sample size of nine in each group needs to be considered before making generalizations. Taguchi et al. (2004) repeated the same experimental study design with 20 Japanese-L1 university students. The experimental group (n = 10) participated in 42 sessions of repeated readings over 17 weeks, whereas the control group (n = 10) read extensively from books of their own choice. Again, both groups improved on their reading rate and comprehension performance, but no group differences were found.

In a larger scale correlational study, Fujita and Yamashita (2014) investigated the relationship between L2 reading fluency and comprehension with 125 first-year high school ELLs whose L1 was Japanese. Results showed a weak but significant relationship between reading rate and comprehension ($r = .24$, $p < .01$). A cluster analysis yielded four groups: low comprehension/reading rate, low comprehension/high reading rate, high comprehension/low reading rate, and high comprehension/reading rate. Fujita and Yamashita surmised that a competition for attentional resources occurred between students' decoding and comprehension processes. However, the negative correlation found in the middle two groups (low rate/high comprehension and high rate/low comprehension) could be due to the way in which the students approached the reading. For example, one interpretation based on Carver 's (1997) rauding theory could be that the slow readers in the middle group were reading at rauding rates, which led to better learning; alternately, the faster readers in the middle group were merely scanning or skimming. Similar to findings from L1 reading fluency research (e.g., Kuhn & Stahl, 2003), the L2 studies above suggest that fluency training, or even simply extensive reading, can increase reading efficiency and could potentially benefit reading comprehension. Therefore, reading interventions that require ELLs to read more than they normally do may increase their reading rate.

Assessing Silent Reading

Although reading comprehension can be difficult to monitor (e.g., Toohey, 2007), maze tests are commonly used as a time efficient, reliable, and valid method of monitoring students' reading competence (i.e., fluency and comprehension) with curriculum material in K-12 schools (e.g., Wayman, McMaster, Sáenz, & Watson, 2010; Wayman, Wallace, Wiley, Tichá, & Espin, 2007). When used as curriculum-based measurement (CBM) for the purpose of formative assessment and instruction improvement, CBM-Maze reading passages have standardized formats (e.g., Espin & Foegen, 1996; Fuchs & Fuchs, 1992; Shinn & Shinn, 2002) similar to cloze exercises: after the first sentence, which is left intact, every seventh word is omitted and becomes a test item. In a maze test, students silently read a passage, usually taken from curriculum-based material, and, at every seventh word, they choose the correct word from three possible answers (the correct one plus two distractors). At the end of the allotted reading time, usually one to three minutes, reading rate and comprehension scores are calculated as the number of words correctly chosen. Another scoring procedure involves the subtraction of incorrect/missing responses from the total score to control for the effects of random guessing. However, Pierce, McMaster, and Deno (2010) found high correlations between these two methods of scoring, and suggested that "technically sound scores" (p. 158) can be obtained from simply counting the raw number of correct responses on a maze.

Because students' performance on a maze task depends on their ability to read fluently and comprehend the text (e.g., Fuchs, Fuchs, Hamlett, & Ferguson, 1992), in theory, successful maze completion requires students to read at Carver's (1990) rauding rate rather than at the more superficial levels of scanning or skimming. Johnson, Semmelroth, Allison, and Fritsch (2013) argued that "maze tasks demonstrate significant promise as reliable and valid indicators of student reading performance, and tap several of the essential areas…needed for academic literacy" (p. 220). In adult education, Gellert and Elbro (2012) found a high correlation (r = .84) between maze tests and a traditional, lengthier, question and answer type of reading comprehension test. In K-12 L1 English education, Muijselaar, Kendeou, de Jong, and van den Broek (2017) found statistically significant correlations between the mean scores of three CBM-Maze tests and the popular, commercial Gates-MacGinitie standardized reading comprehension test (grade 9 students, r = .71) and oral reading fluency (grade 9 students, r = .81). Although Muijselaar et al. cautioned that CBM-Maze tests appeared to favour reading fluency and lower-level reading skills, and may thus be less informative for language comprehension than the Gates-MacGinitie test, they also recognized CBM-Maze as a cost and time-efficient, useful measure.

As ELLs' oral language proficiency is a concern when assessing reading fluency orally, CBM-Maze has been used to measure silent reading fluency with high school ELLs (McMaster, Wayman, & Cao, 2006) and in other studies with a mix of ELLs and NSs in other grades (e.g., Kim, Petscher, & Foorman, 2015; Pierce et al., 2010; Wiley & Deno, 2005). Maze test results have been found to correlate positively with results from standardized reading comprehension tests across all grades (McMaster et al., 2006; Pierce et al., 2010; Wiley & Deno, 2005) and with science content tests. In a study with students (but no ELLs) in grades six to eight, Johnson et al. (2013) found moderate correlations (r = .63-.67; $p < .001$) between results on maze passages constructed from science textbooks and those from standardized state science achievement exams. Of course, subject-specific state achievement exams are not strictly reading comprehension exams, but Johnson et al.'s finding indicate potential predictive abilities of CBM-Maze for content-area achievement.

Role of Text in Science Education

Not only are ELLs learning to read in English, but they are "reading to learn", especially at the secondary education level, where the expectation exists that students will understand increasingly complex expository texts (Chall, 1996). In longitudinal studies of L2 text comprehension with secondary students conducted in the Netherlands, Hacquebord (1994) found that ELLs' reading comprehension lagged behind in content areas. Based on the findings, Hacquebord argued for increased focus on the reading comprehension of course textbooks. As previously noted, some ELLs have been observed to avoid language-heavy courses such as biology (Toohey & Derwing, 2008), as textbooks in content area classes such as science can contain substantial amounts of unfamiliar language.

Norris and Phillips (2003) advocated for the development of reading comprehension as part of science pedagogy. They argued that advances in scientific knowledge have a "constitutive relationship" with text (p. 226); that is, reading and writing are crucial constituents of science. When students read scientific texts, not only do they decode or locate information, but they are also in the process of "comprehending, interpreting, analyzing, and critiquing texts" (p. 229). Reading and writing form a fundamental sense of literacy, and the mental activities involved actually resemble doing science, but without physical manipulatives or experimentations. Students lacking the necessary reading skills will be severely disadvantaged in their acquisition of scientific knowledge. Other researchers (e.g., Krajcik & Sutherland, 2010; Pearson, Moje, & Greenleaf, 2010) have echoed Norris and Phillips' sentiment that the reading of scientific text is akin and complementary to the process of scientific inquiry; scientists read prior inquiries into scientific phenomena before embarking on future investigations. Scientific texts serve as vital sources of background information on which new knowledge is built. Hines, Wible, and McCartney (2010) also expressed the concern that "for students unfamiliar with the language or style of science, the deceptively simple act of communication can be a barrier to understanding or becoming involved with the science" (p. 447). Martin (1993) warned that "diluting scientific discourse" essentially reduces the vigor of the scientific concepts behind the writing (p. 222). Similar to Norris and Phillips (2003), Lee, Quinn, and Valdés (2013) called for improvements in reading education that includes comprehension and interpretation, and they placed the onus on all teachers, including those in content-area subjects.

In an effort to help teachers assist ELLs who struggle in science classes, much of the literature for ELL educators has focused on vocabulary knowledge development or modification of texts (e.g., Duguay, Massoud, Tabaku, Himmel, & Sugarman, 2013; Miller, 2009; Munro et al., 2013; Snow, 2010). Understandably, while vocabulary knowledge clearly contributes to text comprehension in a significant manner, memorization of vocabulary alone cannot resolve ELLs' reading problems, nor does it constitute a good science education. Although modified texts (i.e., simplified syntactic complexity) may help lower proficiency ELLs access texts, the goal of reading interventions for higher proficiency ELLs needs to extend beyond mere vocabulary instruction and the use of simplified texts to teaching students how to interact with authentic texts. In the interest of providing an equitable education (e.g., Lee & Fradd, 1998), educators should strive to help ELLs meet the same standard as their native-speaking peers.

Before ELLs can interpret, analyze, or critique scientific texts, skills which Norris and Phillips (2003) argued as essential for science education, ELLs must first comprehend the information encoded in texts. Thus, I return to the problems associated with reading academic texts. Ideally, content-area teachers would teach not only their subject matter but also reading skills to ELLs. However, as stated earlier, content-area teachers may not have the training to effectively teach English to ELLs; many science teachers are unaware of the need to teach the academic language structures in their subject area, and they may lack sufficient metalinguistic awareness to do so (e.g., Cho & McDonnough, 2009; DelliCarpini & Alonso, 2014; Snow, 2010). Also, ESL teachers typically do not have expertise in science, especially at the high school level; furthermore, some schools may not have any trained ESL specialists on staff. As a result, despite ELLs attending mainstream science classes, science teachers usually view their role as teachers strictly of science concepts, while ESL teachers concentrate on language (e.g., Arkoudis, 2003; Huang, 2004), culminating in the neglect of the language-content relationship in the science classroom (Mohan & Slater, 2006). Understandably, teachers' identities and practices are "inextricably linked" (Norton & Toohey, 2011, p. 414). However, the science classroom offers abundant opportunities for ELLs to encounter English that is academic in nature. To support educators in facilitating ELLs' acquisition of English and reading, the literature on the intersection between science and ESL is further explored in the next section.

Academic Language and the Language of Scientific Texts

Students' ability to read and comprehend complex expository text in high school undeniably provides them with a competitive advantage in their future endeavours (e.g., employment, post-secondary studies). The language used in a content-area course, when juxtaposed against common everyday language, can appear different in "formality of tone, complexity of content, and degree of impersonality of stance" (Snow, 2010, p. 450). As shown in Dicerbo et al.'s (2014) literature review of teaching academic language to ELLs, teachers and researchers alike may conceptualize academic language differently. For the present study, I adopted Nagy and Townsend's (2012) broad definition of academic language: "the specialized language, both oral and written, of academic settings that facilitates communication and thinking about disciplinary content" (p. 92). Their definition also recognizes the "cognitive processing of disciplinary concepts and phenomena" involved when interacting with disciplinary language (p.

92). This definition reflects Norris and Phillips' (2003) notion that the mental activities involved when reading scientific texts resembles doing science.

The literature for teachers of mainstream content-area classes often contains general pedagogical recommendations and strategies for educators to help struggling readers meet the linguistic demands of academic texts. For example, Wexler, Reed, Mitchell, Doyle, and Clancy (2015) provided teachers with a broad, four-step instructional routine that they can implement in their content-area courses. The goal is to prevent the circumvention of textbook usage in classes, which deprives struggling readers of much needed reading practice. The sequential four steps consist of explicit background knowledge instruction, explicit vocabulary instruction, explicit main idea instruction, and peer-mediated practice. While each step in Wexler et al.'s routine is evidence-based and useful, the routine itself falls shy of dealing with academic language commonly used in content-area subjects. Also, Wexler et al. hinted at vocabulary as the main linguistic cause of ELLs' observed underperformance on reading comprehension assessments, but the authors failed to discuss or address other linguistic problems faced by ELLs.

In a qualitative case study of a grade seven science classroom, Bruna, Vann, and Escudero (2010) observed that the teacher's focus on vocabulary as the exclusive representation of academic language severely limited the ELLs' abilities to partake in "genuine scientific inquiry and explanation" (p. 30). Classroom transcripts showed that students' language productions were largely reduced to single word responses. From the perspective of second language acquisition pedagogy, single word utterances do not represent the goal of content-based language instruction wherein the use of subject-matter content serves as a vehicle to L2 learning (Pica, 2002). Snow and Uccelli (2009) also found that grade seven students subjected to vocabulary-only instruction did not appear to acquire the linguistic sophistication (in terms of vocabulary and grammar) of academic English, as evidenced by their writing.

Teaching ELLs to read and comprehend academic language necessitates more than vocabulary instruction. Through corpus analysis, Biber (2006) identified a number of salient features present in academic language. Compared with conversational English, it contains comparatively higher proportions of nouns and noun phrases, more uses of nominalization (nouns created from other word classes with or without a change in morphology), increased informational density, and higher degrees of abstractness. In the aforementioned study by Snow and Uccelli (2009), students who experienced an instructional focus on academic vocabulary without grammar failed to produce writing samples with higher lexical density, lower frequency vocabulary, or lengthy noun phrases.

Establishing that vocabulary knowledge is not the sole determinant of academic language proficiency allows for the exploration of additional interventions for ELLs who struggle with reading. From a linguistics perspective, Halliday (1989/2004a) explained that the "difficulty [of scientific writing] lies more with the grammar than the vocabulary" (p. 161). However, it should also be stressed that in no way is the value of vocabulary being undermined; Halliday (1997/2004b) stated:

> 'grammar' here is short for lexicogrammar. The essential point is that lexicogrammar (the syntax, morphology and vocabulary of a language all taken together) is one stratum within the overall organization of language; and underlying all of it is a network of *systems*. When we look at the grammar in this way, systematically, we are able to see it all as a single resource, a unified force for transforming our experience into meaning. (p. 183)

Halliday's sentiment is echoed by language acquisition researchers who challenged the separability of lexis and grammar (e.g., Alderson, & Kremmel, 2013; Eskildsen, 2012; Römer, 2009; Shiotsu & Weir, 2007). Therefore, instruction in reading academic language will likely benefit from the incorporation of the grammar of science. However, ELLs' verbal communicative competence in everyday routines does not necessarily indicate proficiency in reading comprehension and fluency when reading academic texts (e.g., Hulstijn, 2011, 2015).

According to Hulstijn's (2011, 2015) Basic Language Cognition (BLC) Theory, language proficiency can be divided into two dimensions: basic language cognition and higher language cognition (HLC). In short, BLC pertains to highly frequent linguistic items (vocabulary and grammar) over which all adult L1 speakers would be expected to have automatized productive (speaking) and receptive control in common oral communicative situations. HLC is a complement of BLC that extends to the written mode, consisting of less common lexciogrammatical structures. Hulstijn proposed that one's degree of HLC proficiency is a function of their education level. Therefore, even orally proficient high school ELLs would be simultaneously acquiring BLC and HLC. Exposure to and instruction in the lexicogrammar of science is therefore necessary for improving ELLs' reading comprehension of academic texts.

Expository texts, such as those found in science textbooks, pose significant challenges for ELLs due to the use of unfamiliar discipline-specific academic language (e.g., Snow, 2010). Yet, as many (e.g., Martin, 1993; Norris & Phillips, 2003) have argued, access to this specialized language is vital for science education; text is "one of the most important pieces of technology used by scientists" (Martin, 1993, p. 187). From a vocabulary perspective, Coxhead, Stevens, and Tinkle (2010) analyzed the lexical profile of four secondary level science textbooks used between grades nine to twelve in New Zealand. They found that science texts indeed contain specialized vocabulary that is not found on other words lists such as Coxhead's (2000) AWL or West's (1943) General Service List of English Words (GSL). Furthermore, they found that for the grade 10 textbook alone, 11,000 word families of the BNC with proper nouns were required to reach 98% text coverage; for the entire series of high school science textbooks, 98% coverage was reached at 14,000 word families with proper nouns. These figures are much higher than Nation's (2006) estimate of 8,000 word families that would be required to read novels and newspapers, which could partially explain the difficulty of reading science textbooks, as "a learner needs to know at least 3,000 more words than to read a novel in English" (Coxhead et al., 2010, p. 48).

Beyond the traditional vocabulary distinction, Systemic Functional Linguistics (SFL) forms a framework for analyzing scientific texts and finding explanations for texts' utility for scientific writers: the linguistic register of science is not only a pre-described rigid set of rules, but also a resource for communication and creation of meaning. Therefore, SFL is concerned with texts in context rather than discrete sentences (Halliday & Martin, 1993). SFL-based pedagogy has been gaining momentum as an option for improving the academic literacy of ELLs (e.g., Gebhard, 2010).

A Convergence of Theories

An SFL view of language is compatible with the usage-based learning perspective in SLA (see Tyler, 2010). Usage-based learning theorists (e.g., Tomasello, 2003) propose that humans acquire language not through algebraic computations and permutations of words; rather, language emerges through encounters with concrete exemplars during interactions involving language. Thus, during language use, humans' knowledge develops through making connections between linguistic units and the external world (i.e., knowledge develops through meaningful language use). The context in which language use occurs affects a person's linguistic forms (words or grammatical constructions) of choice. Form-meaning connections strengthen, depending upon the frequency of their co-occurrences in the input (Ellis, 2002). In other words, constructions of form-meaning patterns emerge as a result of an individual's interaction with the world. In education, students predominantly interact with academic language through reading. As Mohan and Slater (2006) stated,

> [Systemic functional linguistics] emphasizes the text or discourse as a whole in relation to the context of social practice and recognizes how science lexis and grammar correlate with science texts and contexts. It provides the tools to investigate and critique how language is involved in the construction of meaning and sees language *as a resource for making meaning.* (pp. 304-305)

This view of academic language as an alternative tool for communication and of the importance of making it explicit for ELLs in the science classroom is gaining momentum with K-12 teachers (e.g., Lee et al., 2013).

The value of explicit language learning in L2 education is recognized by usage-based theorists (e.g., Ellis & Wulff, 2015). Implicit learning, or learning from mere exposure, has limits and may be impractically inefficient when learning some constructions in an L2. Just as skill acquisition begins with explicit instruction and practice, Ellis and Wulff argued that explicit knowledge of form-meaning associations aids in learners' ability to implicitly abstract patterns experienced through later language use.

SFL and Scientific Writing

From an SFL perspective, writers use linguistic resources to accomplish three metafunctions of communication: ideational metafunction pertains to the representation of the writer's ideas through text; interpersonal metafunction concerns the writer's stance in relation to the reader; textual metafunction relates to the organization of ideas in a passage (e.g., Halliday & Matthiessen, 2014). Banks (2008) considered two linguistic resources to be of particular interest in the study of scientific writing: grammatical metaphor and thematic structure. With regard to L2 reading, the ability to decipher grammatical metaphor and recognize thematic structure could help ELLs realize ideational meaning and develop a coherent mental representation of a text being read. Indeed, others (e.g., de Oliveira & Schleppegrell, 2015; Fang & Schleppegrell, 2008, 2010; Fang, Schleppegrell, & Cox, 2006; Jones & Lock, 2011) have incorporated these two concepts into their pedagogical writings for content-area teachers. The use of grammatical metaphor (described in detail below), in particular, is prevalent in academic language, typically in the form of noun phrases (e.g., Banks, 2008; Biber, 2006; Biber & Gray, 2011).

The term grammatical metaphor is closely associated with Halliday (1989/2004a), who listed grammatical metaphor as one of the main characteristics of scientific writing. Analogous to lexical metaphor which is "an alternative use of a word…to express a figurative, transferred meaning which it otherwise does not have" (Taverniers, 2006, p. 328), grammatical metaphor is the substitution of one grammatical class (i.e., part of speech) for another (Halliday, 1989/2004a). For example, *he discovered* could be expressed as *his discovery*, where both mean a person making a finding. *He discovered* consists of a pronoun (he) + verb (discovered), but when reworded as *his discovery* results in a determiner (his) + noun (discovery) combination.

Verbs typically represent processes (i.e., actions) and nouns denote people, places, and concrete or abstract objects, but in *his discovery*, the action of discovering is reformulated as a thing encoded by a noun. Thus, grammatical metaphors are realizations of circumstances (e.g., instead of – adverb), processes (e.g., replaces – verb), qualities (e.g., alternative – adjective), or things (e.g., replacement – noun) by a grammatical class other than one that is congruent with their semantic categories (note that the examples provided are metaphorical alternatives of one another) (Halliday & Matthiessen, 2014).

Nominalization

Writers create grammatical metaphors primarily through nominalization (Halliday & Matthiessen, 2014). Nominalization refers to the process of transforming verbs and adjectives, which typically express actions and qualities, into nouns. For example, the verb *orbit* may be nominalized as the noun *orbital* in physics, or the verb *report* as the noun *report*. Note that the word may or may not undergo morphological changes. Nominalization may also occur with sentences. For example, *they investigated how atoms move* may be reworded as *their investigation of atomic movement*, where the nominalized recast is no longer a complete sentence but is instead a noun phrase that can serve as the head of another sentence (e.g., *Their investigation of atomic movement involved the use of expensive equipment.*). In the previous example, nominalization led to morphosyntactic changes. Once nominalized, one noun group can even be embedded within another to create complex noun clusters (e.g., *lung cancer death rates*).

In scientific writing, grammatical metaphor and nominalization of processes (e.g., *move* to *motion*) and qualities (e.g., *long* to *length*) create phenomena that can adopt additional properties (e.g., *perpetual motion*) due to the syntactic properties of nouns (Halliday & Martin, 1993). Noun phrases, also called nominals, contain a head-noun that can be modified and are typically constructed in the following combination: determiner + premodification + head noun + postmodification and complementation (Biber, Johansson, Leech, Conrad, & Finegan, 1999). For example, in *the potential energy of atoms*, the noun *energy* is the head noun. Noun phrases can take on a variety of syntactic roles including the subject, object, and pre-modifier of other nouns

(Biber et al., 1999). Thus, noun groups can be combined to create increasingly long noun phrases. Fang and Schleppegrell (2008) provided the following example:

> In English, a simple noun like *spot* can be expanded into a complex noun group such as *those two tropical rain forest hot spots in South America that have not been completely explored* by adding both pre-modifiers and post-modifiers. (p. 29)

Another hypothetical example is *these six chemical elements in group 18 of the periodic table that do not react.* Figure 1 shows an analysis of the constituents in this noun phrase following Fang and Schleppegrell's (2008) format.

Premodification			**Head**	**Postmodification**	
these	*six*	*chemical*	*elements*	*in group 18 of the periodic table*	*that do not react*
pronoun	number	adjective	noun	prepositional phrase	embedded clause
which one?	how many?	what kind?	what is it?	where is it?	which one?

Figure 1. Noun phrase constituent example including parts of speech and information type

Compressed Writing

Halliday and Martin (1993) argued that grammatical metaphor and nominalizations in scientific writing arose out of scientists' need to reify abstract and complex ethereal concepts as things, and to package their ideas and experimental findings in a concise manner for communication. A historical exploration of academic writing by Biber and Gray (2011) revealed a preference for a nominal style that is compressed, achieved through the use of "nominalizations, attributive adjectives, nouns as nominal premodifiers, prepositional phrases as nominal postmodifiers, and appositive noun phrases" (p. 229). More importantly, Biber and Gray

(2011) noted that these “written grammatical characteristics…have remained rare in spoken conversational discourse” (p. 247). As such, ELLs will need exposure to these constructions through reading and explicit instruction to become proficient readers of academic texts.

Nominals allow for the establishment of a chain of reference in discourse (Biber et al., 1999) – a necessity for a cohesive paragraph on a scientific topic. McNamara, Graesser, McCarthy, and Cai (2014) utilized computer software (Coh-Metrix) to analyze the Touchstone Applied Science Associates (TASA) corpus, a collection of high-school level texts, and found that science texts have higher cohesion (overlapping words and ideas across sentences) compared with texts from language arts or social studies. In contrast, McNamara et al.’s Coh-Metrix data showed science texts ranking higher in syntactic simplicity than language arts or social studies. Naturally, the use of a compressed writing style would lead to shorter sentences that Coh-Metrix assessed as syntactically simpler, analogous to assertions that “academic writing is better described as ‘compressed’ rather than ‘elaborated’” (Biber & Gray, 2011, p. 228).

Corpus analyses of academic and scientific writing conducted over the past three centuries have revealed a higher incidence of nominalization than in other forms of writing (Banks, 2008; Biber & Gray, 2011). The prevalence of nominalization continues to be prominent in present-day academic writing (e.g., Biber & Gray, 2011; Fang & Schleppegrell 2008; Fang et al., 2006; Gee, 2008; Jones & Lock, 2011). Gee (2008) suggested that the presence of nominalization and complex embeddings in science textbooks contributes to increased text difficulty. Graesser, Cai, Louwerse, and Daniel (2006) found that lengthy, ambiguous noun phrases tax readers’ working memory, as in the following example sentence from Gue et al.’s (2004) grade 10 science textbook, which is approved by Alberta Education: *What is the overall* *effect of this difference in attraction for the electrons?* (p. 75). It contains a lengthy noun phrase with numerous levels of embedding (Figure 2).

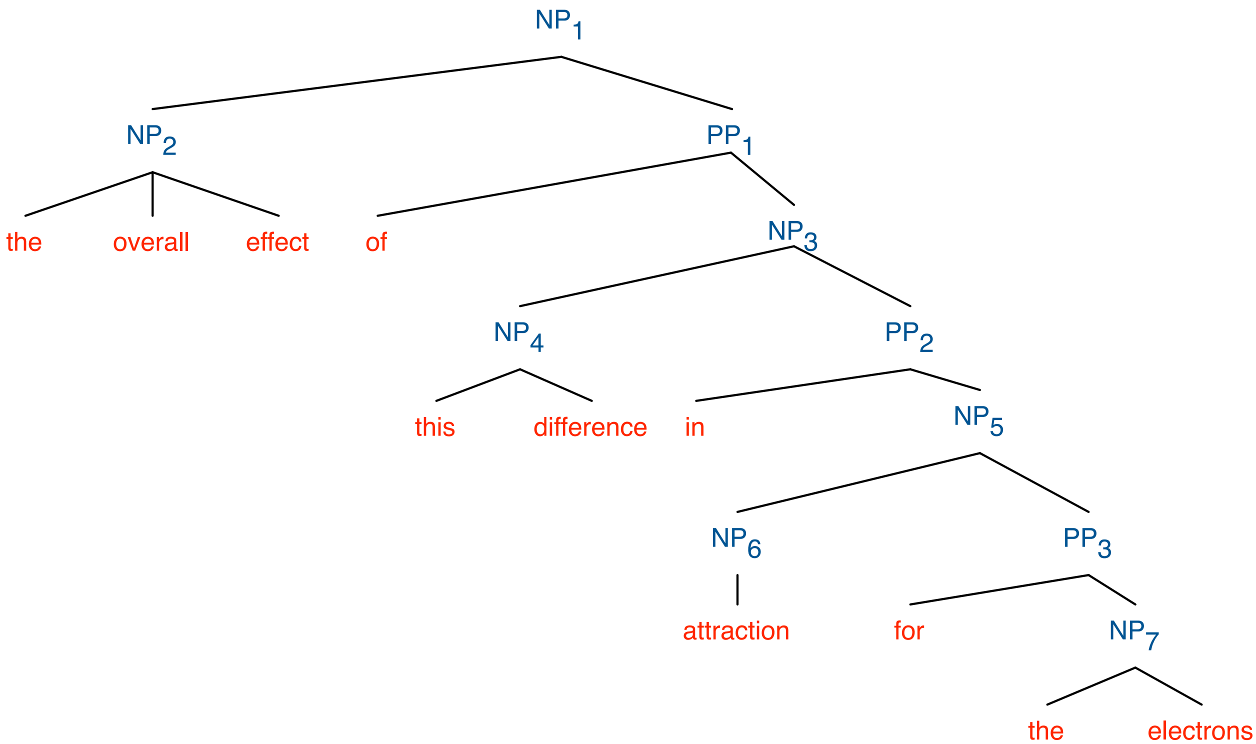


Figure 2. Syntax tree of an example noun phrase containing multiple levels of embedding. NP and PP denote noun phrase and prepositional phrase, respectively.

Text Cohesion, Thematic Structure, and Nominal Groups

In general, texts exhibiting high cohesion have been observed to facilitate reading comprehension in studies with K-12 and postsecondary students (McNamara, Louwerse, McCarthy, & Graesser, 2010). Texts high in cohesion contain textual elements that bridge concepts across sentences for readers; on the other hand, low cohesion texts lack these elements, thus requiring readers to make inferences as to how ideas are linked across sentences. McNamara et al. (2014) coined these linking textual elements “cohesive cues”, which include conjunctions or connectives (e.g., *and, but, so, because*) and overlapping referents or key words. Of concern to the present study are overlapping referents in the form of nouns and noun groups that are common in science texts.

In addition to grammatical metaphor, a feature that contributes to the distinctive style of scientific writing is *thematic structure* (Banks, 2008), which relates to text cohesion. In SFL, thematic structure contains two parts: Theme and Rheme (the capitalization of Theme and Rheme is a convention used in linguistics) (e.g., Halliday & Matthiessen, 2014). Clauses and sentences carry a writer's intended message. The Theme, then, is the writer's chosen "point of departure to guide the [reader] in developing an interpretation of the message" and the Rheme is "the remainder of the message, the part in which the Theme is developed" (Halliday & Matthiessen, 2014, p. 89). For example, in the sentence *All matter is composed of atoms.*, *all matter* forms the Theme while *is composed of atoms* is the Rheme. Within a paragraph, Themes and Rhemes of preceding sentences may reappear in succeeding ones. This cross referencing of information can be described as Theme progression and follows three primary patterns (Eggins, 2004): Theme reiteration, zig-zag, and multiple themes. In Theme reiteration, the first Theme repeatedly serves as the Theme of the succeeding sentences. Contrastively, the other two patterns involve a shift in Theme. Zig-zag describes a linear progression where the Rheme of a sentence becomes the Theme of the next. Finally, in the third pattern, multiple Themes, the Rheme of one sentence contains multiple pieces of information, "each of which is then picked up and made Theme in subsequent [sentences]" (Eggins, 2004, p. 325). To economize space and avoid repetition in writing, nominalization allows Themes and Rhemes to be linguistically repackaged when they reappear in a text passage (e.g., *All matter is composed of atoms. The composition can* *vary.*). Therefore, recognizing overlapping referents that exist in noun groups may potentially help ELLs take advantage of text cohesion to build a coherent mental representation of the text.

Lexical Density and Ambiguity

Although high cohesion texts should generally facilitate comprehension (e.g., McNamara et al., 2010), unfortunately, the use of nominalization also creates the byproducts of lexical density and ambiguity (e.g., Halliday, 1989/2004a). According to Biber et al. (1999), "the lexical density of a text is the proportion of the text made up of lexical word tokens (nouns, lexical verbs, adjectives, and adverbs)" (p. 62). The calculation of lexical density has been described in different ways (e.g., Halliday, 1989; Ure, 1971), but all descriptions include a ratio of words with lexical properties (i.e., nouns, lexical verbs, adjective, and adverbs) to grammatical words (i.e., pronouns, determiners, and finite verbs). Halliday (1989) also argued that the precise boundary of distinction between lexical and grammatical words is unimportant, as long as it is drawn consistently. Ure (1971) used the following formula to calculate lexical density: lexical density = (number lexical items x 100) / total tokens.

The importance of lexical density is seen when comparing the language of conversation with academic language. Corpus analysis using Ure's (1971) method has shown that academic language has a higher lexical density than conversation, as well as a higher frequency of nouns (Biber et al., 1999). Laufer and Nation (1995) suggested that lexical density may be a reflection of syntax; indeed, nominalization can have the effect of increasing the information load per unit of text when employed in academic writing.

Ambiguity in scientific texts may result from the authors' choice of words and syntax. First, the nominalization of processes and qualities results in words that represent abstract concepts; the analysis of the TASA corpus showed that science texts exhibited lower word concreteness than language arts texts (McNamara et al., 2014). Furthermore, the syntactic ambiguity created by nominalization is illustrated in the following example from Halliday (1989/2004a). The sentence *Lung cancer death rates are clearly associated with increased smoking* contains two nominal groups. It is unclear whether *lung cancer death rates* refers to "'how many people die from lung cancer', or 'how quickly people die when they get lung cancer'? Or is it perhaps 'how quickly people's lungs die from cancer'?" (p. 170). The noun phrase *increased smoking* is equally ambiguous, as it could be interpreted as "'people smoke more', or 'more people smoke' or … 'more people smoke more'" (p. 170). Listing all possible interpretations of the two nominal groups (three for the first and three for the second) to calculate the number of possible interpretations for the sentence yields $3^2 = 9$ different permutations (a mathematical concept denoting the combination of things when their order matters). Of course, as Halliday noted, other interpretations, such as "is it the smokers or non-smokers who die?", also need to be considered, resulting in an even greater number of possible interpretations (Halliday stated about 50).

When clauses are nominalized, the original subject and/or object is suppressed, leading to ambiguity for readers (Fang & Schleppegrell, 2008; Halliday, 1989/2004a). For example, when *Lee verified the experiment by Abbott* is nominalized as *verification of Abbott's experiment*, one cannot be certain about (a) who verified Abbott's experiment nor (b) the status of the verification. To further illustrate, in noun (pre-modifier) + noun (head) combinations, because of the absence of function words, 15 different possible types of logical relationships can be expressed (Biber et al., 1999).

According to Biber et al. (1999), approximately 60% of noun phrases in academic writing have either a pre- or post-modifier (25% have pre-modifiers, 20% have post-modifiers, and 12%

have both), whereas in conversations, noun phrases with modifiers are less frequent (15%) in comparison. Fang et al. (2006) and Fang and Schleppegrell (2008), using excerpts from K-12 science books, illustrated the presence of abstraction and lexical density as a result of writers' preference for using nominalization. Biber and Gray (2010) also found a preference for phrasal modifiers in the form of nominal pre-modifiers and nominal post-modifiers in academic writing that students would "encounter in a university education, including textbooks, departmental web pages, and even course syllabi" (p. 9).

Biber and Gray (2011) explored the form-meaning relationship of noun phrases in academic writing, particularly those involving nouns as nominal pre-modifiers and prepositional phrases as nominal post-modifiers. Similar to Halliday's (1994/2004a) argument, noun-noun constructions (as well as multiple noun sequences) produce a wide range of possible meaning relationships. Likewise, noun-modifying prepositional phrases have the potential to express multiple and sometimes abstract meanings. Again using the example sentence *What is the overall effect of this difference in attraction for the electrons?* (Gue et al., 2004, p. 75), one finds three nominal post-modifying prepositional phrases: note that the noun and prepositional phrase *difference in attraction* does not illustrate locative meaning, but rather expresses an abstract idea.

Many researchers (e.g., Biber & Gray, 2010; Fang & Schleppegrell, 2008, 2010; Nagy & Townsend, 2012) have argued that nominalization and lengthy noun phrases in academic language pose problems for ELLs. Biber and Gray (2010) called for more attention to be given to "noun phrase structures with multiple levels of phrasal embedding" because of the possible ambiguity of meaning created from their usage (p. 19). Fang et al. (2006) shared the sentiment that "academic language is only explicit for those who already have tools for unpacking the meanings" (p. 264). Long noun phrases pose a barrier to ELLs' academic performance in content areas (Abedi, 2006; Abedi & Lord, 2001). Abedi and Lord (2001) investigated the effect of linguistic complexity on students' test performance in two studies with 1,174 grade eight students, approximately 31% of whom were categorized as ELLs. The researchers created two 25-item, parallel math assessments; one version contained word problems with simplified language. Linguistic modifications made on the word problems included reduction of nominal length, substitution of low frequency words with high frequency ones, alteration of passive to active voice, deletion or recasting of conditional and relative clauses, simplification of question phrases, and addition of concrete information. The researchers randomly assigned one of the two versions of the tests to the participants. ELLs who wrote the linguistically simplified version of the test performed, on average, 3.7% better than the ELLs who took the non-modified version. Abedi and Lord concluded that language and reading comprehension ability affects achievements in mathematics. Abedi (2006), in reviewing studies performed at the National Center for Research on Evaluation, Standards, and Student Testing, found that "reducing linguistic complexity of test items narrows the performance gap between ELL and non-ELL students in content-based areas such as math and science" (p. 2,284). Specifically, noun phrases were identified as one of the linguistic features that may cause reading difficulties for ELLs.

Analyses of students' and experts' writing have yielded evidence of divergent patterns of noun phrase usage between novice ELLs and NS peers and expert writers (Ai & Lu, 2013; Parkinson & Musgrave, 2014). Ai and Lu (2013) used Lu's (2010) computerized L2 Syntactic Complexity Analyser to assess the complexity of 600 English essays sampled from a NS and a NNS writing corpus. The analysis included 14 measures of syntactic complexity, including complex nominals per clause and per T-unit (a main clause and any dependent clause or attached/embedded non-clausal structures). Results showed a statistically significant difference between the NS and NNS writing samples, with NSs producing, on average, higher numbers of complex nominals per clause (1.222 versus 1.015) and per T-unit (2.089 versus 1.558). In a study that compared the writing of two groups of ELLs – an English for academic purposes (EAP) course and a master's (MA) program, Parkinson and Musgrave (2014) found that the more novice EAP writers exhibited a heavier reliance on adjectives as pre-modifiers in noun phrases than did the MA writers, who utilized a wider array of pre- and post- modifications in their noun phrases. Parkinson and Musgrave interpreted the students' written production as evidence of their acquisition of those forms. While interpreting production (or the lack thereof) as evidence of knowledge has its limitations, it highlights a potential gap in ELLs' knowledge of the language commonly found in academic writing (i.e., lengthy noun phrases).

Academic Language Interventions

The comprehension and production of academic language are requisite skills for academic success, and ELLs should not be sheltered from academic language (e.g., Harper & de Jong, 2004). Exposure to academic language is vital to its acquisition. Miller's (2011) comparison of general adult ESL textbooks with university textbooks showed two notable significant lexico-grammatical differences: the ESL texts had lower coverage of Coxhead's (2000) AWL and less nominal modification. If the process of second language acquisition is indeed usage-based, then ELLs need to be exposed to the same textbooks as their NS peers and provided with scaffolds to comprehend them.

To be clear, vocabulary is a part of academic language, and, as Nagy and Townsend (2012) noted, research in academic vocabulary instruction has yielded positive results that showed gains in students' content and linguistic knowledge, including ELLs' knowledge in science classes (e.g., Brown & Ryoo, 2008; Brown, Ryoo, & Rodriguez, 2010). Furthermore, general academic word knowledge, conceptualized as breadth of knowledge of Coxhead's (2000) AWL, has been observed to explain variance in junior high ELLs' reading comprehension and academic performance across the subjects of math, social studies, and science (Townsend, Filippini, Collins, & Biancarosa, 2012).

In an experiment focusing on a "content-first" approach in science class, Brown and Ryoo (2008) used a website to teach 49 fifth grade minority students the concept of photosynthesis. Students were randomly assigned to use one of two versions of the website. In the first "content-construction" phase of instruction, version 1 of the website explained photosynthesis to the control group using everyday and scientific language (i.e., traditional method). Once scientific language was introduced, it replaced everyday language in subsequent explanations. Version 2 presented the experimental group with the content using everyday language only (i.e., content-first approach). In the second "explicit language phase," the website provided interactive activities on photosynthesis to both groups. However, version 2 introduced scientific language to the experimental group for the first time and provided students with activities that connected everyday language with scientific language. Students in the content-first approach significantly improved in their understanding of photosynthesis, as evidenced by their performance gains on a multiple-choice post-test that used scientific language. Furthermore, the experimental group outperformed the control group in their production of scientific language on the open-ended questions. However, coding criteria to differentiate between the students' use of everyday language (e.g., "Stomata: takes in air from people and animal") and scientific language (e.g., "Stomata: it's a small hole in the leaf where carbon dioxide get in") was unclear, suggesting that the study itself was not strongly grounded within a linguistics framework. While these results are promising, the academic language that was the focus of the explicit language phase appears to have been at the word level, in particular discipline specific technical vocabulary such as "glucose" and "photons" (p. 539). Interestingly, "the software [taught] students that 'a type of sugar that plants produce' is also called 'glucose'" (p. 539). This explanation itself is a noun phrase. Furthermore, the content-first approach may have improved the students' background knowledge prior to the explicit language phase. In other studies, elementary-aged L2 students tended to perform below their native-speaking peers when reading expository texts containing unfamiliar topics, but these group differences disappeared after accounting for the discrepancy in background knowledge (Droop & Verhoeven, 1998; García, 1991). Similar findings regarding the impact of background knowledge were found in studies conducted with older, high school-aged children (Abu-Rabia, 1996; Hacquebord, 1994). Therefore, due to the (a) absence of a robust linguistic framework defining scientific language and (b) the lack of control of the confounding effects of background knowledge, ascertaining the language gains made by participants in Brown and Ryoo's (2008) study becomes problematic.

Other linguistic interventions focusing on academic language have been suggested. For example, Mohan (1986) introduced the Knowledge Framework, which deals with text structure, as a systematic approach to integrate language and content. Early (1990) extended the Knowledge Framework's utility by including key visuals and incorporating it into a language experience approach designed to be successful with beginner ELLs. Also, as already mentioned, vocabulary interventions designed to promote the acquisition of Coxhead's (2000) Academic Word List or discipline-specific technical words have yielded positive results (see Nagy & Townsend, 2012). Specifically related to science, Coxhead and Hirsh (2007) recommended teaching learners the first sublist of the AWL followed by the first sublist of their science-specific word list. However, as highlighted in this literature review, grammatical interventions such as instruction in nominalization also deserve attention as a way to support ELLs' reading of complex texts – an idea shared by teacher-educators of L2 learning (Bunch, Walqui, & Pearson, 2014; Fang & Schleppegrell, 2008; Fillmore, 2014; Palincsar & Schleppegrell, 2014). To my knowledge, the effects of a grammar-based intervention with a focus on noun groups for improving ELLs' reading skills has yet to be tested using an experimental design.

Biber and Gray (2010) argued that "EAP reading courses should provide extensive practice in decoding texts with extensive phrasal modification" (p. 19); however, in general, ESL textbooks appear to provide inadequate exposure to this feature of academic language (Miller, 2011). Research in L2 pedagogy has demonstrated that explicit teaching of grammar within meaningful contexts results in learner gains (e.g., Norris & Ortega, 2000); therefore, a reading intervention that is based on texts written for native speakers and that includes practice in interpreting nominal structures warrants investigation to determine its ability to promote ELLs' comprehension of scientific texts. Also, an intervention that provides practice in processing nominal structures commonly found in scientific writing will potentially lead to increased automaticity (e.g., fluent reading), according to Anderson's (1993, 2007) ACT-R.

Pedagogical suggestions to help ELLs access the grammar of academic language have been recommended (Fang & Schleppegrell, 2008; Fang et al., 2006; Jones & Lock, 2011, Musgrave & Parkinson, 2014; Schall-Leckrone & McQuillan, 2012), but these practices have not been investigated using experimental designs to examine their effects on ELLs' reading comprehension of scientific texts. Fang et al. (2006) provided the strategies to help students analyze nominal group structures, expand nominal structures into everyday language (and vice versa), and synthesize portions of text into noun-phrases to create coherent paragraphs. Fang and Schleppegrell (2008) also recommended similar activities. Jones and Lock (2011) prepared a concrete lesson plan with sample activities and sample worksheets on nominalizations for teachers to use with upper intermediate ESL students. Their activities begin with students noticing and exploring the difference between nominalized and unpacked texts, are followed by a practice phase in which they reconstruct nominalization in context, and then conclude with more practice in unpacking nominalizations. Musgrave and Parkinson (2014) described a task to specifically focus on the noun + noun phrase. The aim of the task is two-fold: the raising of students' consciousness to the noun-noun phrase, and the focusing on form, meaning, and use of the language. Students and their teachers first select noun-noun phrases from a familiar academic text, then the students discuss the meaning of the noun-noun phrases and expand them into clauses/sentences. Next, students engage in a *Jeopardy*-style game in which they have to recode clauses as noun-noun phrases. In the final step of the task, students categorize the noun-noun phrases into predetermined categories (e.g., locative meaning relationship for *kitchen sink*) to improve their inferencing abilities.

In a slightly different line of research, Schall-Leckrone and McQuillan (2012) infused ELL modules into pre-service teacher training in which students were introduced to the academic language of history, including figurative language and nominalization. However, only slightly over half of the cohort were comfortable with deconstructing the language of history (e.g., unpacking nominalizations), possibly due to the students' limited metalinguistic awareness. As the authors pointed out, the effects of teacher training on their instructional practices also needs to be empirically examined. Other researchers have reported positive receptions to SFL-based pedagogy in in-service teacher professional development programs (Achugar, Schleppegrell, & Orteíz, 2007) and a master's degree program in TESL (Gebhard, Chen, Graham, & Gunawan, 2013); this type of training increased teacher's awareness of the linguistic resources employed in the construction of academic texts in content subjects.

Unfortunately, I have been unable to find any experimental studies on the impact of the above activities on ELLs' reading comprehension performance. To be fair, the Sheltered Instruction Observation Protocol (SIOP) model of instruction (Echevarria, Vogt, & Short, 2008) developed for teachers of ELLs in content classes advocates for the identification and instruction of grammatical structures found in science textbooks (Himmel, Short, Richards, & Echevarria, 2009). Implementation of the SIOP as a whole, which also includes components such as intensive lesson planning and strategy building, led to science content knowledge gains for ELLs in middle school (Echevarria, Richards-Tutor, Canges, & Francis, 2011). However, it is not possible to isolate the effects of grammar instruction in Echevarria et al.'s (2011) study. DiCerbo et al. (2014) also noted "a paucity of research in the area of explicit grammar instruction for ELLs" (p. 464), specifically in relation to acquiring academic language.

Methodological Issues and Computer Assisted Language Learning and Research

In Plonsky's (2013) review of quantitative studies in L2 research published between 1990 and 2010, lab-based observational studies were found to far outnumber classroom-based experimental studies. Within the pool of classroom-based experimental studies, individual participants were randomly assigned to experimental conditions in only 23% of them. Although classroom-based studies provide ecological validity, and random participant selection and group assignment constitute two crucial elements of experiments (e.g., Hudson & Llosa, 2015), this design "is not always practical, pedagogically appropriate, or methodologically viable" (Plonsky, 2013, p. 680). To increase the practicability and feasibility of experimental designs in SLA, MacWhinney (2017) suggested that SLA research could benefit from using the online environment and web browsers as a platform for conducting experiments. Ideally, the systems would be openly accessible (without proprietary software) through the use of a coding language such as HTML5 that is understood by any modern web browser that is present on desktop computers or mobile devices. However, MacWhinney noted the financial costs and expertise required to develop such systems as potential limitations. In addition, ensuring that experimental conditions align with pre-planned classroom learning outcomes remains challenging for classroom-based experiments.

Another methodological issue present in SLA research is in the area of statistical analysis. Increasing attention has been paid to improving the quality of statistical analyses in quantitative SLA research (e.g., Larson-Hall, 2016; Larson-Hall & Plonsky, 2015; Norris, 2015; Norris, Plonsky, Ross, & Schoonen, 2015; Norouzian & Plonsky, 2018; Plonsky 2014) regarding three prevalent themes: inadequate information reporting, reliance on null hypothesis significance testing (NHST), and interpretation of effect sizes. Surprisingly, one third of SLA studies published prior to 2010 omitted informative basic descriptive statistics (i.e., means and standard deviations) (Plonsky, 2013, 2014). Plonsky also identified a publication bias towards statistically significant (i.e., $p < .05$) results and omission of non-statistically significant data either in publication volumes or within a manuscript, although the information could help advance our field through inclusion in future meta-analyses. Starting in 2010, however, changes were made and new guidelines were established for reporting quantitative research results in SLA (see Larson-Hall & Plonsky, 2015; Norris et al., 2015). Recommendations include the reporting of all informative statistics, including effect sizes and confidence intervals (conforming to the publication guidelines established by the American Psychological Association). These recommendations also address a call for “new statistics” (e.g., Cumming, 2012, 2014) that encourage the interpretation of data beyond the statistically significant/non-significant dichotomy under traditional NHST. As Norris (2015) cautioned, "there is nothing mystically important about $p < .05$" (p. 122), as this widely adopted cutoff alpha value is arbitrary and statistical significance is affected by factors such as sample size, statistical power, and effect size. Furthermore, statistical significance is only loosely related to practical significance (e.g., Larson-Hall, 2016). Accordingly, with a shifting focus towards reporting effect sizes and confidence intervals in SLA research, Plonsky and Oswald (2014) revisited the utility of dogmatic interpretations of effect sizes based on Cohen's (1988) benchmarks. Plonsky and Oswald found the benchmarks to "underestimate the effects obtained in L2 research" (p. 878) and provided "new field-specific benchmarks of small ($d = .40$), medium ($d = .70$), and large ($d = 1.00$)" (p. 889) when interpreting intergroup differences. However, Plonsky and Oswald also cautioned against rigid interpretations of *d* values using these guidelines. For example, out of necessity, L2 researchers often study comparison groups (alternate interventions) rather than a true control (no intervention); in comparisons of alternative interventions, smaller effect sizes are to be expected than when an intervention is compared with a null condition. Finally, Norouzian and Plonsky (2018) reminded researchers to carefully label measures of effect sizes (i.e., eta-squared versus partial eta-squared), which previously have led to misinterpretations.

As MacWhinney (2017) indicated, affordances provided by technology could facilitate classroom-based experimental studies. However, as noted in the next section, in which I describe research in the area of computer assisted language learning, previous studies have been plagued by inadequate reporting practices (e.g., Plonsky, 2015).

Computer Assisted Language Learning

Schools in Canada have increasingly embraced technology as a vital component of students' learning experiences. For example, Alberta Education's (2013b) Learning and Technology Policy Framework reflects the belief that technology will "enable [Alberta Education's] new vision of education" (p. 4). The ministry argued that digital technology allows for student-centered learning by providing students with "equitable access to learning opportunities" that empowers students to "learn at their own pace, anywhere, at any time" (p. 15). Implied in the policy is the expectation for schools and teachers to utilize digital educational technology. The use of computers to deliver the reading comprehension interventions developed for this study is appropriate in light of Alberta Education's emphasis on the use of technology to improve learning outcomes.

According to Taylor (1980), computers in education can take the role of a tool, tutor, or tutee. As a tool, the computer is used for specific purposes (e.g., word processing). The task of tutoring can also be delegated to the computer where the machine not only presents content matter, but also collects, evaluates, and provides feedback to students' responses. The computer becomes the tutee when the user programs it to function as tool or tutor. In language teaching, the computer has predominately been used as a tool and a tutor to varying degrees of success in achieving learning outcomes. In the present study, the computer plays the role of a tutor.

Warschauer and Healey (1998) identified three historical stages of computer assisted language learning (CALL) that correspond to the dominant SLA theories of the time. In the 1960s, behaviouristic practices (e.g., commonly called 'drill-and-kill') were the norm, which led to the computer taking on the role of a mechanical tutor. The popularity of communicative approaches in the 1970s and early 1980s, and the associated focus on using language, contributed to students' use of the computer, either individually or with others, to discover language. Finally, socio-cognitive SLA theories gained prominence in the 1990s, ushering in the integration of CALL and the use of language skills in authentic contexts. With four decades of history, there exists a wealth of studies examining the efficacy of CALL for language learning.

Macaro, Handley, and Walter (2012) performed a systematic review of CALL studies in the K-12 L2 context that were published between 1991 and 2010. Of interest to the present study are the insights that Macaro et al. revealed regarding the linguistic and other benefits that CALL yields. Because of advancements in online technology, the authors restricted the scope of their analysis to research conducted since 2000.The global popularity of studying English as an L2 in K-12 led them to focus exclusively on English CALL. In total, Macaro et al. analyzed 47 studies in depth. While the topics of inquiry for the 47 studies collectively encompassed a number of language skills (reading, writing, listening, speaking, vocabulary, grammar, and pronunciation), the researchers could not find "clear or sufficient evidence of [technology's] effectiveness" (p. 24) in promoting gains in any of those skills over traditional non-CALL methods. Macaro et el. criticized the body of CALL research for lacking rigour in design and methodology, inadequacy in reporting details, and failure to connect technology use with SLA theory. Despite the absence of evidence showing a direct impact on language learning, the authors concluded that linguistic gains may result as a secondary benefit from students' positive attitudes towards CALL; moreover, Macaro et al. argued that the use of CALL "may produce different learning behaviours which are beyond linguistic outcomes but are no less educationally valid" (p. 25).

Somewhat contrary to Macaro et al.'s (2012) conclusion, Grgurović, Chapelle, and Shelley (2013) found that, overall, CALL was "at least as effective" (p. 165) as non-CALL instruction. In their meta-analysis of CALL research spanning 1984-2006, Grgurović et al. identified 37 studies that met the following inclusion criteria: the studies assessed language performance (reading, writing, vocabulary, grammar, pronunciation, communication, or integrated skills), employed at least a post-test in an experimental or quasi-experimental design, and contained sufficient statistical information to calculate effect sizes. This pool of studies included participants in education settings ranging from primary to post-secondary, with a spectrum of proficiency levels from beginner to advanced. Grgurović et al. found that, overall, CALL environments showed small but statistically significant effect sizes; however, they were unable to determine "the effects of instructional conditions, characteristics of participants, and conditions of the research design…because of the small number of effect sizes…used as the basis for conclusion" (p. 191). Similar to Macaro et al. (2012), Grgurović et al. found a "lack of methodological rigor in many studies" (p. 192) leading to their exclusion from the meta-analysis. Also, only a small proportion of studies were conducted in the K-12 context, prompting Grgurović et al. to call for additional research examining the effectiveness of technology in language teaching with primary and secondary students.

The results of the large-scale reviews in CALL provide adequate justification to conduct an intervention study using CALL as a tutor to deliver online modules. The development of online modules grounded in SLA theory and aligned to the classroom (i.e., relevant to the curriculum) imparted pedagogical strength, relevance, and value to the study. A CALL intervention study realizes the additional beneficial pragmatic by-product of meeting government and school administrations' implied expectations of technology use. For example, practice with on-screen reading may provide students an additional edge in exam preparation, considering Alberta Education's adoption of a digital format for its Provincial Achievement Tests, Diploma exams, and other assessments (see Quest A+ and Student Learning Assessment at https://questaplus.alberta.ca). From a research perspective, a CALL-based experiment allows for tighter control of sources of variances (e.g., teacher effects), and students can be randomly assigned to different intervention conditions. Therefore, in my dissertation study, the interventions were delivered online to increase scientific rigour and to facilitate a classroom-based experimental design. Furthermore, detailed attention was paid to statistical analyses and reporting methods that met the inclusion criteria for potential future meta-analyses and allowed me and potentially other researchers to better evaluate the effectiveness of the interventions being studied, thus answering the call for better SLA and CALL research.

Review Summary

ELLs can benefit from additional linguistic support (e.g., reading instruction) in content-area classes to achieve parity in academic performance with their native-speaking peers. Linguistic barriers (e.g., independent L2 reading skills) may hinder ELLs' progress, especially as they prevent access to written information found in academic texts. Research in reading and SLA has shown that knowledge of English grammar, vocabulary, and potentially fluency (accuracy and/or efficiency), contributes to success in reading comprehension. Examinations of written academic language have revealed key differences from conversational language; the language of science, in particular, favours the use of noun phrases. Furthermore, scientific writing follows a cohesive Thematic structure, with noun phrases serving as co-referents across sentences. Noun phrases have also been demonstrated to be problematic for ELLs. Theoretically, explicit instruction in unpacking the meaning of nominal structures should benefit ELLs. Unfortunately, not all teachers have the linguistic knowledge to effectively teach ELLs the language of their subject discipline beyond vocabulary. Although pedagogical suggestions concerning nominalization exist, their effectiveness is in need of empirical testing. CALL offers a potential, practical delivery method for ELL interventions. Research examining decades of CALL suggests that at the very least, the use of technology is associated with positive learner attitudes.

Study Overview

In my dissertation study, I attempted to address the gap in research examining reading comprehension interventions designed for high-school aged ELLs. I investigated the impact of vocabulary and grammar instruction on ELLs' performance in reading comprehension, reading fluency, and noun phrase comprehension, as well as the students' experiences with the interventions. To begin, I analyzed two Alberta Education approved grade 10 Science textbooks to understand the type of text (vocabulary, grammar, cohesion) to which ELLs are exposed in high school science classes. I created three different Internet-based modules/reading interventions using 30 passages from the grade 10 science textbooks. Module 1 required students to read each passage and then complete traditional reading comprehension activities (i.e., answering comprehension questions and learning Science 10 technical, discipline-specific vocabulary). Module 2 required students to read each passage and then complete AWL vocabulary activities, and Module 3 required students to read each passage and then complete scientific grammar activities. Each module consisted of 30 sessions. All three modules used the same science passages in each session. A pre-test, intervention, post-test design and multivariate statistical analysis were used to examine the relative effects of the three types of online modules on ELLs' reading comprehension, silent reading fluency, and ability to interpret noun phrases. Field notes and a follow-up survey provided additional information that I thematically analyzed (Braun & Clarke, 2006) to understand the students' experiences using the modules, namely, the perceived benefits and challenges of module use. These rich multiple data sources provided deeper insights regarding the effects of my interventions on the ELLs' knowledge, skills, and attitudes.

Chapter Three – Method

Introduction

As presented in the previous chapter, pedagogical practices in the area of L2 reading of content area texts can benefit from classroom-based experimental research. A reading intervention study was designed to address the aforementioned research questions. In the following sections in this chapter, I describe the research design, the data collection instruments and interventions, the participants and the setting, the procedures involved in the textbook analysis and the experiment, and the data analysis methods.

Overall Research Design

A mixed methods study of a multiphase design (Creswell & Plano Clark, 2011) formed the framework of this dissertation study. The interventions developed for the study were informed by the quantitative results of a textbook analysis and qualitative feedback from in-service teachers and teacher educators. Following the collection of quantitative data from an intervention experiment, qualitative data was gathered to augment and enhance the interpretation and discussion of the quantitative results.

The study began with a corpus analysis of two Ministry of Education approved grade 10 science textbooks, followed by the creation of the three reading interventions/modules and two assessment instruments that were based on the content of the textbooks. Science 10 course material was chosen because the course is a prerequisite for all other higher level (i.e., post-secondary stream) science courses. A randomized pre-post experimental study was conducted in two phases: a Pilot Phase and a main study. Participants were randomly assigned to one of three groups, each group was assigned one of the modules, and then each group of students used a website that I had created to practice reading science passages extracted from one of the government-approved Science 10 textbooks. Following each reading, the three groups completed different follow-up activities, corresponding to their assigned intervention/module.

I decided not to include a true control group that would have received no interaction with the science texts and activities for a number of reasons. Considering that the act of silent reading alone would likely have an effect on the dependent variables, assignment to a control group where no reading is done would likely put these students at a disadvantage compared with peers in an intervention group. On ethical grounds, I could not justify purposefully excluding students from interventions that might benefit them academically. Furthermore, as I was responsible for the classroom activities during each experiment timeslot, I was required to occupy all of the students with meaningful work. Finally, the focus of my study was the comparative effects of the different instructional approaches. The probability of an ELL receiving no instruction at all in school is low; therefore, the traditional comprehension question and answer module serving as a comparison group allows for more meaningful analyses.

Another ethical consideration resulting from an experimental study involving educational interventions is that those who use a more effective intervention may have an unfair advantage in school (e.g., higher grades, greater success in scholarship competitions). Differential group assignment may also lead to resentful demoralization (Creswell, 2008). Therefore, to address this potential problem, all three modules were made accessible to all the participants after the post-test administration.

The pilot study, essentially a small scale “dress rehearsal” (Dörnyei, 2007, p. 75), allowed me to ensure the stability of the interventions on my website when simultaneously accessed by multiple users. The pilot study also provided me with the opportunity to observe the participants while they were completing the research instruments (background questionnaire, pre- and post-tests, interventions) and to fine-tune them if necessary. Participants provided feedback on the usefulness of the online modules through a post-intervention questionnaire. Following the pilot study, I recruited additional ELLs from the same population of ELLs in one local school district, and I conducted the same experiment and administered the identical post-intervention questionnaire as the pilot. Because the website was stable and the instruments remained identical from the pilot to the main study, the data from both phases were combined to maximize the sample size for the statistical analyses.

Textbook Analysis – Tools and Procedures

To gain a more nuanced understanding beyond the grade level text readability (i.e., Flesch-Kincaid Grade Level) of the two Alberta Education approved Science 10 textbooks (*Addison Wesley Science 10* [Sandner et al., 2004] and *ScienceFocus 10* [Gue et al., 2004]), each chapter was digitalized and individually analyzed using a variety of online text analysis tools, described in detail below. Picture captions and sidebars were excluded from the corpus, as they are tangential to the continuous reading passages and their inclusion would render "misleading" results (Dowell, Graesser, & Cai, 2016). Note that both texts contained 12 chapters and presented identical topics in the same sequence (likely a result of both textbooks having been tailor-made for the Alberta Science 10 Program of Studies).

Three freely available automated online text analysis tools provided a number of different indices of text complexity measured at the word, sentence, and passage levels. At the word level, the Compleat Lexical Tutor (Lextutor) by Cobb (n.d.) includes a function called VocabProfile that analyzes words in an inputted text and returns the amount of text coverage by vocabulary frequency bands (commonly called the lexical frequency profile or LFP). For the present analysis, text coverage was measured against the British National Corpus and the Corpus of

Contemporary American English (BNC-COCA). An analysis of coverage by Coxhead's (2000) Academic Word List (AWL) was performed as well. Lextutor was accessed with a web browser (https://www.lextutor.ca/vp/comp/); each chapter was individually copied and pasted into the browser window.

A program called TAASSC: Tool for the automatic analysis of syntactic sophistication and complexity (Kyle, 2016) was used to analyze syntactic complexity. TAASSC allows for batch processing of text and includes indices of clausal and phrasal complexity. Two measures of phrasal sophistication of particular relevance to my study are complex nominals per clause and complex nominals per T-unit (see Lu's [2010] syntactic complexity analyzer in TAASSC). Complex nominals are defined as nouns with modifiers, nominal clauses, and infinitives and gerunds in the subject position. The TAASSC was downloaded free of charge from the developer's website (http://www.kristopherkyle.com/taassc.html) and installed on my personal computer.

Another automatic text analyzer I used was Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014). Six indices of passage-level text readability calculated by Coh-Metrix were examined: narrativity, syntactic simplicity, word concreteness, referential cohesion, deep cohesion, and Flesch-Kincaid grade level. The system provides percentile scores on the first five indices referenced against the Touchstone Applied Science Association (TASA) corpus. These five indices are considered together as a multidimensional measure of text readability. Narrativity relates to the degree the language resembles oral language. The syntactic complexity index in Coh-Metrix includes considerations for numbers of words and clauses per sentence and left-embeddedness before the main verb; lower scores on these sub-indices lead to higher syntactic simplicity scores. The system also analyzes sentence construction across sentences within a text with the assumption that passages with higher syntactic uniformity are easier to read. Word concreteness pertains to the ease with which mental images are evoked by words in the text. Referential cohesion is a measurement of the amount of overlap in words and ideas across sentences throughout a passage. Deep cohesion indicates the presence of connectives that help readers make links between logical relationships. In contrast to the Coh-Metrix readability measure, the commonly used Flesch-Kincaid grade level readability index is only a single dimension approach and a less accurate measure of readability, as the formula is based only on the lengths of words and sentences in a passage. Coh-Metrix was used through its web tool interface (http://www.cohmetrix.com) by copying and pasting the individual texts into the browser window.

Pilot Phase – Participants and Setting

In Alberta, the government provides school jurisdictions with additional funding based on the number of coded Canadian-born and foreign-born ELLs enrolled. The maximum number of years of ELL funding for each student is 5 years (Alberta Education, 2016). For the present study, participants consisted solely of ELLs with an assigned code of 301 (foreign-born), who are predominantly children of recent immigrants or refugees.

Participants ($N = 34$) in the pilot study were enrolled in a high school that caters to the academic upgrading needs of students and has specific ESL classes for coded ELLs. The students' ages ranged from 18 to 20 years ($M = 18.77$, $SD = .62$). Students self-identified as either male (44.1%) or female (55.9%). These students received instruction in three separate cohorts, each taught by a different teacher (three in total) in sheltered (ELL-only) ESL classes, with the goal of bolstering their English language proficiency to meet the academic demands of the pre-requisite courses for high-school graduation. In addition to English language and academic content instruction, the school provided the students with academic skills and strategies support instruction (e.g., time management, study skills) and post-high school transition and employment support in the form of a work placement program.

All of the participants reported a place of birth outside of Canada (supporting ELL code 301 status) including China, Eritrea, Hong Kong, Mexico, Nepal, Philippines, Portugal, and Uganda. The majority (73.5%) stated Filipino/Tagalog as their L1, followed by Ilocano (8.8%). Other L1s represented were Chinese (2.9%), Lusoga (2.9%), Nepali (2.9%), Portuguese (2.9%), Spanish (2.9%), and Tigrinya (2.9%). The amount of education in their first language varied widely from 0 to 12 years ($M = 10.68$, $SD = 3.15$). The participants' length of ESL instruction received in Canada ranged from 0 to 15 months ($M = 6.47$, $SD = 4.30$).

Pilot Phase – Instruments

Background Questionnaire (Appendix A). Participants reported their personal background information on a researcher-administered, paper-based questionnaire. The questionnaire contained six questions that asked students to provide their age, gender, country of origin, first language, educational level in their L1, and amount of ESL instruction in Canada.

Post-intervention Questionnaire (Appendix B). Participants' perceptions of the intervention/website were gathered through a 12-item, paper-based questionnaire that included Likert-type and open-ended responses. Topics explored included participants' interest and enjoyment in using the website, perceived difficulty of the readings used in the intervention, evaluation of the utility of the interventions in improving their reading comprehension, and suggestions for improving the intervention/website.

Knowledge of Nominal Groups Test (Appendix C). To the best of my knowledge, a widely used/published test that measures an ELL's ability to comprehend written complex noun phrases does not exist. Consequently, I developed a test that focuses on form-meaning connections in an attempt to assess the participants' reading comprehension of complex noun phrases. The test contained 10 items; each question included a complex sentence with a complex noun phrase underlined, followed by three possible interpretations. Students were asked to judge each interpretation as possessing a similar or different meaning to the underlined noun phrase. Each correct judgement was worth one mark, for a total possible score of 30. The 10 sentences in the test are sampled from one of the two Alberta Education approved textbooks (Sandner et al., 2004, *Addison Wesley Science 10*); the source textbook was chosen to avoid exposure to test-item sentences during the intervention (which used passages from the other textbook, *ScienceFocus10*). Both texts have similar readability indices.

Due to the novelty of the test, face and content validity was established through a review by an applied linguistics expert and two in-service teachers: an ESL specialist and a science teacher. No alternate version was developed; therefore, the same test was used as the pre- and post-test. Internal consistency as determined by Cronbach's alpha was .66. Because the instrument was a short test with only 30 items, the reliability of the measure was considered moderate but acceptable for the exploratory purposes of the study. This decision is consistent with interpretations of Cronbach's alpha by science education researchers when developing research instruments for which alpha values above .64 are described as adequate (Taber, 2017).

Silent Reading Comprehension and Fluency Test (Appendix D). Curriculum-Based Measures (CBM) maze tests were developed to concurrently assess participants' silent reading fluency and comprehension of science-related texts. The silent reading fluency tests were constructed based on the standardization guidelines for maze passages (Espin & Foegen, 1996; Fuchs & Fuchs, 1992; Shinn & Shinn, 2002). Beginning with the second sentence, every seventh word (or the next word if the seventh word is an article or proper noun) is replaced by a multiple-choice item containing the correct word accompanied by two distractors. The distractors have the following characteristics: they do not fit the sentence semantically; one distractor is of the same part of speech as the correct answer while the other is not, but is present elsewhere in the passage; neither distractor rhymes with the correct response; and they are visually dissimilar to the correct choice. Following scoring procedures employed in previous research with maze passages (Johnson et al., 2013), the number of correct responses was counted and the raw score served as the silent reading comprehension and fluency metric. Shin, Deno, and Espin (2000) reported CBM maze test-retest mean reliability coefficient to be .81 when the gap between testing is one to three months.

I created three maze passages using text from the *Addison Wesley Science 10* textbook (Sandner et al., 2004). To ensure comparability, Coh-Metrix, Coh-Metrix Text Easability Assessor, and Lextutor's Vocabprofile were used to analyze various passages. The first two, Maze Forms A and B, served as alternate forms, with Form A as the pre- and Form B as the post-test. The alternate form was initially developed to test whether the effects of the intervention could be transferred to a similar text and students' comprehension of the text. Forms A and B share similar technical specifications on the following indexes, respectively: Flesch-Kincaid grade level (10.1 and 10.6), Coh-Metrix L2 readability (15.2 and 15.7), noun phrase density (442.7 and 425.9), word count (398 and 378), AWL coverage by word families (13.8% and 13.1%), lexical density (.58 and .62), and 95% vocabulary coverage reached by the sixth-thousand frequency band. Both texts were extracted from the same unit of study to improve alternate form equivalence. Maze Forms A and B contains 47 and 51 test items, respectively. Internal consistency of both forms calculated by Cronbach's alpha, respectively, are .87 and .88.

After futile attempts to locate two passages that also shared similar Coh-Metrix text ease and readability scores (syntactic complexity, word concreteness, referential cohesion), which McNamara et al. (2014) argued would yield even more comparable texts than the use of readability indexes alone, I decided to include a third maze passage (Form C) also from the same unit of study as a pre- and post-test. Although repeated maze reading may lead to practice effects (Johnson et al., 2013; Tolar et al., 2012), such effects are unlikely, due to the complexity of the maze format. It would be difficult to remember the items. This third maze was expected to reveal potential group differences while controlling for differences in content and readability indices that were difficult to control across maze passages. The technical specifications for Maze Form C are as follows: Flesch-Kincaid grade level of 9.5, Coh-Metrix L2 readability of 25.2, noun phrase density of 410.6, word count of 414, AWL coverage by word families at 16.5%, lexical density of .53, 95% vocabulary coverage reached by the third thousand frequency band, 56 test items, high in referential cohesion and syntactic simplicity, and low in word concreteness. Internal consistency for Maze Form C as determined by Cronbach's alpha is 0.89.

Global Reading Comprehension Test. The commercially available paper-based IDEA Proficiency Test (IPT3) Reading Test, English version (Ballard & Tighe, 2008a), provided an additional measure of reading comprehension. Targeted for ELLs in grades seven to twelve, the IPT3 Reading test measures overall English reading proficiency, with subsections titled "vocabulary, vocabulary in context, reading for understanding, reading for life skills, language usage, and writing conventions." The entire test consists of multiple-choice items with a maximum score of 55. Two alternate IPT3 forms, 3C and 3D, exist and were used as pre- and post-tests, respectively. Allotted time for the test is 70 minutes. The reliability (Cronbach's alpha) of both forms and their subsections is reported to range between .77 and .95 (Ballard & Tighe, 2008b).

Interventions (Appendix E). The reading interventions were delivered online through a researcher-created website (http://206.12.25.125). Self-contained computerized modules controlled for teacher effects that may otherwise have resulted from teacher-delivered instruction. All of the authoring tools for building the intervention are freely available: *WordPress* (https://en-ca.wordpress.org) was used to create the main website; *Namaste! LMS* (http://namaste-lms.org), a WordPress plugin, expanded the site's functionality to incorporate learning management capabilities; and *H5P* (https://h5p.org) uses html5 technology for the creation of interactive activities that run on any computing device with a modern web browser, whether on a traditional desktop computer or a mobile smartphone. The intervention is accessible to students without the need to purchase licenses for specialized software. With minimal training, teachers should be able to use and/or tailor the activities to their own content-area subject matter. The intervention was hosted on a university computing server during the study but has since been transferred and preserved on the researcher's personal website. Participants accessed their assigned module on a school computer; to reduce the possibility of treatment diffusion between groups during the intervention, researcher-generated usernames and passwords were provided and required. Once the study was completed, the three modules were made accessible to all students.

All three online modules contained 30 lessons. Each lesson required the participants, regardless of intervention group, to read the same passage taken from *ScienceFocus 10* (Gue et al., 2004), one of the two Alberta Education approved grade 10 science textbooks. The first lesson included a brief, two-minute video explaining the focus of the module. Key technical science vocabulary (as identified by the textbook authors) in the passages were e-glossed with their definitions from the textbook glossary. Follow-up activities varied, depending on the intervention condition (traditional question and answer; AWL vocabulary; scientific grammar). Each of the 30 lessons was designed to be completed in 30 minutes, for a total of 15 hours of intervention.

The traditional question-and-answer condition mimicked a typical science class reading activity. A set of multiple-choice reading comprehension questions accompanied each reading passage. Also included was a drag-and-drop, definition matching activity with the technical science vocabulary appearing in the passage.

The AWL vocabulary condition focused on vocabulary in the passage that is included in Coxhead's (2000) AWL. Similar to the traditional question-and-answer group, participants read the passage, then matched the AWL vocabulary with their dictionary definitions (Longman Dictionary of Contemporary English Online, https://www.ldoceonline.com). In total, 222 unique AWL families were present throughout the 30 lessons; only 8 AWL families present on the maze assessments (Maze forms A, B, C) were not covered in the intervention.

The scientific grammar condition was developed according to the pedagogical suggestions identified in the literature review (see Fang & Schleppegrell, 2008; Jones & Lock, 2011). Complex noun phrases in the reading passages were identified with the help of the Stanford Parser (http://nlp.stanford.edu/software/lex-parser.html), a computer program that automatically generates the grammatical structures of inputted sentences. The activities in this module included combining separate clauses into one by utilizing nominalized verbs, matching complex noun phrases with plain English equivalents, identifying head nouns in noun phrases, and deconstructing/unpacking complex noun phrases to find relevant information.

A former ESL consultant at a local school district provided feedback on the first iteration of the interventions. The traditional and vocabulary conditions were well received, but she advised against the use of metalanguage such as “head noun” and “pre/post-modifiers” in the grammar condition. The consultant feared that teachers might judge the activities as too difficult (because of the metalanguage) which would result in denial of permission to conduct the experiment in the classroom. In response to the consultant’s comment, some of the grammar activities were reworded without the use of metalanguage. For example, instead of asking students to identify the head noun, they were given questions such as “in one word, what is this phrase about?”

Pilot Phase – Procedures

After receiving approval from the university’s Research Ethics Office, I applied for permission to conduct my research in two local school jurisdictions through the Cooperative Activities Program in the Faculty of Education. The application was forwarded to two school boards for their review; one of the boards was interested in the study due to its potential to improve the academic outcomes of its ELL population. As a result, I received ethics approval from that board and permission to contact the district ELL consultants. In an initial meeting with the consultants, I provided them with an introduction to the study and an explanation of the time and participant requirements for my pilot study; in turn, the consultants recommended two potential sites suitable for the experiment and connected me with the principals and ELL teachers at the schools. Two similar meetings were held with the school principals and ELL teachers. Fortunately, they agreed that the study addressed their schools’ learning goals and was relevant to Alberta Education’s Program of Studies. Although both principals provided permission to recruit ELL teachers and their students to participate in my research project, due to last minute staffing changes, one school withdrew two weeks before the pilot study was scheduled to begin. I was fortunate that the sole participating school provided a classroom and laptops so I was able to conduct my research in an ESL class. I was allotted two 60-minute periods for participant recruitment/consent and pre-post testing, and was allowed to attend the teachers' ESL classes two to three times a week (depending on the school schedule), for 30 minutes each time to assist and observe the students while they were completing the interventions/modules.

Although the board approved my ethics application, which included the administration of a standardized reading comprehension test, unfortunately, the teachers denied my request to administer the test. Their decision was based on their desire to avoid sacrificing instructional time for testing and their concerns regarding the additional stress it would place on students. As compensation, however, the teachers offered to share with me the results of the IPT3, which they administered at the beginning and end of the term.

The pilot study was conducted during the winter semester with ELLs from two different classes ($N = 34$) at the school. To control for background knowledge, students who had previously taken Science 10 were identified by the teachers and excluded from the study. After obtaining informed consent from the teachers and their students (consent forms in Appendix F), I administered the background questionnaire and the battery of pre-tests: the silent reading fluency tests Maze Forms A and C (three minutes per form), and the knowledge of nominal groups test (untimed, but completed by all students within 20 minutes). Participants were then randomly assigned to one of three interventions: traditional question and answer (Traditional; $n = 11$), AWL vocabulary (Vocab; $n = 11$), and scientific grammar (Grammar; $n = 12$). During my next visit, the participants received the website address for their assigned intervention and their individual login credentials. I provided technical assistance where necessary. The website functioned as intended, and the students began working on their assigned modules independently. I was present for the entirety of the experiment to monitor, to ensure that the students were on task, and to troubleshoot if needed. Participants completed one lesson each time I attended their class; the intervention phase lasted for three months, with the duration affected by factors such as student absences during my visits and school scheduling conflicts (e.g., fieldtrips, guest speakers, holidays, special events). The post-tests (the silent reading fluency test Maze Forms B and C, knowledge of nominal groups test) and the post-intervention questionnaire were group-administered upon completion of the last lesson.

Phase Two – Participants and Setting

Two additional cohorts of ELLs (*N* = 47) from another local high school within the same district (grades 10 to 12) participated in the second phase of this study. The school offered ESL programming in the form of locally developed courses – in Alberta, school jurisdictions have the option of providing specially designed, ministry approved, credited courses that complement the regular provincial programs of study. Participants in Phase Two were enrolled in two locally developed courses, ESL Expository English 15 and 25, to further develop their oral and written academic English proficiencies.

Phase Two participants' reported ages ranged from 15 to 18 years (*M* = 16.13, *SD* = .95). They self-identified as female (51.1%) and male (44.7%). The majority (76.6%) of the participants stated Philippines as their country of birth; the remaining 23.4% were evenly divided amongst Colombia, Cuba, Israel, Kosovo, Mexico, Nepal, Saudi Arabia, and Vietnam. Participants identified their L1s as Tagalog (70.2%), Spanish (10.6%), Cebuano (6.4%), Vietnamese (4.3%), Arabic (2.1%), Hiligaynon (2.1%), Nepali (2.1%), and Shqipe (2.1%). Their amount of schooling in the L1 varied from three to 11 years (*M* = 8.66, *SD* = 1.63). Participants also reported that they had received between one and 49 months of ESL instruction in Canada (*M* = 11.13, *SD* = 11.37).

Phase Two – Instruments

The background questionnaire, post-intervention questionnaire, knowledge of nominal groups test, silent reading comprehension and fluency tests, and online interventions remained unchanged from the Pilot Phase of the study.

Phase Two – Procedures

The intent of the second phase of this study was to maximize participant enrollment and to increase the power of the data analyses and the reliability of the results. With the help of a school district ESL consultant, five high schools with sizable ELL populations were identified and invited to participate. Two schools responded with interest; unfortunately, the principal at one of the schools would consider interventions totalling only a maximum of two and a half hours of class time. As a result, only one school from the district could be considered as a potential experiment site. At an introductory meeting with the ESL department, I explained the purpose and requirements for the study, and the teachers agreed that the intervention aligned with their school's ESL programming needs. Similar to the teachers in the Pilot Phase, they were apprehensive of any formal testing of their students, due to potential negative affect and stress. After assurances that the purpose of pre- and post-tests was to assess the effectiveness of the intervention, they reluctantly agreed to the administration of the knowledge of nominal nouns test and the silent reading fluency and comprehension test.

Phase Two of this study occurred during the fall semester with ELLs enrolled in two locally developed high school ESL courses: ESL Expository English 15 (*n* = 22) and ESL Expository English 25 (*n* = 25). Each course was taught by a different teacher. I followed the same data collection procedures outlined above in the Pilot Phase. After obtaining informed consent and administering the background questionnaire and pre-tests, I randomly assigned the participants to one of the three intervention conditions, Traditional (n = 15), Vocabulary (n = 15), and Grammar (n = 17). Students worked independently on their assigned modules during regular class time in two school computer labs under the supervision of the researcher and both teachers. The students completed 15 hours of intervention over two and half months. Post-tests and the post-intervention questionnaire were group-administered during the class after students had completed the final lesson.

Data Analysis

The corpus of digitalized textbooks was analyzed using three automatized text analysis tools (TAASC, Lextutor, and Coh-Metrix). Each chapter was individually inputted into each of the tools. The tools returned quantitative measures of linguistic sophistication which are presented in detail in the next chapter.

The quantitative data gathered from the experiment were analyzed using SPSS version 25.0 and R version 3.4.1. Descriptive statistics were calculated for the pre- and post-test results from the knowledge of nominal groups test, silent reading fluency tests, and global reading comprehension test. Repeated measures analysis of variance was used to examine statistically significant changes over time, from pre-test to post-test. Gains were also calculated by subtracting the students' pre-test scores from their post-test scores, and these served as the dependent variables for subsequent inferential statistical analyses. Multivariate analysis of variance was used to identify statistically significant group differences on all dependent measures (gain scores on the knowledge of nominal groups test, silent reading fluency tests, and global reading comprehension test) as a result of the independent variable: the intervention conditions of traditional, vocabulary, and grammar. Confidence intervals and measures of effect sizes were also calculated.

Qualitative data from field notes and post-intervention questionnaires were quantified as well as thematically analyzed following steps outlined by Braun and Clarke (2006). After a thorough review of the data, initial codes were generated for interesting features. Preliminary themes were created by collating the related codes. The process of theme identification was also facilitated by the structured nature of the post-intervention questionnaire, given that the participants' responses were already centred on specific topics (e.g., cause of difficulty in reading). The coded data and themes were reviewed for congruence. The qualitative data were reported in a separate section from the quantitative data, but they were also integrated in the discussion chapter to provide additional insights and to augment interpretations of the quantitative data.

Chapter Summary

In Chapter Three, I have provided an overview of the design of the study, the data collection instruments, the interventions, and the procedures. Based on the previous arguments for the inclusion of grammar teaching as part of a reading intervention program presented in Chapter Two, a grammar intervention and two other sets of interventions (a traditional practice condition and a vocabulary-focused condition) were created for this dissertation study to compare their effects on high school ELLs' reading comprehension and fluency. By collecting multiple data sources from a pre-test post-test classroom-based experiment, a post-intervention questionnaire, and field observations, I aimed to uncover findings that could inform the practices of teachers of high school ELLs. In Chapter Four, I present the findings from my data analyses.

Chapter Four – Results

Introduction

To better understand the linguistic qualities of content-area texts, two grade 10 science textbooks were digitalized and then analyzed using automated text analysis tools. Using passages from the textbooks, three reading interventions were created. Two randomized, pre-post classroom-based experiments (pilot study and main study) were conducted to examine the comparative effects of the interventions on high school ELLs' reading comprehension and fluency. Responses from post-intervention questionnaires and field notes provided additional insight into ELLs' attitudes towards the intervention. The study results are reported in six sections, organized according to the research questions below:

1. What are the linguistic characteristics of the two Alberta Education approved Science 10 textbooks in terms of lexicon, syntactic construction, and passage-level readability?
2. What are the comparative effects of three online modules (traditional reading comprehension, AWL vocabulary, scientific grammar) on ELLs' passage-level reading comprehension
    a. as measured by a commercialized standardized exam?
    b. as measured by maze passages generated from a Science 10 textbook?
3. What are the comparative effects of three online modules (traditional reading comprehension practice, academic vocabulary practice, or grammar practice) on ELLs' passage-level reading fluency, as measured by performance on maze passages (accuracy in three minutes) constructed from a grade 10 science textbook?

4. What are the comparative effects of three online modules (traditional reading comprehension practice, academic vocabulary practice, or grammar practice) on ELLs' understanding of nominal structures as measured by a noun phrase interpretation activity?
5. What are the ELLs' attitudes (positive or negative) toward the use of the modules in their classes?
6. What are the ELLs' conceptions of the role of the online modules in their learning?

Textbook Analysis

The first research question concerns the linguistic characteristics of the two textbooks that are approved for use in Alberta high school science classrooms. Automated text analyzers were used to generate statistics related to lexis, syntax, and passage-level readability. Results from the analyses are presented in the ensuing subsections.

Vocabulary Profile

Results from the Lextutor Lexical Frequency Profile (LFP) textbook analyses are presented in three parts: word counts and lexical density; lexical coverage by AWL; and lexical coverage by frequency bands. Types (number of unique words), tokens (total number of words), and lexical density by chapters are shown in Table 1. Although the number of types and tokens varied across chapters and books, lexical density appears to remain relatively consistent. Lexical coverage by the AWL, as seen in Table 2, shows that the AWL words constitute a sizable portion of the words included in all of the chapters in both texts, ranging from 5.48% to 12.8% (M = 8.64%) of total tokens. In all, 435 AWL families were found in both books (375 in *Addison Wesley Science 10*; 373 in *ScienceFocus10*). Some AWL words occurred more frequently than others in certain chapters. For example, a chapter on the topic of energy systems contained 294 tokens of "energy", but a chapter on biology included only two tokens of the same word.

However, some AWL families appeared across most, if not all, of the 24 chapters of the two texts. The top ten common AWL families found in both texts are as follows (total number of chapters of both textbooks combined in parentheses): section (24), process (24), similar (23), involve (23), energy (23), occur (22), release (21), require (20), react (20), create (20), and chemical (20).

Table 1

Word Counts by Types and Tokens, and Calculated Lexical Density

	Addison Wesley Science 10			*ScienceFocus 10*		
Chapter	Types	Tokens	Lexical Density	Types	Tokens	Lexical Density
1	1240	4960	0.58	1074	5817	0.61
2	1766	12189	0.58	1035	5432	0.59
3	1031	6091	0.56	866	4883	0.59
4	735	4905	0.54	1324	7084	0.57
5	1027	5375	0.56	991	6106	0.53
6	1359	6914	0.58	728	3037	0.57
7	1248	5231	0.58	1068	4286	0.60
8	1191	6047	0.55	900	4628	0.58
9	1335	7071	0.55	925	4858	0.59
10	834	3556	0.57	1226	7171	0.58
11	1288	9499	0.58	1130	4990	0.59
12	1063	4772	0.61	1385	6781	0.60

Table 2

Lexical Coverage by the AWL

	Addison Wesley Science 10				*ScienceFocus 10*			
Chapter	Families	Types	Tokens	Token (% of Text)	Families	Types	Tokens	Token (% of Text)
1	108	155	374	7.54%	116	173	572	9.83%
2	140	207	882	7.23%	97	144	461	8.49%
3	88	133	675	11.08%	90	129	623	12.76%
4	64	88	328	6.69%	124	176	529	7.47%
5	84	125	688	12.80%	95	139	520	8.52%
6	137	201	743	10.75%	70	98	364	11.99%
7	127	184	439	8.39%	102	140	310	7.23%
8	132	191	526	8.70%	94	134	371	8.02%
9	137	203	464	6.56%	91	122	266	5.48%
10	85	111	257	7.23%	113	171	472	6.58%
11	115	172	757	7.97%	110	154	340	6.81%
12	132	198	453	9.49%	164	241	657	9.69%

Note. Token (% of Text) = percentage of text covered by AWL tokens

The distribution of lexical coverage by each frequency band, including off-list words (e.g., proper nouns) from the LFP (see Figure 3), revealed that the vocabulary used by the textbook authors spanned frequency bands from K-1 (first thousand most frequent word families) to K-25 (25th thousand most frequent word families – or very rarely encountered words). An examination of the cumulative token coverage revealed an inflection point at the K-3 band, showing that most of the words in the textbooks consist of the first 3,000 most common word families; however, 95% coverage, which is required for adequate comprehension, is not reached with only the K-1 to K-3 words in any of the chapters. In fact, 95% coverage only begins to appear when the K-4 band is included, but only for chapter six in *ScienceFocus10*. For the majority of the chapters in the textbook, 95% coverage is reached at K-7 or above. Also shown in Figure 3 is the unique token coverage by each frequency band. With the exception of the K-1 band, all of the frequency bands individually cover less than 16% of the text. The sharpest decline comes after the K-3 band. It can be seen that the AWL words cover a higher percentage of the tokens in the texts compared with the K-4 to K-25 bands.

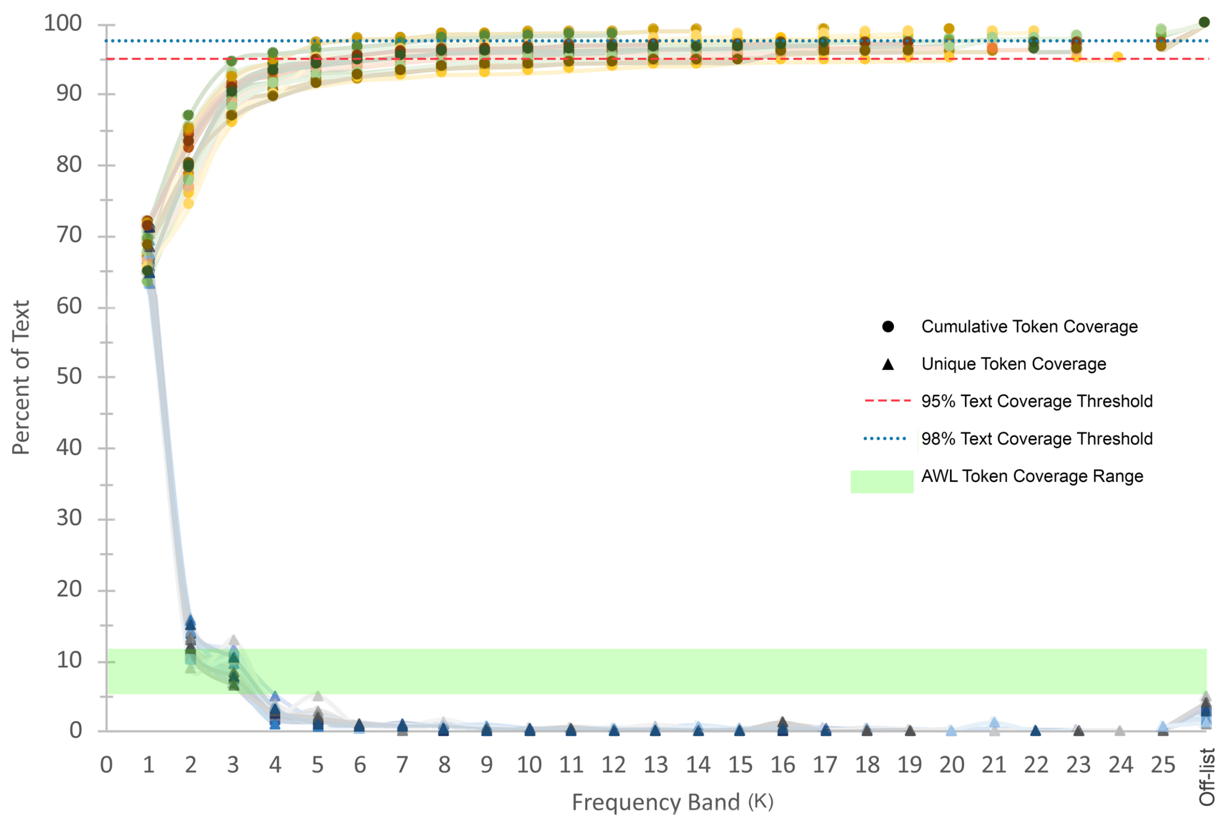


Figure 3. Lexical frequency profile of both textbooks by cumulative and unique token coverage. Each individual node depicts a separate chapter and the range of coverage by AWL tokens is overlaid for comparison.

Syntactic Sophistication and Complexity

Two measures from the textbook analysis using the TAASSC provided indications of the presence of complex noun phrases: complex nominals per T-unit and complex nominals per clause (Figure 4). The results show the average number of complex nominals per T-unit/clause found in each chapter, with higher numbers of complex nominals found per T-unit than per clause. In addition, the *Addison Wesley Science 10* text shows a pattern of increasing use of complex nominals from the beginning to the final chapters, whereas the number in *ScienceFocus 10* remain relatively stable across all chapters.

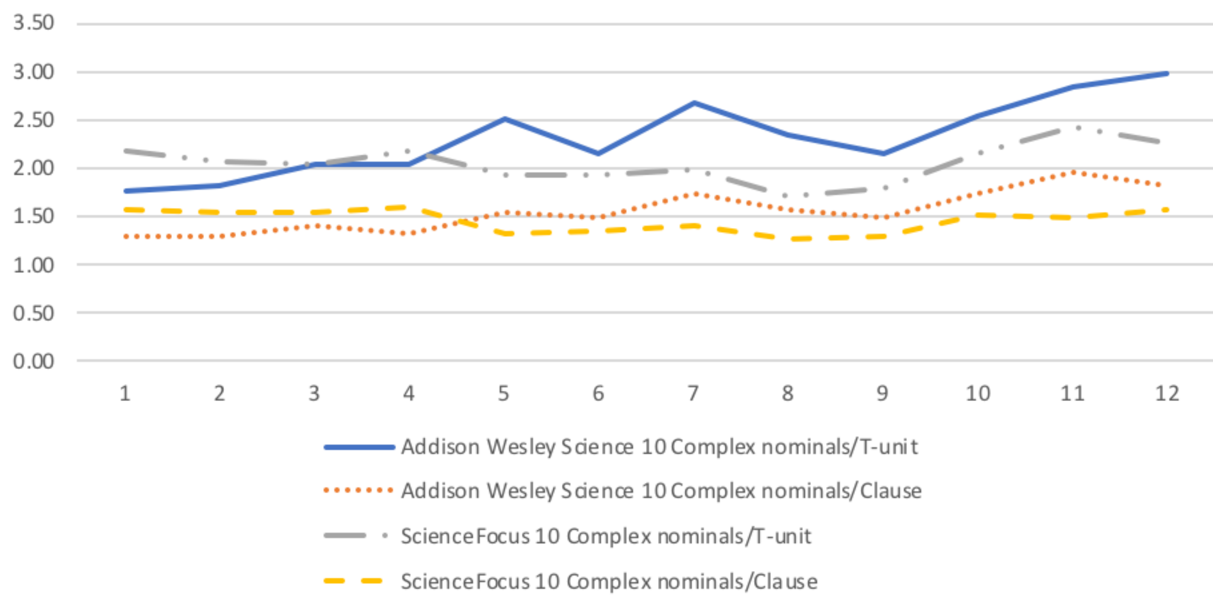


Figure 4. Syntactic complexity measured as average of number of complex nominals per T-unit and clause. Chapter numbers are plotted on the horizontal axis.

Passage-Level Readability

Results of the text readability analysis from Coh-Metrix (Figure 5) show that, overall, both textbooks ranked low in narrativity, with most of the chapters below the 20th percentile (*M* = 17.92, *SD* = 5.07) as measured against the TASA corpus. Conversely, syntactic simplicity was found to be relatively high, ranging from approximately the 60th to 90th percentile (*M* = 83.13, *SD* = 8.65). Word concreteness was lower, with a mean at the 52nd percentile (*SD* = 19.56). The texts ranked relatively high on referential cohesion (*M* = 73rd percentile, *SD* = 15.75) and slightly higher than average on deep cohesion (*M* = 56th percentile, *SD* = 11.27). As shown in Figure 4, although variability exists across the chapters, both books appear to exhibit similar patterns of variation. Traditional readability, as measured by the Flesch-Kincaid grade level for all chapters in both books combined, was 9.6 (*SD* = 0.97; range = 8.1 to 11.8). However, as with the other readability measures presented above, variation in the Flesch-Kincaid index existed across chapters and books, with the *Addison Wesley Science 10* text showing greater variability (*SD* = 1.07) than the *ScienceFocus* text (*SD* = 0.64). In fact, the *Addison Wesley* text again showed a trend of gradual increase in difficulty across chapters from the beginning to the end of the book.

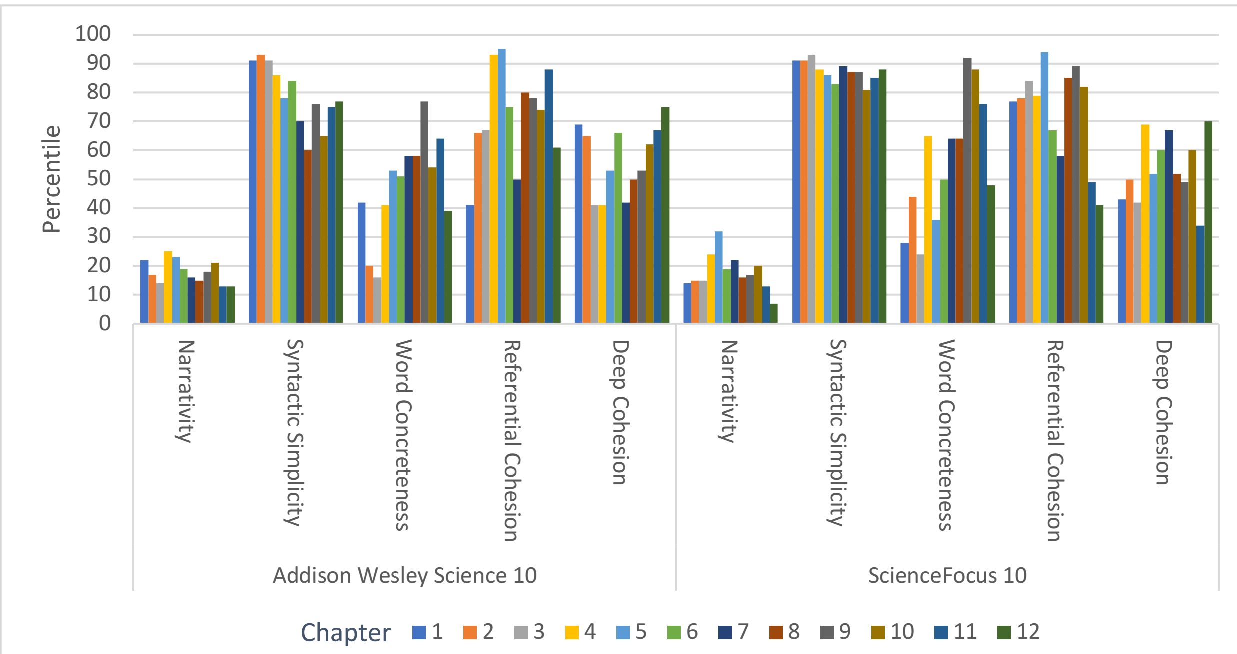


Figure 5. Coh-Metrix measures of text readability. Percentile scores are normed against the TASA corpus.

Comparative Effects of the Intervention Conditions

The focus of research questions two, three, and four was the comparative effects of the traditional question and answer module (Traditional), the AWL vocabulary module (Vocabulary), and the scientific grammar module (Grammar) on the participants' passage-level reading comprehension, silent reading fluency, and comprehension of complex noun phrases. Results of the experiment are presented in two parts: the Pilot Phase and Phase Two (the main study). Due to the use of multivariate analysis techniques, results to all three research questions are presented together.

Pilot Phase Results

Pilot Experiment. The data were analyzed using SPSS version 25.0 and R version 3.4.1. Descriptive statistics were calculated for the pre- and post-tests (the knowledge of nominal groups test, silent reading fluency tests [Maze A, B, C], and global reading comprehension – IDEA Proficiency Test [IPT3]). Corresponding gain scores were computed by subtracting pre-test scores from the post-test scores. The results, summarized in Table 3, show that all three groups made overall gains in mean scores on all four measures, with the exception of the vocabulary group on the IPT3 exam (a minor decrease of 0.64 point). Repeated measures ANOVA showed the gains over time were statistically significant for two versions of the silent reading fluency tests: Maze A and B, $F(1, 30) = 82.042$, $p < .000$ and Maze C, $F(1, 33) = 142.946$, $p < .000$. No statistically significant differences over time were observed for the knowledge of nominal groups test $F(1, 30) = 2.284$, $p < .141$ or the IPT3 $F(1, 33) = 2.830$, $p < .103$.

No statistically significant differences were found between the groups on ANOVAs conducted on any of the pre-tests: knowledge of nominal groups, $F(2,31) = 0.064$, $p = .938$; Maze Form A, $F(2,31) = 1.850$, $p = .174$; Maze Form C, $F(2,31) = 2.801$, $p = .076$; IPT3, $F(2,31) = 0.286$, $p = .753$. Of interest in the present study was the comparative gains made by groups; therefore, the following gain scores (inherently accounting for pre-test scores) served as the scores for the dependent variables in the following multivariate analysis: gains on knowledge of nominal groups test (GainNouns), gains made from Maze Forms A to B (GainMazeAB), gains on Maze Form C (GainMazeC), and gains on the IPT3 (GainIPT).

The data were checked for suitability for using multivariate analysis of variance. An examination of box plots of the gain scores revealed one to two univariate outliers in each group; after verification that they were not the result of data entry error, these data points were retained in the analysis. Shapiro-Wilk's test of normality ($p < .05$) confirmed normal distributions of GainMazeAB scores for all three groups, GainMazeC for Vocabulary and Grammar groups, and GainIPT for the Vocabulary group; MANOVA is robust to minor deviations from normality in the dependent variables (Tabachnick & Fidell, 2013). The absence of multivariate outliers was confirmed through Mahalanobis distance values using $\chi^2\ (4, N = 34) = 18.47, p > .001$. No multicollinearity was found with Pearson correlation. Scatterplots were generated for each pair of gain scores; approximately linear relationships were present between the GainMazeAB, GainMazeC, and GainIPT, but absent for GainNouns. Homogeneity of variance-covariances matrices was assessed and verified by Box's test of equality of covariance matrices ($p = .387$).

A one-way MANOVA was performed to determine the existence of statistically significant group differences on GainMazeAB, GainMazeC, and GainIPT. Multivariate test results showed statistically significant differences between the groups on the combined dependent variables, $F(6, 60) = 2.832$, $p = .017$; Pillai's Trace = .441; partial eta squared $\eta_p^2 = .221$, 90% CI = [0.024, 0.298]. Follow-up one-way ANOVAs indicated that GainMazeC ($F(2,31) = 4.875$, $p = .014$, $\eta_p^2 = 0.239$, 90% CI = [0.030, 0.398]) and GainIPT ($F(2,31) = 3.465$, $p = .044$, $\eta_p^2 = 0.183$, 90% CI = [0.003, 0.341]) were significantly different between the three groups, but no statistically significant differences were found for GainMazeAB ($F(2,31) = 1.095$,

Table 3

Descriptive statistics of group performance on dependent variables from the pilot study

Test	Group[a]	Pre-test *M*	Pre-test *SD*	Pre-test 95% CI	Post-test *M*	Post-test *SD*	Post-test 95% CI	Gain score *M*	Gain score *SD*	Gain score 95% CI
Knowledge of nominal groups	Traditional	16.33	2.92	[14.94, 20.51]	20.00	3.54	[17.59, 22.04]	2.09	4.55	[-0.96, 5.15]
	Vocabulary	18.09	4.44	[15.11, 21.07]	18.55	3.64	[16.10, 20.99]	0.45	4.08	[-2.29, 3.20]
	Grammar	17.90	3.32	[15.39, 19.61]	18.60	2.32	[15.85, 19.99]	0.42	3.75	[-1.97, 2.80]
Silent reading fluency and comprehension (Maze Forms A & B)	Traditional	13.33	3.50	[10.64, 16.02]	21.22	5.67	[16.86, 25.58]	7.91	4.81	[4.68, 11.14]
	Vocabulary	10.36	6.22	[6.19, 14.54]	17.91	7.20	[13.07, 22.75]	7.55	6.44	[3.22, 11.87]
	Grammar	9.40	5.95	[5.15, 13.65]	19.50	5.28	[15.73, 23.27]	10.58	4.87	[7.49, 13.68]
Silent reading fluency and comprehension (Maze Form C)	Traditional	19.22	5.19	[15.23, 23.21]	29.89	5.01	[16.82, 30.78]	10.73	2.83	[8.82, 12.63]
	Vocabulary	17.09	7.01	[12.38, 21.80]	22.64	9.33	[16.37, 28.90]	5.55	4.95	[2.22, 8.87]
	Grammar	13.30	5.96	[9.03, 17.57]	23.80	9.76	[16.82, 30.78]	9.92	4.50	[7.06, 12.78]
Global reading comprehension (IPT3)	Traditional	46.36	6.49	[42.01, 50.72]	49.00	4.96	[45.67, 52.33]	2.55	2.81	[0.66, 4.43]
	Vocabulary	46.73	5.73	[42.88, 50.58]	46.09	6.61	[41.65, 50.53]	-0.64	3.20	[-2.79, 1.51]
	Grammar	43.00	8.50	[37.60, 48.40]	43.67	7.90	[38.65, 48.69]	0.67	2.54	[-0.94, 2.25]

[a] Traditional, $n = 11$; Vocabulary, $n = 11$, Grammar, $n = 12$

p = .347, η_p^2 = 0.066, 90% CI = [0.000, 0.198]). Tukey HSD post-hoc tests (Table 4) revealed that the Traditional group made statistically significant greater gains (α = .05) than the Vocabulary group on the IPT3 and Maze Form C, and the Grammar group experienced statistically significant greater gains than the Vocabulary group on Maze Form C.

A separate ANOVA for GainNouns as the dependent variable was performed due to the lack of a linear relationship with the other three dependent variables, GainMaze AB, GainMazeC, and GainIPT (thus reducing the power of the MANOVA). No outliers were found upon inspection of the boxplots. GainNouns scores were normally distributed for the Traditional, Vocabulary, and Grammar groups as assessed by Shapiro-Wilk's test (p > .05). Homogeneity of variance was assessed by Levene's test (p = .722). The differences between the groups on their GainNouns scores were not statistically significant (F(2, 31) = 0.599, p = .556, η_p^2 = 0.066, 90% CI = [0.000, 0.147]).

Post-Intervention Questionnaire Results

Preliminary analysis of the post-intervention questionnaire results showed favourable reception of the interventions by the students. Student responses to the questions were combined with the additional qualitative data gathered in Phase Two of the study and thematically analyzed (Braun & Clarke, 2006).

Table 4

Pilot phase post hoc Tukey HSD multiple comparisons

Comparison	Mean difference	*SE*	*p*	95% CI	Cohen's *d*
Maze AB					
Traditional-Grammar	-2.67	2.26	.471	[-8.23, 2.88]	-0.55
Traditional-Vocabulary	0.36	2.31	.986	[-5.31, 6.04]	0.06
Grammar-Vocabulary	3.04	2.26	.381	[-2.52, 8.59]	0.53
Maze C					
Traditional-Grammar	0.81	1.76	.889	[-5.13, 3.51]	0.22
Traditional-Vocabulary	5.18	1.79	.019	[0.77, 9.59]	1.29
Grammar-Vocabulary	4.37	1.76	.047	[0.05, 8.69]	0.92
IPT3					
Traditional-Grammar	1.88	1.19	.270	[-1.05, 4.81]	0.70
Traditional-Vocabulary	3.18	1.10	.035	[0.19, 6.17]	1.06
Grammar-Vocabulary	1.30	1.19	.524	[-1.63, 4.23]	0.45

Phase Two Results

The data collected in Phase Two was amalgamated with the data from the Pilot Phase for quantitative analysis ($N = 81$) using SPSS version 25.0 and R version 3.4.1. A summary of the descriptive statistics for the pre- and post-tests of the knowledge of nominal groups test, silent reading fluency tests, and corresponding gain scores (computed by subtracting pre-test scores from the post-test scores) are presented in Table 5. Overall, all three groups exhibited gains on all measures, with the Grammar group achieving the highest gain scores on all measures. A multivariate one-way repeated measures ANOVA confirmed statistically significant differences between the pre-test and post-test scores on all three measures, $F(3, 29) = 97.463$, $p < .0005$; Pillai's = 0.789, partial eta squared $\eta_p^2 = 0.789$.

Table 5

Descriptive statistics of group performance on dependent variables from Phase Two

Test	Group[a]	Pre-test *M*	Pre-test *SD*	95% CI	Post-test *M*	Post-test *SD*	95% CI	Gain score *M*	Gain score *SD*	95% CI	Cohen's *d*
Knowledge of nominal groups	Traditional	17.35	4.02	[15.72, 18.97]	18.77	3.55	[17.34, 20.20]	1.42	3.57	[-0.02, 2.86]	0.37
	Vocabulary	19.08	3.36	[17.72, 20.43]	19.19	3.07	[17.95, 20.43]	0.12	3.22	[-1.18, 1.41]	0.03
	Grammar	17.34	3.55	[16.00, 18.69]	19.07	3.21	[17.85, 20.29]	1.72	3.59	[0.36, 3.09]	0.51
Silent reading fluency (Maze Forms A & B)	Traditional	10.88	4.30	[9.15, 12.62]	20.31	7.82	[17.15, 23.47]	9.42	6.89	[6.64, 12.21]	1.49
	Vocabulary	9.19	4.83	[7.24, 11.14]	17.35	5.80	[15.00, 19.69]	8.15	5.55	[5.91, 10.39]	1.53
	Grammar	8.03	4.34	[6.38, 9.68]	19.14	7.23	[16.39, 21.89]	11.10	7.54	[8.24, 13.97]	1.87
Silent reading fluency (Maze Form C)	Traditional	16.27	6.94	[13.46, 19.07]	25.92	10.99	[21.48, 30.36]	9.65	5.71	[7.35, 11.96]	1.05
	Vocabulary	15.54	6.45	[12.94, 18.14]	22.69	8.46	[19.23, 26.15]	7.15	5.21	[5.05, 9.26]	0.95
	Grammar	13.07	5.24	[11.07, 15.06]	24.28	8.63	[20.99, 27.56]	11.21	5.00	[9.30, 13.11]	1.57

[a] Traditional, *n* = 26; Vocabulary, *n* = 26, Grammar, *n* = 29

Following recommendations by Larson-Hall (2017), “data accountable” graphics were prepared to visually represent the data and help reveal additional patterns that may be hidden in the summary statistics. Scatterplots of individual gain scores against pre-test scores were overlaid with regression lines for the Knowledge of Nominal Groups test (Figure 6) and the Silent Reading Fluency and Comprehension tests (Figure 7). The regression lines in Figure 6 show that, in all three groups, participants who initially scored lower on the nouns test made the most gains over time. A different pattern emerged on the plot of combined scores from the two versions of the maze test. As shown in Figure 7, the slopes and directions of the regression lines are different across the groups. It appears that students with lower pre-test scores experienced higher gains if they were in the Vocabulary group. Contrarily, the Traditional and Grammar groups exhibited a reverse pattern: the higher their pre-test performance, the greater their gains on the silent reading fluency and comprehension measure after the intervention.

In addition to the fitted regression lines on Figures 6 and 7, two lines were superimposed on the plots. The first is a horizontal baseline at a gain of zero. A second vertical line segment is drawn between the zero baseline and the regression lines at an x-coordinate equal to the participants’ mean pre-test scores. In effect, the length of this line segment shows the expected gain of the “average participant” independent of different group pre-test means (Jensen & Vinther, 2003). Overall, it appears that the Grammar group made the greatest gains.

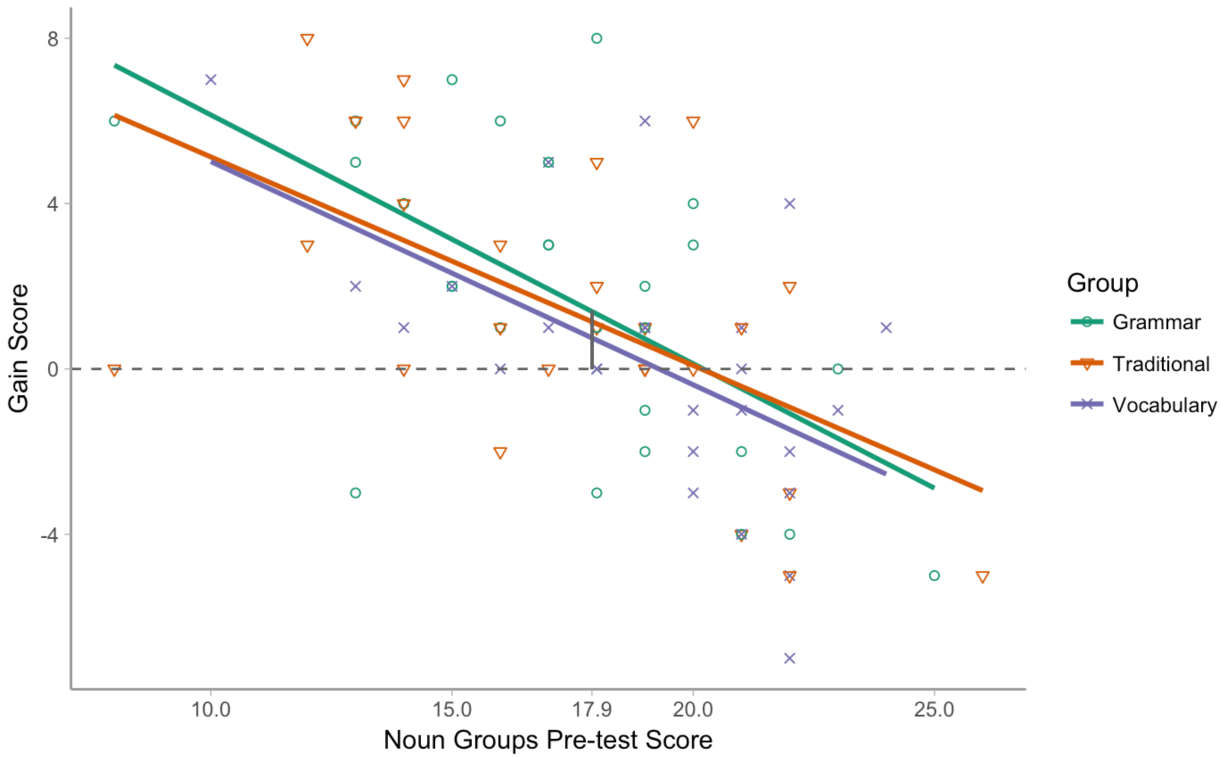


Figure 6. Scatterplot of individual gain scores versus pre-test scores on the knowledge of nominal groups test. Fitted regression lines are added for each group. The length of the vertical line segment between the regression lines and the horizontal dashed line (zero gain) at point 17.9 (the overall pre-test mean score) can be interpreted as the average participant gain.

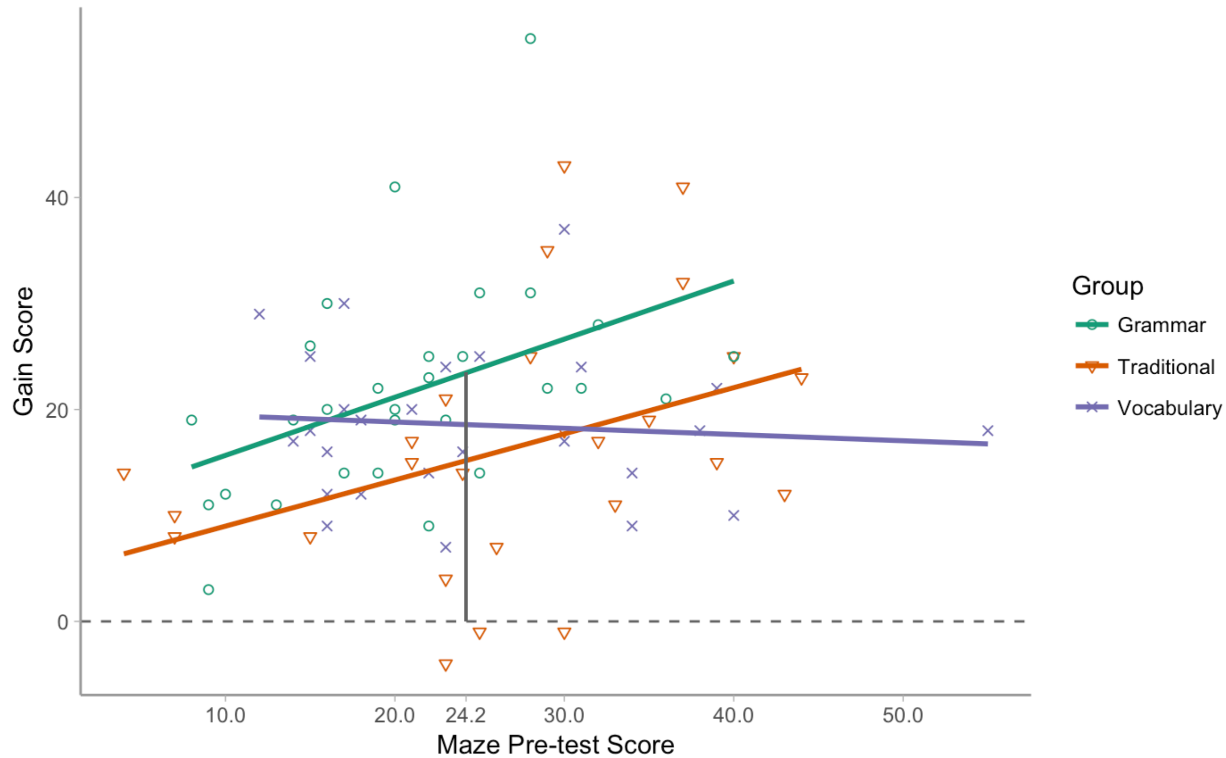


Figure 7. Scatterplot of individual gain scores versus pre-test scores on the silent reading fluency and comprehension test (both forms combined). Fitted regression lines are added for each group. The length of the vertical line segment between the regression lines and the horizontal dashed line (zero gain) at point 24.2 (the overall pre-test mean score) can be interpreted as the average participant gain.

Separate ANOVAs conducted on the pre-test scores failed to reveal statistically significant differences between the three groups: knowledge of nominal groups, $F(2,78) = 1.988$, $p = .144$; Maze Form A, $F(2,78) = 2.779$, $p = .068$; Maze Form C, $F(2,78) = 2.026$, $p = .139$. As with the analysis for the Pilot Phase, the gain scores (GainNouns, GainMazeAB, GainMazeC) served as dependent variables to determine the comparative effects of the interventions.

Prior to conducting the multivariate analysis, assumptions of normality and homogeneity of variance were checked. Normality was assessed and confirmed with Shapiro-Wilk's test ($p < .05$). Pearson product moment correlations between the dependent variables did not reveal multicollinearity. Absence of multivariate outliers was determined through assessment of Mahalanobis distance using $\chi^2\ (3, N = 81) = 16.27$, p > .001. Linear relationships appeared in scatterplots between pairings of GainNouns, GainMazeAB, and GainMazeC. Box's test of equality of covariance matrices confirmed homogeneity of variance-covariance ($p = .908$).

A one-way MANOVA was conducted to determine the effects of comparative online reading modules on GainNouns, GainMazeAB, and GainMaze C. As shown in Table 6, students in the Grammar and Traditional groups achieved greater gains than the Vocabulary group on all three dependent measures. The difference between the groups on the combined dependent variables was statistically significant, $F(3, 77) = 3.433$; Roy's Largest Root = .134; $p = .021$, $\eta_p^2 = 0.118$, 90% CI = [0.010, 0.210]. Follow-up one-way ANOVAs showed statistically significant differences in GainMazeC scores across the three groups, $F(2, 78) = 4.044$, $p = .021$, $\eta_p^2 = 0.094$, 90% CI = [0.008, 0.192], but not for GainNouns, $F(2, 78) = 1.632$, $p = .202$, $\eta_p^2 = 0.040$, 90% CI = [0.000, 0.116], or GainMazeAB, $F(2, 78) = 1.328$, $p = .271$, $\eta_p^2 = 0.033$, 90% CI = [0.000, 0.103]. Tukey HSD post-hoc test (Table 4) indicated that the Grammar group ($M = 11.21$, $SD =$

5.00) made statistically significant greater gains than the Vocabulary group (*M* = 7.15, *SD* = 5.21) on GainMazeC.

Table 6

Phase Two post-hoc Tukey HSD multiple comparisons

Comparison	Mean difference	*SE*	*p*	95% CI	Cohen's *d*
Nouns test					
Traditional-Grammar	-0.30	.94	.945	[-2.54, 1.94]	-0.08
Traditional-Vocabulary	1.31	.96	.367	[-0.99, 3.60]	0.38
Grammar-Vocabulary	1.61	.94	.205	[-0.63, 3.85]	0.47
Maze AB					
Traditional-Grammar	-1.68	1.82	.628	[-6.03, 2.67]	-0.23
Traditional-Vocabulary	1.27	1.82	.777	[-3.20, 5.74]	0.20
Grammar-Vocabulary	2.95	1.82	.243	[-1.40, 7.30]	0.45
Maze C					
Traditional-Grammar	-1.55	1.43	.527	[-1.87, 4.98]	0.29
Traditional-Vocabulary	2.50	1.47	.212	[-1.01, 6.01]	0.46
Grammar-Vocabulary	4.05	1.43	.016	[0.63, 7.48]	0.80

Post-Intervention Questionnaire Results

Research questions 5 and 6 involved an investigation of ELLs' attitudes toward the use of the modules, as well as their conceptions of the role of the interventions in their learning. Student responses to the post-intervention questionnaires administered during both phases of the study were analyzed together. Quantitative data (Likert-type scale ratings) were summarized and presented as descriptive statistics; qualitative data (respondent comments) were thematically analyzed (Braun & Clarke, 2006). In total, 55 students from both phases of the study completed the questionnaire, yielding a 68% response rate. The majority of respondents (87%) indicated positive interest in using the intervention website, and 91% of the students agreed that the website helped them become better readers in English. However, a slightly lower percentage of students (75%) stated that they were interested in science in school. No discernable differences in response patterns were found between groups.

Students were asked to rate their overall difficulty in their ability to use the website. Based on their experience with the website, the majority of them rated it as either very easy (42%) or easy (38%) to use. The remaining 20% rated the website as somewhat difficult to use. Two themes emerged from my examination of the participants' explanations of their responses. These themes coincided with an easy/difficult dichotomy. Those who found the website easy to use ascribed their reasons to the ease of the website use. For instance, some welcomed the simplicity of the layout, describing it as "not that flashy to the eyes [which] made it easy to navigate" (Participant 6) or "it's easy [to use] because the interface is very simple" (Participant 24) and "straightforward" (Participant 43). Others found the website easy to access "because [you] just go in the web" (Participant 51), although a few indicated a dislike for the additional login process (for the purpose of individual progress tracking and restricting access to assigned experimental conditions). Participant 32 equated using the website to "basic use of a computer or laptop", but this idea was not universally shared, as Participant 45 cautioned that "it might be confusing for people who are inexperienced with using computer[s]." Those who found the website difficult to use attributed their difficulty to the content of the website (i.e., the difficulties they experienced when reading the science passages).

When asked about their perceived difficulty of the reading passages used in the intervention, 44% of the respondents found the readings "somewhat easy", 37% "somewhat difficult", 15% "very easy", and 4% "very difficult". Unfortunately, due to the anonymous nature of the questionnaire, I was unable to connect these responses to the participants' pre-test scores to determine if proficiency influenced the responses. The prevalent theme for perceived difficulty was related to vocabulary. For example, participants attributed the difficulty of the readings to vocabulary, through statements such as:

“[The readings were] kind of difficult because I have never seen the words before” (Participant 8);

“It's a little bit difficult for me sometimes because there [are] some words that [are] really hard to read” (Participant 15);

“There is some word that I can’t understand” (Participant 21).

“Some words are unfamiliar to me” (Participants 34 & 38).

Conversely, some students found that their knowledge of words led them to feel that the readings were “not that hard”. Others found the readings relatively easy because of the vocabulary scaffold (in-line gloss) built into the online modules: Participant 24 stated that “some words are new to me, but I can just hover over the words and it’ll show the definitions.” Aside from vocabulary, some participants believed that the locus of difficulty lay within themselves because they were (un)familiar with the topics or they possessed inadequate reading skills (e.g., “I’m not good [at] reading” [Participant 41]).

The majority (46%) of participants found the intervention activities to be “somewhat easy”, 33% “somewhat difficult”, 20% “too easy”, and 2% “too difficult”. Self-reported reasons for the students’ ratings varied widely, but the three most prevalent themes were (1) the answers to the module activities could be found by reading the passages or questions carefully; (2) students’ interest (or the lack thereof) in science affected their motivation; and (3) the students liked the activity formats. With relation to the third theme, students appreciated the availability of immediate feedback in the follow-up activities. For example, Participant 5 found the interventions “easy to use…you can see clearly the answers right away”. Others found the interactive nature of the activities (e.g., drag-and-drop) to be fun. Some students found the activities “a bit repetitive, but that helps you remember” (Participant 23). Participant 13 echoed this sentiment: “The activities are fun. After doing it again and again, it felt easier to do some of it.”

Despite the differing levels of perceived difficulties, 85% of the students expressed a desire to continue using the website independently for additional practice. The same percentage of students thought that a similar set of online activities would help them cope with readings in other academic subjects. The positive responses were accompanied by comments that reflected the enjoyment that students derived from using the website and their belief that the website contributed to improvements in their English reading abilities. They also believed that it might bolster their future reading abilities and academic performance in science. For example, Participant 6 wrote, “It is enjoyable especially when you are aiming to take a science heavy course.” Similarly, Participant 14 stated, “Yes, I would love to continue using this website. It helped me a lot about my grammar and reading!” This feeling of progress also enhanced feelings of enjoyment from using the website. The students appreciated the instant feedback provided on the activities. The 15% of participants who did not intend to continue using the website cited their busy schedule outside of school as a barrier to using it on their own.

Finally, students provided critical feedback and suggestions of potential improvements to the website and the interventions. Their comments were related to two themes: visual aesthetics and additional content. Some students found the website’s black and white color palette to be visually plain and suggested adding user-customizable color options, including a “nightmode since most [students] might use it at night” (Participant 13). They also suggested adding video tutorials, more lessons, and readings on additional subjects beyond science.

The comments received from the questionnaire are consistent with researcher field note observations and informal conversations with students and teachers. All participants believed that the interventions complemented Alberta Education's Science 10 Program of Studies and would help them improve their academic standings. At both collection sites, teachers and students constantly offered remarks indicating a feeling of advancement in reading ability and increased confidence in their content area courses. Participant 47 offered, "I loved the website because it helped me to learn better and I would recommend other people to use this website." The school district ESL consultant shared with me feedback she had received that "[the school staff] are very happy with your intervention and feel like your work with these students has been integral in helping them to progress and improve their English language skills, especially with your use of non-fiction texts."

Results Summary

The textbook analysis provided a nuanced examination of the lexical, syntactic, and passage-level features of two grade 10 science books and provided evidence to support the necessity to develop a grammar intervention at this grade level. The focus of the experiment was to discern the comparative effects of the three forms of intervention provided during this study; results showed that all three were effective in promoting reading comprehension and fluency, but to varying degrees. These results are discussed in further detail in the next chapter.

Chapter Five – Discussion, Summary and Conclusions

Overview

The textbook analysis together with the experimental study and the post-intervention questionnaire have provided insights that may lead to pedagogical advancements for improving high school ELLs' reading comprehension. Results presented in the previous chapter will be discussed in six sections that correspond to the study research questions: (a) the linguistic characteristics of two grade 10 science textbooks, (b) the comparative effects of three online modules (traditional reading comprehension, AWL vocabulary, scientific grammar) on ELLs' passage reading comprehension, (c) the comparative effects of the three online modules on ELLs' silent reading fluency, (d) the comparative effects of the three online modules of the three modules on ELLs' comprehension of complex noun groups, (e) the ELLs' attitudes towards the use of the modules, and (f) the ELL's conceptions of the role of the online modules in their learning.

Following the discussion centred on the study research questions, I address the contributions of the study to L1 and L2 reading research and SLA theory, and I discuss the implications of the study for pedagogy and the use of technology in language research and teaching. I conclude with the limitations of the study, suggestions for future research, and a brief summary of the main findings.

Discussion of Results

Research Question 1:

What were the linguistic characteristics of the two Alberta Education approved Science 10 textbooks in terms of lexicon, syntactic construction, and passage-level readability?

The exploratory analysis of two grade 10 science textbooks was conducted to gain a nuanced understanding of the linguistic characteristics of the books: their lexical frequency profile (LFP), syntactic complexity in terms of complex noun groups, and passage readability indices. As discussed in the literature review, vocabulary knowledge contributes to L2 reading comprehension (Bernhardt, 2011; Jeon & Yamashita, 2014), and many educators have prioritized vocabulary instruction as the remedy for high school ELLs who struggle with academic English. Participants in the present study cited vocabulary as a barrier to their comprehension of the scientific texts used in the interventions. Results from the lexical frequency analysis support teachers' and students' beliefs that vocabulary is the crux of ELLs' reading problems. Assuming that the 98% text coverage threshold for independent reading (Hu & Nation, 2000; Nation, 2006) holds true for high school science texts, then the LFP constructed from the two Alberta Education approved Science 10 textbooks (Figure 3) shows unequivocally that these two books require students to possess a wide breadth of vocabulary (over 25,000 word families) to reach the requisite 98% comprehension of the all the words in the text and to fully understand the text without aid. Similar to findings by Coxhead et al. (2010), it appears that the vocabulary requirement for independent reading of tenth grade high school science texts far exceeds the estimated 8,000 to 9,000 word families for novels and newspapers (Nation, 2006) and other expository writing rated at 10.0 on the Flesch-Kincaid grade level scale (Schmitt et al., 2011).

The wide breadth of rare vocabulary in the texts coincides with participants' feedback regarding the presence of many novel words contributing to reading difficulty. This observation is not surprising when viewed from a usage-based learning perspective. Students require many encounters with words in the mid- to low-frequency bands to gain at least receptive knowledge (form-meaning recognition) of them. Even when a more liberal 95% text coverage (e.g., Schmitt et al., 2011) estimate for instructional purposes is considered, reading and comprehending both science texts in their entirety requires knowledge of K-1 to K-16 bands. This amount of vocabulary knowledge is high for grade 10 NSs and ELLs alike, yet multiple encounters, possibly at least 20 (e.g., Pellicer-Sánchez & Schmitt, 2010), are necessary to learn the mid- to low-frequency and the higher frequency words; the textbook is one source of contextualized, meaningful exposure at the teachers' disposal.

Limited instructional class time prohibits focusing on the learning of all of the vocabulary necessary to reach 95%-98% coverage. The present study shows that for the two grade 10 science textbooks, a 95% threshold requires approximately 16,000 word families; learning all these words is untenable, even in a language course. However, as Schmitt et al. (2011) found, some degree of reading comprehension is still attained even with knowledge of only 90% of the words in a text. As the LFP analysis results showed, 90% coverage is reached starting at the K-3 frequency band, which is a more realistic target for ELLs in high school. The question remains, however, as to which words teachers and students should devote time in the classroom. The appeal of using principled word lists such as the AWL becomes apparent. Not only does the AWL have cross-subject utility and does knowledge of it appear to contribute to ELLs' reading comprehension (e.g., Townsend et al., 2012), but the AWL also covers a relatively high percentage of the words in the two science textbooks in comparison with any frequency band lower than K-3 (i.e., more infrequent). Also, whereas each frequency band encompasses 1,000 word families, the AWL contains only 570 words; vocabulary learning based on the AWL for the purpose of reading academic texts appears to have potential for a high return on investment. I will return to this point later in my discussion of the results from the experiment conducted as a part of the present study.

It is widely accepted that knowledge of vocabulary contributes to L2 reading comprehension in conjunction with other sources of knowledge, such as grammatical functions in the L2 (Bernhardt, 2011; Jeon & Yamashita, 2014). With the understanding that L2 grammatical knowledge functions as an important subcomponent, it is worthwhile to examine text characteristics beyond the word level that contribute to reading comprehension. In particular, authors of scientific writing have been found to prefer a compressed writing style that uses many complex noun phrases, resulting in high informational density per sentence (e.g., Biber & Gray, 2011; Halliday, 2004). The lexical density statistics found in my analysis of the science texts indicate a high proportion of content words to grammatical words, suggesting that the two grade 10 texts were composed with a compressed style writing. This interpretation is further corroborated by the examination of complex nominals: on average, the sentences in the texts contained more than one complex nominal per T-unit and per clause. One of the books, *Addison Wesley Science 10*, showed an increasing trend in the mean number of complex nominals as the chapters progressed, whereas *ScienceFocus 10* exhibited a more stable pattern throughout. Although it may be possible that different subject matter across the chapters affected the style of writing, that is unlikely to be the explanatory factor here, considering that both texts covered the same topics. The *Addison Wesley Science 10* authors (Sandner et al., 2004) may have expected the reading abilities of high school students to increase throughout the academic year, thereby justifying having written the later chapters in a manner that more closely resembles higher level scientific writing – such as the university undergraduate level materials examined by Biber and Gray (2010).

The lexis and syntax were found to be sophisticated and complex in the two grade 10 science textbooks, serving as evidence that the authors' writing styles were more representative of academic writing than conversational language. Analysis of readability showed that both books indeed ranked low in narrativity in comparison with other high school texts in the TASA corpus. As expected, science texts contain abstract subject-specific science concepts, hence their word concreteness tends to be low. The variability in word concreteness found in both books (lower in beginning, higher in later chapters) also relates to the topics of the chapters; the beginning chapters focused on chemistry and physics, whereas the latter chapters dealt with biology and ecology. The finding of high rankings in syntactic simplicity and cohesion is congruent with the observation made by McNamara et al. (2014) that science texts are relatively simple in syntax to support students' uptake of novel concepts in the absence of background knowledge for support. This relative simplicity does not mean, however, that the sentences are simple for ELLs to comprehend; complex nominals abound, as confirmed by the TAASC analysis and the researcher's ability to locate and use them for the Grammar intervention. As illustrated by SFL scholars (e.g., Fang & Schleppegrell, 2008), the use of complex nominals by textbook authors likely contributed to the high degree of cohesion following a thematic structure (Eggins, 2004), the use of words that carry higher levels of abstraction, and the compression of ideas into shorter sentences in the texts examined in the present study.

Results from the textbook analysis would appear to justify the deployment of linguistic interventions focusing on vocabulary as well as grammar to help ELLs improve their reading comprehension of content area material. Although the efficacy of vocabulary instruction has been well documented for ELLs and thus endorsed, quantitative evidence in support of grammar instruction for improving the reading comprehension of high school ELLs has been more difficult to locate. A pre- post-test experiment was conducted as part of the present study with the goal of filling this gap.

Research Question 2a:

What were the comparative effects of three online modules (traditional reading comprehension, AWL vocabulary, scientific grammar) on ELLs' passage-level reading comprehension as measured by a commercialized standardized exam?

The second research question focused on the comparative effects of the three intervention conditions (Traditional, Vocabulary, Grammar) on ELLs' passage-level reading comprehension. Two measures were used to assess reading comprehension performance: a commercialized reading assessment and maze passages generated from one of the two science textbooks. Pilot study results suggest that the type of intervention received had an effect on students' post-intervention achievement on the IPT3 exam. The students in the Traditional group made the greatest gains, followed by those in the Grammar group. The Vocabulary group, however, saw a decrease in mean scores. The mean gain scores on the IPT3 seem very modest, ranging from 2.55 to -0.64 points (of 55 total possible marks on the test). Only the difference in gains between the Traditional and Vocabulary groups were statistically significant. These differences may be attributed to a form of practice effect. The response format on the IPT3 is entirely multiple-choice. The activities in the Traditional intervention condition also consisted of multiple-choice questions. As a result, the students in the Traditional group had an additional 15 hours of practice reading texts and answering multiple-choice questions. More importantly, the Vocabulary and Grammar groups also improved over time, albeit not as much as the Traditional group. Although the gains were modest and may seem inconsequential, the IPT3 includes a set of norm-referenced categorical benchmarks (i.e., beginning, early intermediate, intermediate, early advanced, advanced). The terms used for the benchmarks carry connotations and may affect ELL program placements. For students on the cusp of a proficiency band, an increase of one mark may change their designation (e.g., intermediate instead of beginner, or advanced versus intermediate) and lead their teachers to view the ELL in a different light (i.e., more competent). In terms of practical significance, teachers' and schools' perceptions of their students' abilities may affect the ELLs' academic trajectories (Kanno & Kangas, 2014). Therefore, the value of the traditional practice of assigning students a reading, then having them answer multiple-choice comprehension questions for the purpose of summative exam preparation, cannot be ignored and should continue to be provided.

Research Questions 2b & 3:

2b) What were the comparative effects of three online modules (traditional reading comprehension, AWL vocabulary, scientific grammar) on ELLs' passage-level reading comprehension as measured by maze passages generated from a Science 10 textbook?

3) What are the comparative effects of three online modules (traditional reading comprehension practice, academic vocabulary practice, or grammar practice) on ELLs' passage-level reading fluency, as measured by performance on maze passages (accuracy in three minutes) constructed from a grade 10 science textbook?

Two versions of a maze test were also administered to concurrently assess ELLs' silent reading comprehension and fluency. The results from the aggregate data collected from both phases of the study will be the focus of the following discussion. In regard to the comparative effects of the interventions on the participants' reading fluency and comprehension of science texts, the Grammar group made the greatest gains on the alternate forms (Maze A and B) and the identical form (Maze C). However, only the differences between the Grammar and Vocabulary groups on Maze C were found to be statistically significant (alpha $< .05$). The lack of statistical significance could first be explained from an examination of the effect sizes from the post-hoc multiple comparisons (Table 6). When interpreted with Plonsky and Oswald's (2014) revised effect size benchmarks for SLA, only the effect sizes between the Grammar and Vocabulary groups on Maze C would be deemed as medium. Although the participant sample size in the present study is larger than in some L2 classroom-based experimental studies, three intervention groups led to smaller cell sizes for statistical analysis; as statistical significance is related to sample size, alpha level, and effect size, it may be that the group sizes and/or the observed effect sizes were simply not large enough to obtain significance at $p < .05$.

The fact remains, however, that, interpreted using Plonsky and Oswald's (2014) guidelines for within-group contrast, all three groups made statistically significant gains from the pre- and post-intervention administration of the maze tests, with medium to large effect sizes, even if some of the differences between groups were not statistically significant. Stated conservatively with respect to traditional null hypothesis significance testing, the comparatively novel Grammar intervention was no less effective than the more established methods. Recall that a true control group was excluded in favour of comparison groups, and that all groups received interventions that had the potential to help ELLs improve. In light of this study design choice, the finding of statistical significance holds promise for Grammar instruction in which grammar is treated as a meaning-making resource. In particular, the Vocabulary and Grammar groups received interventions on two of the biggest correlates of L2 reading comprehension; however, effect sizes between the interventions were small on the alternate form version of the maze and medium on the identical version of the maze, with the Grammar group experiencing greater gains than the Vocabulary group.

Several explanations exist for the patterns of observed effectiveness on the maze tests. The different size of effects derived from the two versions of the maze (alternate and identical) may partially be a result of the difficulty of the passages, testing effects, or a combination of both factors. All of the students performed better on the identical form (Maze C) post-test. As explained in the methodology chapter, I was unable to find two passages within the textbook that were identical in terms of topic and that shared the same scores on all indices of linguistic and readability. Maze C had a lower readability, lower noun phrase density, and higher AWL coverage than Maze A or B. Although effect sizes differed depending on the form of the maze (alternate or identical), the pattern remained consistent (Vocabulary < Traditional < Grammar), suggesting that the different versions of the instrument did not influence the comparative effects of the interventions.

Interestingly, even with AWL coverages on the maze tests that ranged from 13.1% to 16.5%, the Vocabulary group did not perform as well on these as the Traditional and Grammar groups. Although the Vocabulary group focused on AWL form-meaning mapping exercises, eight AWL words on the maze tests were not found in any of the passages used in the intervention phase. In other words, the maze contained novel AWL vocabulary, even for the participants in the Vocabulary group. Even if the Vocabulary group developed improved abilities in deducing meanings of words from context, this process does not appear to have been as successful as others in helping them achieve success in reading comprehension and fluency, as measured by a timed maze test. Also, as Eskildsen (2012) argued, the concurrent acquisition of vocabulary and grammar is possible, even if were an unintended pedagogical outcome. The other groups were, at the very least, exposed to the same words through reading the passages in the intervention and may have gained some AWL knowledge, thus narrowing the advantage of the Vocabulary group in terms of the AWL. While a form dictionary-meaning connection is important to vocabulary learning (Nation, 2001), the Grammar group may have even deepened their knowledge of some of the nominalized words through the process of unpacking complex noun phrases. For example, in an exercise deconstructing the sentence “the evaporation of water from leaves is called transpiration”, incidentally, the Grammar group increased their understanding of the word transpiration. The Grammar group saw the word used in context, learned a colloquial definition (“water is lost”), associated it as a specialized form of evaporation, and understood that the use of the word is constrained to certain locations (in this example, leaves). Consequently, the Grammar group’s ability to learn words from reading may also have improved from learning to unpack complex noun phrases.

Technical vocabulary is another area of academic language addressed by the interventions. The Traditional group, for example, completed form-meaning matching activities with technical vocabulary (i.e., glossary words), yet this group also exhibited fewer gains on the maze tests than the Grammar group but more than the Vocabulary group. Similarly to the Vocabulary group, the other groups received exposure to the technical vocabulary with the addition of online glosses; in other words, participants in all three groups received form and meaning exposure to the technical vocabulary. Two preliminary interpretations may be drawn from these results. First, having ELLs practise matching words to their definitions is relatively ineffective when compared with the other interventions. Second, knowledge of technical vocabulary appears to contribute more to ELLs’ reading comprehension of science texts than to their knowledge of AWL words. Technical words carry more information about a certain topic than does AWL vocabulary; knowledge of discipline-related words can be related to an increase of knowledge in a content area.

To be clear, the results of the present study in no way diminish the importance of vocabulary instruction; after all, the textbook analysis revealed the wide breadth of vocabulary required to comprehend them. Rather, the within-subjects comparisons and large effect sizes from the pre- and post-tests serve as evidence that the value of vocabulary instruction is consistent with previous research confirming the contributions of the AWL and technical vocabulary to success in secondary education contexts (e.g., Nagy & Townsend, 2012). The Vocabulary group participants may have improved in their word recognition speed, which indirectly facilitated their reading comprehension and fluency (e.g., Yamashita, 2013), thus explaining the improvement on the maze tests. Nevertheless, the results from the current study provide support for the claim that vocabulary instruction alone is insufficient in helping to improve ELLs' reading comprehension. It may be that learning only a portion of the AWL limited the Vocabulary group's gains; however, the amount of vocabulary introduced during the intervention period, averaging approximately 20 words per week, reflects a realistic classroom goal. With due respect to educators who devote instructional time to vocabulary instruction, the unfortunate current situation is that more needs to be done in the classroom to further support the development of ELLs' academic language (Wong Fillmore, 2014).

Vocabulary knowledge is only one of the factors that influence L2 reading comprehension (Bernhardt, 2011). The findings from the current study, which showed the vocabulary group as making the smallest gains on measures of reading comprehension, support Geva and Farnia's (2012) argument that word recognition alone may be inadequate to fully comprehend high school level texts. Whereas every participant was exposed to the AWL and technical vocabulary to some degree, the students in the Traditional and Grammar groups received other unique forms of interventions that may have contributed to the ELLs' improvements on the reading comprehension measures beyond vocabulary instruction alone.

The Traditional group completed comprehension questions that required them to read (and possibly re-read) the passages in each lesson. From the perspective of Carver's (1990) rauding theory, the participants were likely engaged in the rauding gear (reading and comprehending) and possibly the learning gear. The nature of the interventions forced the students to practice reading at a deeper level than mere skimming and scanning. According to skill acquisition theory (Anderson, 2007), the Traditional group's silent reading rate and accuracy may have improved more than those of the participants in the Vocabulary group as a result of practicing rauding. The multiple-choice questions used in the activities for the Traditional group were designed to be answerable by different levels of reading engagement. In every lesson, one or two answers could be found by skimming and scanning (so-called easy questions) to promote feelings of success, but the majority of the questions required careful reading and understanding (without inferencing) to arrive at the correct answer. The questions themselves were written as complete sentences, which provided additional reading practice. Also, as a result of rauding and answering comprehension questions, ELLs in the Traditional group may have increased their background knowledge of Science 10 concepts (as some participants indicated on the post-intervention questionnaire). The maze passages were essentially a subset of the Science 10 curriculum content; the increased performance in reading comprehension on the maze passages may have been aided by improved background knowledge, as suggested in Bernhardt's (2011) compensatory L2 reading model. The Vocabulary group, on the other hand, focused on the definition of the words. Judging from student comments, some AWL words in the passages were known prior to the experiment, and others were learned as the lessons progressed and some words were recycled. With known words, students did not have to reread the word in context to complete the exercise (i.e., match the word to its dictionary definition).

Complex noun phrases were extracted from the science texts for the Grammar group ELLs to analyze. The participants were guided in the unpacking of the complex noun phrases, which led to mapping the correct meaning to different parts of a noun phrase. Some of the students noted that the activity was confusing at first, and somewhat repetitive, but the repetition eventually resulted in perceptions of improvements in their knowledge of complex noun phrases, ability to successfully complete the activities, and reading comprehension of the scientific text. The results from the knowledge of nominal groups tests support the students' perceptions, as the Grammar group exhibited a larger effect size (0.51) than the Traditional (0.37) and the Vocabulary (0.03) groups. From the perspective of skill acquisition theory, students may have been reporting abilities that reflected the automatization of their abilities to unpack complex noun phrases. As shown in previous research (e.g., Jeon & Yamashita, 2014), improved grammatical knowledge should lead to enhanced L2 reading comprehension (although the present study may be one of the first to investigate the effects of L2 grammar instruction on high school ELLs' L2 reading comprehension). The Grammar group's performance on the maze tests indicated improvements in reading comprehension and fluency of scientific texts, modestly superior to the Traditional and Vocabulary groups' performance, as evidenced by the effect sizes. This difference in performance may be an attribute of improved efficiency in the participants' processing of complex noun phrases. In other words, the ELLs in the Grammar group were able to more quickly understand the information that was encoded and compressed in the complex nominals in the passages.

Improved efficiency in processing complex noun phrases is beneficial on two levels. First, faster processing should free up working memory for other processes required for reading. Second, improved processing of noun groups could have assisted the ELLs in maintaining coherence (Gernsbacher, 1991; Grabe, 2009) through greater capacity to utilize the textual characteristics of the science texts (relatively "simple" sentence structures, high cohesion; reiterative and zig-zag Theme structures [Eggins, 2004]). For example, the recognition of the Rheme (an idea) from a previous sentence restated as a complex nominal as the Theme of the next is useful. After all, cohesive texts should facilitate reading comprehension (McNamara et al., 2010), but only if ELLs can recognize overlapping referents that are often in the form of nominalizations and noun groups. It may be that the ELLs in the Grammar group were becoming more familiar with the sentence structures (complex noun phrases) commonly used in the science books as a result of repeated exposure from the intervention. Perhaps, as Ellis and Wulff (2015) suggested, helping ELLs make explicit form-meaning connections affected their implicit processing of the complex nominals that they encountered. Or it may be that the participants' abilities to unpack complex noun phrases were coming closer to being automatized due to practice. As evidenced by the results from the timed maze tests, their syntactic processing of the types of sentences used in the science textbooks became faster than that of the other two groups, who did not receive the same amount of processing practice.

Research Question 4:

What were the comparative effects of three online modules (traditional reading comprehension practice, academic vocabulary practice, or grammar practice) on ELLs' understanding of nominal structures as measured by a noun phrase interpretation activity?

Results from the knowledge of nominal groups test provided some substantiation for the Grammar group's greater understanding of complex nominals than that of the Traditional or Vocabulary groups'. The Grammar group experienced the greatest effect size from pre-test to post-test; furthermore, the 95% confidence interval of the Grammar group's gain score is the only one that does not incorporate zero. However, this finding needs to be considered alongside another observation. As shown in Figure 6, an unanticipated pattern occurred on the scatterplot of gain scores against pre-test scores. Considering that the identical form was used for pre- and post-testing, one would not expect negative gain scores. Yet, it appeared that for all three groups, the ELLs who scored the highest on the pre-test experienced negative gains on the post-test, and this effect was the most pronounced for the Grammar group (as exhibited by the slope of the fitted regression line). This finding may seem concerning to teachers, but a plausible, reassuring explanation for this exists. The untimed nature of the knowledge-of-nominals test meant participants had ample time to analyze the sentences. Similar to language learners who exhibit a U-shaped learning trajectory in which production accuracy decreases in spite of advancement in proficiency, the possibility exists that the advanced learners began to overanalyze the questions to the detriment of their performance.

A final observation from the experiment phase is the potential effect of student proficiency on the gains that participants experienced from the interventions. In Figure 7, scores from both versions of the maze tests (alternate and repeated forms) were combined into a single set of maze scores (pre-test, post-test, gain scores), and the combined pre-tests scores were plotted against the combined gain scores. Two distinct patterns emerged. The initially high performers in the Traditional and Grammar groups appeared to have benefitted the most from the interventions, as they achieved the greatest gains. This pattern seems slightly more exaggerated for the Grammar group than for the Traditional group. Conversely, in the Vocabulary group, the students who scored the lowest on the pre-test made greater gains than those with high pre-test scores; however, this effect appears much less pronounced than for the other two intervention conditions.

The patterns described above suggest that the degree of effectiveness of the interventions may be partially dependent on the ELLs' reading proficiency levels. With the exception of a short introduction video, the interventions were entirely text-based (notwithstanding pictorial figures that were part of the textbook). High school students are expected to learn through reading (Chall, 1990); the same expectation was held for the ELL participants who were learning to read by reading to learn. In light of Hulstjin's (2015) BLC Theory, one could argue that NS students could benefit from the same practice with reading and uncommon lexicogrammar (HLC) if they wished to achieve automatized control of reading academic language. With regard to the ELL interventions, although the online modules provided scaffolds, learning was done independently. It may be that the text-based, independent-learning nature of the computerized interventions meant that the less proficient students did not benefit as much because of their lower reading skills. This phenomenon has been reified as the Matthew Effect (Stanovich, 1986), according to which proficient readers gain more from education and continue to outperform less proficient readers.

Perhaps the Matthew Effect was to be expected from having ELLs simply read texts and then answer questions (the traditional intervention condition). Lower proficiency students likely did not possess sufficient lower level reading skills to fully cope with the demands of the science texts. For these students, vocabulary instruction with a focus on increasing the breadth of the form-meaning aspect of word knowledge seemed the most effective. This finding is unsurprising, as it is consistent with the established notion that L2 vocabulary knowledge contributes to reading comprehension for K-12 ELLs (August & Shanahan, 2006). A previously unexplored finding, however, suggests that Grammar instruction is more worthwhile for lower proficiency readers than the activities used in the Traditional condition. The linguistic instruction and practice with unpacking noun phrases likely provided lower proficiency students with an advantage over their peers in the Traditional group who received no such linguistic support. However, higher proficiency students may be better equipped in their developmental stages for interventions focusing on syntax. As Geva and Farnia (2012) found, only when ELLs become older (equated with higher language proficiency) did the effect of syntactic knowledge become apparent. In the Vocabulary group, the higher proficiency students may also have encountered a ceiling effect with the AWL, hence the somewhat stable pattern (as opposed to any sharp increases or decreases). ELLs in the Traditional group, having received no linguistic support, appear to better benefit from traditional instruction only after a certain proficiency level has been achieved.

Research Questions 5 & 6:

5) What were the ELLs' attitudes (positive or negative) toward the use of the modules in their classes?

6) What are the ELLs' conceptions of the role of the online modules in their learning?

The final two research questions concerned the students' attitudes towards and conceptions of the interventions. The computerized interventions were well received by the participants and regarded as having made contributions to the improvement of their reading skills. In short, the students enjoyed the online modules. The favourable comments shared by the students are consistent with previous CALL research (Macaro et al., 2012) that found students to have positive attitudes towards computerized instruction. Students were observed to be at ease with the online nature of the interventions, perhaps due to the ubiquity of computers and the Internet in students' lives and the tight integration of technology into regular classroom routines. This observation is corroborated by the students' perception of the website as easy to use as long as one has basic computing skills. Participant 13, in particular, believed that "everything was simple [and] very user friendly [and that] even my grandma can use it" (with the assumption that Participant 13's grandmother possessed minimal computing skills). In addition to comfort with computer technology, students' positive reception to the modules may be credited to factors related to affordances provided by technology and to the value that participants placed on the contribution of the module to their learning.

Reading practice with technology in this study provided two main advantages over paper-based reading that likely promoted students' positive attitudes. First, participants attributed reading difficulty to unknown vocabulary; this issue was partially addressed through in-line glossing of the technical science vocabulary, which the participants appreciated. Reading on paper would have required students to find glossed words elsewhere, usually in the back of a textbook, and lead to greater interruptions in the reading process. In-line glossing minimized the interruptions and was less laborious, as the students could "just hover over the words" (Participant 24). Second, the computer as a tutor model (Warschauer & Healey, 1998) was recognized by participants as helpful. They valued the interactivity and availability of immediate feedback (correct/incorrect) while completing the activities. I observed that rather than simply giving up when they were wrong, the students paused to re-process the questions and tried to understand why an answer was correct or incorrect. They were also motivated to reattempt the entire set of questions, striving for the computer feedback stating "You got 100%", even though it was made clear to the students that performance during the experiment had no effect on their academic standings. Such positive attitudes towards to the activities may also be related to a perceived value in learning.

Participants believed that the practice gained from the online modules contributed to improvements in their reading ability. In addition to perceiving an immediate return on their time/effort investment, participants believed that module completion would be of benefit to them when they took science courses in the future. In fact, participants requested readings on other subjects because they thought the extra practice would help them comprehend the academic English used in other content area classes. As a result, these conceptions regarding the online modules facilitated a positive feedback loop for the students: the more they practiced, the better they felt they could read, and the more enjoyment they received through practice.

Discussion of Contributions to the Research Community

L1 and L2 Reading

This dissertation study has been situated within a cognitive, skill-based perspective of reading. As such, importance has been placed on meaningful practice as a means to help learners develop the automatized execution of reading skills. L2 reading is understood as a complex, interactive process involving a myriad of factors including L1 literacy, L2 linguistic knowledge, and individual differences (Bernhardt, 2011). The contribution of the separate components to L2 reading comprehension have been previously studied, with L2 grammar knowledge and L2 vocabulary knowledge emerging as two of the largest correlates (Jeon & Yamashita, 2014). The results from the reading comprehension measures in this study showed differential effects from instruction (Traditional, Vocabulary, Grammar) that targeted different components of L2 reading. In other words, it appears that different sources of knowledge contributed to the participants' L2 reading comprehension performance. Congruent with results from the meta-analysis by Jeon and Yamashita (2014), findings from the current study suggest that L2 passage level reading comprehension may be more strongly related to L2 grammatical knowledge than L2 vocabulary knowledge, although the answer to the question of precisely to what degree was beyond the scope of this dissertation project. Also, as posited in Bernhardt's (2011) L2 reading model, L2 linguistic knowledge alone did not explain 100% of the observed variance in reading comprehension performance; the Traditional group may have demonstrated greater gains than the Vocabulary group, potentially as a result of growth in areas tangential to L2 language knowledge (i.e., comprehension strategy use, reading fluency).

Although the relationship between L2 grammar knowledge and reading comprehension has been established (e.g., Jeon & Yamashita, 2014) and it appears that L2 syntactic knowledge may affect L2 reading as a function of sentence processing time (van den Bosch et al., 2018), the mechanism through which grammatical knowledge contributes to L2 reading comprehension has not been well understood or explored. From the L1 reading research literature, grammar knowledge has been proposed as aiding the maintenance of coherence (Gernsbacher, 1997) or affecting processing time (Fender, 2001). The Grammar group in this study exhibited the greatest gains on reading comprehension and fluency assessments constructed from science texts that were high in cohesion, as is typical with high school science texts (McNamara et al., 2014). These results support Gernsbacher's view of the function of grammar in L1 reading, as discussed above. Alternatively, it may be possible that, through repeated practice with processing complex noun phrases, the Grammar group became more efficient at reading them, leading to more fluent reading, which also contributes to reading comprehension (e.g., Grabe, 2009). Unfortunately, to my knowledge, how L2 readers process sentences containing complex nominals has yet to be studied in depth. However, grammatical knowledge matched to grammatical features that are prevalent in a text appears to facilitate reading comprehension.

The utilization of grammatical knowledge for reading was evidently more effective for participants who were better readers at the onset of the study (see Figure 7), echoing findings by Geva and Farnia (2012). These results support the argument of a skills hierarchy in L2 reading that places prominence on lower-level skills such as word recognition (Nassaji, 2014). That is, participants with lower pre-test reading scores likely possessed weaker lower-level reading skills that hindered the execution of other processes such as sentence parsing. Although compensation may occur (as explained previously in my discussion of Bernhardt's [2011] L2 reading model), cognitive resources (e.g., working memory) would have been taxed, as execution of lower-level process was inefficient and became laborious, resulting in decreased reading comprehension and a slower rate of reading fluency. Also, lower-level skills help readers derive intended meaning from the text and avoid overreliance on the situational model (background knowledge) that may lead to erroneous interpretations (Kintsch, 1998; van den Broek & Kendeou, 2008). Comments from the participants repeatedly suggested that their attention was focused on vocabulary; in other words, the lack of word recognition hindered their reading comprehension, fostering a sense of reading difficulty. While it may seem intuitive that insufficient vocabulary debilitates reading comprehension (e.g., Schmitt et al., 2011), the meanings of some of the words in the texts used in the interventions, particularly the technical science vocabulary, were explained within the passages or understandable in context. In other words, learning from reading (Chall, 1996) was required, which was a goal of this study. This observation suggests that not only is vocabulary knowledge important, but rapid, automatized recognition of a wide breadth of vocabulary is also vital to cope with the academic demands of reading in high school.

L2 reading problems stem from issues with language proficiency or reading processes (Alderson, 1984), and I argue that the two are inextricably bound. This study has shown that teaching ELLs to parse sentences (a reading process) containing complex noun phrases (a linguistic feature) has the potential to improve reading comprehension beyond instruction with (Vocabulary Group) and without (Traditional Group) a language focus. L2 reading comprehension requires L2 language comprehension, as shown by researchers (Gottardo & Mueller, 2009; Verhoeven & van Leeuwe, 2012; Yamashita & Shiotsu, 2017) working within the Simple View of Reading (Hoover & Gough, 1990). Extending the finding that grammar instruction is a worthwhile pursuit for ELLs, there is reason to believe that it would also be effective for L1 English students, as well. It is recognized that academic English differs from conversational English (Cummins, 2008; DiCerbo et al., 2014) and that native-speakers and non-native speakers alike can continue to develop higher language cognition (i.e., academic English) (Hulstijn, 2015). If academic English is indeed a specialized, often unfamiliar form of language with yet to be learned lexicogrammatical features (e.g., prevalent use of complex nominals in scientific discourse), then high school students (NSs and NNSs) who experience reading problems with academic writing may benefit from both reading and language-based interventions.

SLA

This dissertation was grounded in a number of theories related to applied linguistics, namely systemic functional linguistics (SFL), usage-based learning, and skill acquisition theory. The overall study results support SFL as a viable foundation on which to build L2 pedagogical practices, which I discuss in a later section. According to usage-based learning, exposure to language is required for learning, and I heeded the call from educators who advocated for ELLs to receive instruction in and exposure to academic discourse (Harper & de Jong, 2004; Wong Fillmore, 2014). The significance of exposure lies at the centre of corpus-based vocabulary studies reviewed in Chapter Two, and words in the mid- to low-frequency vocabulary bands are assumed to be difficult for ELLs. The Lexical Frequency Profile results from the textbook analysis, coupled with participants' statements that attest to their limited vocabulary knowledge, support the usage-based notion that vocabulary learning (form-meaning mappings), and, by extension, language learning, depend on frequency of input (e.g., Ellis & Wulff, 2015). Therefore, the idea of the reciprocal relationship between vocabulary and reading (Nation, 2001; Stanovich, 2000) is also upheld. Learners' ability to engage in "statistical abstraction of patterns latent within and across form and function in language usage" (Ellis & Wulff, 2015, p. 89) requires exposure to language forms; yet, for example, conversational language contains far fewer instances of nominalizations and complex noun phrases than written language (Biber & Gray, 2011). As observed in the present study, ELLs' comprehension of academic English may be enhanced from the increased frequency of exposure provided by the reading of academic texts and form-focused follow-up activities.

To further discuss potential contributions of this study to SLA theories, I turn to the concept of lexicogrammar (Halliday, 1997/2004b), which formed the basis of the understanding of vocabulary and grammar in this research. The boundary between vocabulary and grammar in the interventions may appear somewhat blurry, reflective of the nature of language. For example, Kremmel et al. (2017) found that foreign language readers parse multiword units (i.e., formulaic phrases) as chunks; this observation can be interpreted to indicate that boundaries between words, phrases, and sentences are not clear-cut. Although formulaic phrases were not a part of this investigation, and no particular academic phrases were explicitly taught, it could be argued that the structure of a complex noun phrase (Biber et al., 1999) is formulaic, particularly in case of academic writing, where writers favour the use of "nouns as nominal premodifiers and prepositional phrases as nominal post-modifiers" (Biber & Gray, 2011, p. 223). Through repeated practice with unpacking complex noun phrases, it may be possible that the Grammar group more readily recognized the constituents of complex noun phrases and parsed them as chunks. This behaviour would also explain the greater growth in reading fluency experienced by the Grammar group, as discussed above. This growth may be attributable to two processes, both stemming from explicit practice over the intervention period. First, in support of usage-based learning, explicit attention drawn towards complex noun phrases could have facilitated implicit learning during subsequent encounters, as learners subconsciously "update the statistical tallying" (Ellis & Wulff, 2015, p. 89) of the form-meaning connection of complex noun phrases. Alternatively, as proposed by ACT-R (Anderson, 2007), explicit practice with unpacking complex noun phrases led to proceduralization of the process; in turn, the speed with which this happens would appear similar to chunking. Therefore, meaningful practice with unpacking complex noun phrases has the potential to result in automatization.

The seemingly paradoxical observation, where the participants felt the interventions were "a little bit boring" yet were motivated to continue, supports the notions of a complementary relationship between investment and motivation and reflects the power of imagined communities (e.g., Norton & Toohey, 2011). In informal conversations, the participants conveyed a desire to enrol in higher level academic courses. Students possessed ambitions to become members of an imagined community of higher learning; thus, they were invested in and motivated to learn the requisite academic language necessary for entry into the community, as demonstrated by their persistence during the entire intervention period. However, as Kanno (2014) found, students must display a certain level of academic performance for that to happen; teachers' perceptions of students' potential could also affect future academic trajectories. One of the reasons I was welcomed into the schools was that the teachers did not conceptualize an artificial ceiling for their students; they, too, were invested in language learning and determined to do all that was within their power to help their pupils reach their academic goals.

CALL as a Research Platform

The decision to deliver the interventions via an online platform satisfied not only pedagogical requirements but also my research needs. From a research perspective, using the computer as a tutor (Taylor, 1980) controlled for potential confounding teacher effects that may have resulted from human delivered interventions; furthermore, it facilitated the classroom-based experimental design with random assignment, which has been found to be relatively rare in L2 research (Plonsky, 2013). I also believe that I have addressed the criticisms of past CALL studies (Grgurović et al., 2013; Macaro et al., 2012) by applying methodological rigor, fully reporting pertinent data and statistics, and, most importantly, connecting the use of technology to SLA theory.

Pedagogical Implications

A goal of this study was to identify ways to help ELLs improve their academic standings. In the interest of equitable education and opportunities, the situation in which ELLs continue to avoid language-heavy courses (Toohey & Derwing, 2008) is unacceptable. Often these courses are "provincially examinable courses [without which] students…may have limited opportunities for post-secondary education" (Toohey, 2007, p. 268). Not only is science language-heavy, but results from the textbook analysis confirmed that the language used in the science texts indeed differs from oral language (e.g., Cummins, 2008; DiCerbo et al., 2014), as indicated by the low ranking in narrativity. ELLs (and NSs alike) need to learn to understand academic language, but, with the exception of ESL classes, the only course in K-12 schools with an instructional focus on the English language is English Language Arts (ELA). However, it would be unreasonable to expect students to become proficient readers of scientific writing from enrollment in ELA courses where the reading of literature (which ranks high in narrativity) is emphasized above other text genres. Instructional practice with reading forms of writing other than stories should be included as part of K-12 education. In Alberta, some school jurisdictions have begun to recognize this need, as evidenced by the existence of ESL Expository English 15 and 25; however, this type course is neither universally offered nor required, and, to the best of my knowledge, no plans nor resources have been devoted to the creation of a provincial Program of Studies (curriculum). Therefore, the preparation of ELLs to read and understand academic writing across subject areas does not always occur in ESL and ELA courses, and sole reliance on ESL and ELA teachers for this task is ill-advised; teaching ELLs to read needs to be a shared responsibility. This study has provided evidence to support targeted linguistic instruction that all teachers of ELLs can replicate (and Science 10 teachers can readily adopt) to fill this gap.

To be clear, my intent has never been to take a reductionist approach and to identify and promote “a best way for teaching ELLs to read” (i.e., spend all L2 instructional time on grammar). Rather, my study provides additional evidence to justify the inclusion of L2 grammar instruction and repeated practice as part of language instruction for high school ELLs.

As discussed above, the patterns shown in Figure 7 suggest that ELLs at varying levels of proficiencies may benefit differently from diverse forms of pedagogical activities (the Matthew Effect). For lower proficiency students, who likely have insufficiently developed lower level reading skills, the return on investment appears to be greater for vocabulary instruction.

Providing some linguistic support in the area of grammar may also be beneficial for these students. At the same time, improving one's reading skills requires one to read; every high school ELL needs to practice passage-level reading. Research in L1 reading shows that even struggling readers can improve from extensive reading (Kuhn & Stahl, 2003); the present study provides support that practice in reading academic material can help, but that the beneficial effect can be augmented with additional meaning-focused linguistic activities.

Although each of the intervention conditions individually led to within-group improvements, providing a combination of all three in a reading course or content-area classroom could lead to improved reading performance on multiple-choice global reading exams and curriculum-based measurement (CBM) maze assessments. School institutional practices often include outcomes from both types of assessments in decision-making processes (i.e., placement or additional interventions) regarding ELLs. A lexicogrammar perspective of language underlies this study; the intervention conditions were isolated in this study only to study their comparative effects. Having established that grammar instruction is also viable and beneficial to L2 reading comprehension, effective reading instruction for ELLs should include vocabulary and grammar, in addition to the traditional instructional practice of reading and answering comprehension questions. Much has been written on the topic of L2 vocabulary instruction, and educators teaching ELLs would be wise to incorporate techniques beyond the form-meaning matching activity used in this study. The results from this experiment serve as evidence to support devoting classroom time to SFL-based grammar instruction. It is conceivable that this type of instruction may be even more effective in a face-to-face environment in which a teacher could clarify how to unpack complex noun phrases.

Based on the results of the textbook analysis, teachers should use caution when interpreting readability indexes captured as a single number for an entire book. Although a text book may be rated as a certain grade level, individual chapters or even passages may not necessarily correspond to the given rating. Furthermore, a unidimensional readability measure (often the Flesch-Kincaid grade equivalent), while useful, may be even more informative when interpreted in conjunction with other indices of readability (McNamara et al., 2014).

With the understanding that infrequent vocabulary and complex noun phrases pose barriers to reading for ELLs, teachers can pre-identify potential linguistic problems in readings and incorporate the teaching of them into their lesson plans. Automatic text analysis tools such as the Vocabprofile module on Cobb's (n.d.) Lextutor (https://www.lextutor.ca) can help teachers quickly identify the frequency of the vocabulary in a text. In addition, the Concordance module on Lextutor can index a user's own text and identify the collocates; Coxhead et al. (2010) suggested the use of this function for teachers to find the words that co-occur with the vocabulary to be learned so that they can be better prepared to help students fully understand the words and how they are used. Teachers are also encouraged to use the webtool provided by the Coh-Metrix team (http://www.commoncoretera.com) to analyze texts that they plan to use with ELLs. The tool returns clear bar graphs of the five readability indices, along with an explanation and interpretation of the indices. For teachers who may not yet have the linguistic knowledge or confidence to identify and unpack the linguistic structure of complex nominals themselves, the automated natural language processing tool (http://nlp.stanford.edu:8080/parser/) or other parsers with web interface (e.g., http://tomato.banatao.berkeley.edu:8080/parser/parser.html) are easily accessible and can help. These tools could also be used to analyze students' production of academic language in content-area subjects, and could “inform [future] pedagogy and curriculum design” (Norton & De Costa, 2018, p. 105).

Similar to findings from previous studies (e.g., Cho & McDonnough, 2009; DelliCarpini & Alonso, 2014), the teachers with whom I consulted and worked in this study had limited proficiency with metalanguage. This gap in knowledge is problematic, in the sense that it limits the range of L2 language-focused pedagogical activities that can be orchestrated in the classroom. Specifically, metalanguage is an integral part of L2 pedagogy literature grounded in SFL (e.g., Palincsar & Schleppegrell, 2014); it is the means through which teachers and students can effectively discuss language and its role in meaning-making. ESL teacher training/professional development could benefit from the inclusion of linguistics components based on SFL, which have been reported to be positively received by teachers (Achugar et al., 2007; Gebhard et al., 2013). With an increased linguistic understanding of academic English, teachers can create pedagogical activities similar to those developed in this study for use with other science courses. These activities may even accommodate other subjects, although the effectiveness of the grammar activities focusing on complex nominals may be contingent on the textual characteristics of the texts that students are expected to read in those classes (for ideas on incorporating functional grammar in education settings, see de Oliveira & Schleppegrell, 2015; Fang & Schleppegrell, 2008; Jones & Lock, 2011).

On the surface, engaging students in repetitive practice (i.e., drills) might be considered an outdated practice; however, as argued here, meaningful, contextualized repetitive practice appears to be warranted. It potentially contributes to vocabulary retention (Willis & Ohashi, 2012) and, as demonstrated in this study, growth in reading comprehension and fluency. However, while students stated that the activities became easier with repetition, it led some to describe the interventions as boring (although they were still motivated by their sense of progress). Therefore, for teachers who incorporate repetitive practices in their lesson plans, a preamble and constant reminders of the benefits of practice, as well as providing students with evidence of growth, may be useful in maintaining student motivation. It is also possible that the medium used for practice (online) played a role, as discussed in the following section.

CALL as a Tutor

Although the popularity in language learning and in CALL has shifted away from a mechanical tutor concept (Warschauer & Healey, 1998), the positive reception demonstrated by the participants in this dissertation study, along with the reading comprehension and fluency results from the experiment, suggest that some merit exists in a mechanical tutor model. Students generally appreciated the repetitive nature of the exercises and perceived their progress (which they self-monitored through immediate feedback during every session) as the result of practice. Unlike the mechanical drills of the behaviourist era, however, the interventions created for this study were contextualized, meaningful, and relevant to the participants' current and future academic needs; i.e., the modules aligned with curricular goals and used genuine content course material. Whether the practice was conducted in the Traditional, Vocabulary, or Grammar condition, each intervention focused on the reading of scientific texts that students will be expected to understand in the near future if they choose to pursue science courses beyond the grade 10 level. The skills practiced in each condition also have cross-curricular value, in that reading is a key component in academic subjects. From a skill acquisition theory (Anderson, 2007; DeKeyser, 2007) perspective, these skills are most efficiently developed through repetitive, meaningful practice, as was the case in this study.

The use of computer technology provided advantages over traditional, paper-based instruction. Students were not required to wait for teachers to grade their worksheets; the tutor-like online modules were able to provide all of the participants with immediate, timely feedback on their performance, which was greatly appreciated by the students and teachers alike. Also, the ability to quickly access definitions for glossary words saved the participants time and led to minimal interruption in their reading processes. The students also expressed their preference for the interactive nature of the activities, perhaps reflective of their comfort with computer user interfaces. It appears, however, that the allure of computer-based activities still remains for many of these technology-savvy participants, who described them as “fun”, and this provided some justification for devoting resources to their creation. Moreover, practice with reading on a computer screen for school purposes could familiarize and prepare students for high stakes exams, such as Alberta Education’s Quest A+ system (https://questaplus.alberta.ca), that are increasingly administered online.

The computer technology used in the present study to create the interventions can also be used by educators. During the planning stage, I purposely searched for online content creation tools that were free, easy to use, required no coding knowledge, and produced interactive activities that were accessible with all modern computing devices. HTML5 technology met the criteria and was beginning to gain popularity at the time, while support for the more popular Flash platform (used for Alberta Education’s Quest A+), was quickly declining (coincidentally, MacWhinney [2017] would later recommend the use of HTML5 for SLA research). The platform used, H5P (https://h5p.org), is recommended for its ease of use and integration with popular content/learning management systems such as Moodle. With respect to classroom teachers’ limited lesson preparation time, I recommend that schools, school boards, or even government departments in charge of education devote resources to the creation of online modules similar to the ones created for this study. Unlike generic commercialized ESL teaching materials, these tailor-made solutions would then have a direct relationship with curricula and programs of study, using “real” text that students in their jurisdiction read in content-area classes. These materials could be deployed as supplementary materials. Moreover, once such a CALL-based intervention is developed, it would cost little to maintain (electricity consumption and Internet bandwidth), would be easily accessible, and would remain relevant until the curriculum changed. Teachers could then offer assistance to students while they are engaging with the online modules, providing an added level of scaffolding (similar to the present study).

Alternatively, teachers could involve their students in the production of content for future instructional use. For example, students in groups could create think-aloud tutorial videos on unpacking noun phrases that teachers could then use as a teaching aid for other cohorts of students. Additional benefits from such an activity would include insights into ELLs’ mental comprehension process, additional interaction opportunities, and student empowerment. Toohey et al. (2015) reported that, through the process of digital videomaking, ELLs engaged with their NS peers “in a variety of multimodal literacy activities in ways that did not privilege native speakers” (p. 481).

Despite the affordances provided by technology, drawbacks also exist. For example, some students struggled with correctly inputting their usernames and passwords during the first week. As I was on site for support and troubleshooting, I noticed that the problem was attributable to keyboard entry errors. Understandably, some students are likely accustomed to using a single set of login credentials for multiple websites or relying on biometric login technology. Nonetheless, the amount of time required for logging in would be better devoted to instruction. Secondly, based on my experience in one of the participating schools, computer labs serve the computing needs of an entire school, which creates logistic hurdles (i.e., booking availability). I was fortunate that the participating teachers were able to negotiate the use of the labs for the study. For the purpose of control, participants used the online interventions on school computers only. In future applications, however, teachers may wish to consider permitting students to use their own personal computing devices (e.g., smart cellular phones or tablets), as HTML5 functions on these devices, students are familiar with them, login credentials could be saved on them, and they would likely be the devices that students use for additional practice outside of the classroom. Some participants explicitly stated intent and willingness to use the modules on their own. Online learning programmers should anticipate this possibility and make improvements to the system, heeding the request of participants in the current study for the inclusion of options for users to customize the visual interface, for example, a “night-mode” that inverts colours to white text against a black background.

Limitations and Future Directions

Despite the positive reception of the interventions by the teachers and students, and the positive outcomes of the experiment, a number of limitations exist. First, it is arguable whether or not the participant sample was representative of the population, which may limit the generalizations that can be made. Participation by most of the ELLs in the schools, with the exception of those who had previously taken Science 10, lends greater generalizability to the findings presented in this dissertation. Although the participants’ backgrounds were diverse, Filipino comprised the majority of students’ ethnicity. This composition is reflective of the ELL population in this school jurisdiction, but other Canadian ESL classrooms might not be dominated by members of this particular group. Other researchers (e.g., Odo, 2012) question the utility of the ELL label because it represents such diversity in terms of first language and culture; however, in Canada, ESL classrooms are typically as diverse as our immigrant population, so controlling for L1 may not always be possible for researchers who conduct studies in Canadian classrooms. At the time of data collection, the composition of the ELL population in both high schools was dominated by one ethnic group. According to the teachers, the other high schools within the district were experiencing a similar trend. Therefore, I can claim that the participant sample was representative of current high school ESL classrooms only in the one school jurisdiction in which the data were collected.

A major limitation of the study was the restricted time that I was allotted to collect data in the ESL classes. Because of this restriction, I was unable to administer additional assessments (e.g., a standardized reading assessment with ELL norms in the second phase, other methods of reading comprehension assessment such as retelling, a vocabulary knowledge measure, a science knowledge measure, a working memory measure) that would have added further insight and clarified some of the current findings. The teachers welcomed me into their classes, but they did so with caution. Although they were receptive to exposing their students to interventions that had the potential of helping the ELLs, losing class time to testing and exposing students to the stress of testing was acceptable only to the extent of what was negotiated and done in this study.

The knowledge of nominal groups test created for this study showed merely an acceptable level of reliability, and the findings from this measure should be interpreted with caution. The creation of valid and reliable assessment could constitute an entire study unto itself, but, for the purpose of this research, it provided an indication of how well the participants could interpret the meaning of complex noun phrases. In future, such a test could include other methods of gauging ELLs' degree of comprehension of complex noun phrases and other linguistic structures (e.g., think-aloud protocols). Furthermore, investigations examining ELLs' processing of complex noun phrases and their ability to unpack them could contribute to the refinement of grammar and language teaching pedagogy.

The use of a comparison group rather than a control group was a decision based on my own ethical principles. Knowing the importance of reading, I did not want to assign one group of ELLs to a control group with no reading practice. However, this decision resulted in what might be considered an uninteresting result under the NHST paradigm. Finding *p*-values below .5 in statistical analyses of group differences becomes much more difficult when all of the intervention/comparison groups are effective.

In future studies, I propose the adoption of a lexicogrammar approach to vocabulary and grammar in L2 pedagogy research. For example, instruction beyond the AWL in the form of the academic formulas list (Simpson-Vlach & Ellis, 2010) may further support ELLs' comprehension and understanding of how meaning is created through language use. It would also be useful to examine the combined effects of all three forms of intervention on ELLs' reading comprehension compared with a true control. In my study, reading practice was structured as an individual activity; however, incorporating the linguistically focused activities into a collaborative learning environment may also yield fruitful results. As I argued above, the Grammar intervention may also benefit NS high school readers, and, based on the current findings, an investigation with NSs may be warranted. Also, as ELLs and NSs may employ different patterns of noun phrase usage in their writing (Ai & Lu, 2013; Parkinson & Musgrave, 2014), an examination of the effects of Grammar instruction in which ELLs and NSs are grouped together is also appealing to explore how they co-construct an understanding of complex noun phrases. Finally, it would be worthwhile to conduct a similar experiment using texts from other content-area subjects to determine the effects on reading comprehension performance.

As with other studies, increasing the number of participants would increase the power and generalizability of the results. The administration of the additional assessments identified above would improve the quality of the study by providing more nuanced understandings of ELLs' reading comprehension and depth of knowledge in the lexicogrammar features. However, such classroom-based intervention research would likely require the extensive cooperation of classroom teachers. Research often is an intrusion into regular classroom routines; therefore, researchers need to be sensitive and communicate clearly to teachers that the negative effects of interruption could be offset by benefits to the students. Designing interventions that complement curricular goals or address real classroom needs is one avenue to facilitating researchers' entry into classroom. Throughout the data collection phase, the teachers expressed their appreciation for this study. They wanted interventions for their ELLs that targeted academic English, and upon learning of my study, they recognized its potential in fulfilling a classroom need. At our initial meetings, a few of the teachers remarked, "This [type of intervention] is what our students need." Their appreciation grew as the study progressed and when they saw that students were benefitting from their participation. The next step is to persuade teachers that testing could result in the positive aspect of washback (the improvement of students' learning), and that assessment in classroom-based research is for the purpose of improving instruction, ultimately enhancing students' achievement of learning outcomes. As a result of conducting this research, I believe that I have begun to build the trust of classroom teachers and that this trust will grow and allow me to continue developing and researching effective L2 pedagogical practices.

Conclusion

The ELLs who participated in this study were randomly assigned to one of three different forms of reading intervention based on either traditional pedagogy, the learning of AWL words, or the unpacking of complex noun phrases. In a pre-test, intervention, post-test experiment, participants completed 15 hours of reading practice and follow-up activities. Regardless of the intervention group, the students reported enjoyment and a sense of improvement in their reading proficiency and increased knowledge of science from using the online modules. Within-group comparisons showed statistically significant improvements on measures of noun group knowledge, reading comprehension, and reading fluency. Between-group comparisons from pre- to post-testing indicated that the Grammar group experienced the greatest gains, although not all of the gains were statistically significantly different from those in the other two groups (Traditional and Vocabulary). Despite the lack of statistical significance in some of the between-group comparisons, the current study provides classroom-based experimental evidence to support arguments made by proponents of SFL-based ESL pedagogy (e.g., Fang & Schleppegrell, 2008; Gebhard, 2010; Mohan & Slater, 2006): Grammar instruction and practice that focus on grammar as a meaning-making resource can be an effective method for helping ELLs better understand the academic language of science.

References

Abbott, M. L. (2006). ESL reading strategies: Differences in Arabic and Mandarin speaker test performance. *Language Learning, 56*(4), 633-670. https://doi.org/10.1111/j.1467-9922.2006.00391.x

Abedi, J. (2006). Psychometric issues in the ELL assessment and special education eligibility. *Teachers College Record, 108*(11), 2282-2303. https://doi.org/10.1111/j.1467-9620.2006.00782.x

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219-234. https://doi.org/10.1207/ S15324818AME1403_2

Abu-Rabia, S. (1996). Druze minority students learning Hebrew in Israel: The relationship of attitudes, cultural background, and interest of material to reading comprehension in a second language. *Journal of Multilingual & Multicultural Development, 17*(6), 415-426. https://doi.org/10.1080/01434639608666293

Abu-Rabia, S., & Sanitsky, E. (2010). Advantages of bilinguals over monolinguals in learning a third language. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education, 33*(2), 173-199. https://doi.org/10.1080/15235882.2010.502797

Achugar, M., Schleppegrell, M., & Oteíza, T. (2007). Engaging teachers in language analysis: A functional linguistics approach to reflective literacy. *English Teacher: Practice and Critique, 6*(2), 8-24. Retrieved from https://files.eric.ed.gov/fulltext/EJ832185.pdf

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Adams, M. J. (1999). The science and politics of beginning reading practices. In J. Oakhill, R. Beard, & D. Vincent (Eds.), *Reading development and the teaching of reading: A psychological perspective* (pp. 213-217). London, England: Blackwell.

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier, & P. Thompson (Eds.), Automatic treatment and analysis of learner corpus data (pp. 249-264). Philadelphia, PA: John Benjamins.

Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning, 53*(2), 207-231. https://doi.org/10.1111/1467-9922.00216

Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics, 29*(1), 175-193. https://doi.org/10.1017/S0142716408080089

Alberta Education. (2009). *High school completion longitudinal study.* Edmonton, Canada: Accountability and Reporting Division. Retrieved from http://education.alberta.ca/media/1079006/hscdataanalysis_report_%20final.pdf

Alberta Education. (2013a). *Ministerial order on student learning (#001/2013).* Retrieved from http://www.education.alberta.ca/media/6950988/mostudentlearning.pdf

Alberta Education. (2013b). *Learning and technology policy framework*. Retrieved from https://education.alberta.ca/media/7792655/learning-and-technology-policy-framework-web.pdf

Alberta Education. (2016). *Funding manual for school authorities 2016/2017 school year: For school jurisdictions, accredited funded private schools and private ECS operators with students/children in ECS to grade 12.* Retrieved from https://education.alberta.ca/funding-framework-for-k-12/

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson, & A. H. Urquhart, A. (Eds.), *Reading in a foreign language* (pp. 1-27). London, England: Longman.

Alderson, J. C. (2000). *Assessing reading*. Cambridge, England: Cambridge University Press.

Alderson, J. C., & Kremmel, B. (2013). Re-examining the content validation of a grammar test: The (im)possibility of distinguishing vocabulary and structural knowledge. *Language Testing*, 30(4), 535-556. https://doi.org/10.1177/0265532213489568

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, *48*(1), 35-44. https://doi.org/10.1037/0003-066X.48.1.35

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.

Anderson, R. C., & Freebody, P. (1979). *Vocabulary knowledge and reading* (Technical Report No. 136). Retrieved from http://files.eric.ed.gov/fulltext/ED177480.pdf

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp.77-117). Newark, DE: International Reading Association.

Arkoudis, S. (2003). Teaching English as a second language in science classes: Incommensurate epistemologies? *Language and Education, 17*(3), 161-173. https://doi.org/10.1080/09500780308666846

Alberta Education. (2017). *Guide to education: ECS to grade 12 2017-2018.* Retrieved from https://education.alberta.ca/guide-to-education/

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth.* Mahwah, NJ: Lawrence Erlbaum.

August, G. (2010). So, what's behind adult English second language reading? *Bilingual Research Journal: The Journal of the National Association for Bilingual Education, 30*(2), 245-264. https://doi.org/10.1080/15235882.2006.10162876

Baddeley, A. D. (2006). Working memory: An overview. In S. J. Pickering (Ed.), *Working memory and education* (pp. 1-31). Burlington, MA: Academic Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York, NY: Academic Press.

Baddeley, A. D. (2015) Working memory in second language learning. In Z. Wen., M. Borges, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 17-28). Bristol, England: Multilingual Matters.

Ballard & Tighe. (2008a). *IPT3 Reading: Grades 7-12. English* (2nd ed.). Brea, CA: Ballard & Tighe.

Ballard & Tighe. (2008b). *Technical manual. IPT 3 reading & writing: Grades 7-12. English* (2nd ed.). Brea, CA: Ballard & Tighe.

Banks, D. (2008). *The development of scientific writing: Linguistic features and historical context.* London, England: Equinox.

BC Ministry of Education. (2016a). *Overview of class size and composition in British Columbia public schools 2016/2017: Provincial overview.* Retrieved from http://www.bced.gov.bc.ca/reports/pdfs/class_size/2016/public.pdf

BC Ministry of Education. (2016b). *Six-year completion and grade 12 graduation rates - 2015/16. Province - public and independent schools combined.* Retrieved from http://www.bced.gov.bc.ca/reports/pdfs/graduation/prov.pdf

BC Ministry of Education. (2016c). *Provincial required examinations - 2015/2016. Province - public and independent schools combined.* Retrieved from http://www.bced.gov.bc.ca/reports/pdfs/exams/req/prov.pdf

Bell, A. B., & Baecher, L. (2012). Points on a continuum: ESL teachers reporting on collaboration. *TESOL Journal, 3*(3), 488-515. https://doi.org/10.1002/tesj.28

Bell, J., & Burnaby, B. (1984). *A handbook for ESL literacy.* Toronto, ON: OISE Press.

Bernhardt, E. B. (2005). Progress and procrastination in second-language reading. *Annual Review of Applied Linguistics, 25*, 133-150. https://doi.org/10.1017/S0267190505000073

Bernhardt, E. B. (2011). *Understanding advanced second-language reading.* New York, NY: Routledge.

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics, 16*(1), 15-34. https://doi.org/10.1093/applin/16.1.15

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Philadelphia, PA: John Benjamins.

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*(1), 2-20. https://doi.org/10.1016/j.jeap.2010.01.001

Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language and Linguistics, 15*(2), 223-250. https://doi.org/10.1017/S1360674311000025

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* London, England: Longman.

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading, 9*(1), 43-61. https://doi.org/10.1207/s1532799xssr0901_4

Bigelow, M., & Vinogradov, P. (2011). Teaching adult second language learners who are emergent readers. *Annual Review of Applied Linguistics, 31,* 120-136. https://doi.org/10.1017/S0267190511000109

Bossers, B. (1991). On thresholds, ceiling, and short circuits: The relation between L1 reading, L2 reading, and L1 knowledge. *AILA Review, 8*(1), 45-60. Retrieved from http://www.aila.info/download/publications/review/AILA08.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Brisbois, J. E. (1995). Connections between first- and second-language reading. *Journal of Reading Behavior, 27*(4), 565-584. https://doi.org/10.1080/10862969509547899

Brown, B. A., & Ryoo, K. (2008). Teaching science as a language: A "content-first" approach to science teaching. *Journal of Research in Science Teaching, 45*(5), 529-553. https://doi.org/10.1002/tea.20255

Brown, B. A., Ryoo, K., & Rodriguez, J. (2010). Pathway towards fluency: Using 'disaggregate instruction' to promote science literacy. *International Journal of Science Education, 32*(11), 1465-1493. https://doi.org/10.1080/09500690903117921

Bruna, K. R., Vann, R., & Escudero, M. P. (2007). What's language got to do with it? A case study of academic language instruction in a high school "English learner science" class. *Journal of English for Academic Purpose, 6*(1), 36-54. https://doi.org/10.1016/j.jeap.2006.11.006

Bruna, K. R., Vann, R., & Escudero, M. P. (2010). "You got the word now": Problematizing vocabulary-based academic language instruction for English learners in science. *The Tapestry Journal, 2*(1), 19-36. Retrieved from http://journals.fcla.edu/tapestry/article/view/81786/78902

Bunch, G. C., Walqui, A., & Pearson, D. P. (2014). Complex text and new common standards in the United States: Pedagogical implications for English learners, *TESOL Quarterly, 48*(3), 533-559. https://doi.org/10.1002/tesq.175

Carrell, P. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics, 12*(2), 159-179. https://doi.org/10.1093/applin/12.2.159

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19*(2), 246-251. https://doi.org/10.1016/j.lindif.2008.10.002

Carver, R. P. (1977). Toward a theory of reading comprehension and rauding. *Reading Research Quarterly, 13*(1), 8-63. https://doi.org/10.2307/747588

Carver, R. P. (1990). *Reading rate: A review of research and theory.* New York, NY: Academic Press.

Carver, R. P. (1997). Reading for one second, one minute, or one year from the perspective of rauding theory. *Scientific Studies of Reading, 1*(1), 3-43. https://doi.org/10.1207/s1532799xssr0101_2

Carver, R. P. (2000). *The causes of high and low reading achievement.* Mahwah, NJ: Lawrence Erlbaum.

CBC News. (2016, Jan 12). Language turning out to be key hurdle in welcoming Syrian refugees. Retrieved from http://www.cbc.ca/news/canada/new-brunswick/syrian-refugees-language-arabic-moncton-1.3399907

Chall, J. S. (1996). *Stages of reading development* (2nd ed.). Fort Worth, TX: Harcourt-Brace College Publishers.

Cho, S., & McDonnough, J. T. (2009). Meeting the needs of high school science teachers in English language learner instruction. *Journal of Science Teacher Education, 20*(4), 385-402. https://doi.org/10.1007/s10972-009-9136-9

Cobb, T. (n.d.). *Compleat Lexical Tutor*. Accessed at https://www.lextutor.ca

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: Dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256. https://doi.org/10.1037/0033-295X.108.1.204

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238. https://doi.org/10.2307/3587951

Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. *Revue Française de Linguistique Appliquée, 2*(12), 65-78. Retrieved from http://www.cairn.info/revue-francaise-de-linguistique-appliquee-2007-2-page-65.htm

Coxhead, A., Stevens, L., & Tinkle J. (2010). Why might secondary science textbooks be difficult to read? *New Zealand Studies in Applied Linguistics, 16*(2), 37-52. Retrieved from https://search.informit.com.au/documentSummary;dn=893694181605406;res=IELNZC

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson.

Creswell, J., & Plano Clark, K. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York: Routledge.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*(1), 7-29. https://doi.org/10.1177/0956797613504966

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9,* 1-43. Retrieved from http://files.eric.ed.gov/fulltext/ED125311.pdf

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222-251. doi:10.3102/00346543049002222

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70-89). New York, NY: Cambridge University Press.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In N. H. Hornberger (Ed.), *Encyclopedia of language and education* (pp. 487-499). New York, NY: Springer Science+Business Media. https://doi.org/10.1007/978-0-387-30424-3_36

Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing: An Interdisciplinary Journal, 14*(5), 549-568. https://doi.org/10.1023/A:1011100226798

Daneman, M., & Merilke, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422-433. https://doi.org/10.3758/BF03214546

de Oliveira, L. C., & Schleppegrell, M. J. (2015). *Focus on grammar and meaning.* Oxford, England: Oxford University Press.

DelliCarpini, M., & Alonso, O. B. (2014). Teacher education that works: Preparing secondary level math and science teachers for success with English language learners through content based instruction. *Global Education Review, 1*(4), 155-178. Retrieved from http://ger.mercy.edu/index.php/ger/article/view/62/84

DeKeyser, R. M. (2007). The future of practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 287-304). New York, NY: Cambridge University Press.

Demont, E., & Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of*

Educational Psychology, 66(3), 315-332. https://doi.org/10.1111/j.2044-8279.1996.tb01200.x

DiCerbo, P., Anstrom, K. A., Baker, L. L., & Rivera, C. (2014). A review of the literature on teaching academic English to English language learners. *Review of Educational Research, 84*(3), 446-482. https://doi.org/10.3102/0034654314532695

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies.* New York, NY: Oxford University Press.

Dowell, N. M., Graesser, A. C., & Cai, Z., (2016). Language and discourse analysis with coh-metrix: Applications from educational material to learning environments at scale. *Journal of Learning Analytics, 3*(3), 72-95. https://doi.org/10.18608/jla.2016.33.5

Droop, M., & Verhoeven, L. T. (1998). Background knowledge, linguistic complexity, and second-language reading comprehension. *Journal of Literacy Research, 30*(2), 253-271. https://doi.org/10.1080/10862969809547998

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*(1), 78-103. https://doi.org/10.1598/RRQ.38.1.4

Duff, P.A. (2001). Language, literacy, content, and (pop) culture: Challenges for ESL students in mainstream courses. *Canadian Modern Language Review, 58*(1), 103-132. https://doi.org/10.3138/cmlr.58.1.103

Duguay, A., Massoud, L., Tabaku, L., Himmel, J., & Sugarman, J. (2013). *Implementing the common core for English learners: Responses to common questions* (Practitioner Brief). Washington, DC: Center for Applied Linguistics. Retrieved from http://www.cal.org/resource-center/briefs-digests/briefs/implementing-the-common-core-for-english-learners

Early, M. (1990). ESL beginning literacy: A content-based approach. *TESL Canada Journal, 7*(2), 82-93.

Echevarria, J., Vogt, M. E., & Short, D. (2008). *Making content comprehensible for English learners: The SIOP model* (3rd ed.). Boston, MA: Allyn & Bacon.

Echevarria, J., Richards-Tutor, C., Canges, R., & Francis, D. (2011). Using the SIOP model to promote the acquisition of language and science concepts with English learners. *Bilingual Research Journal, 34*(3), 334-351. https://doi.org/10.1080/15235882.2011.623600

Eggins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). London, England: Continuum.

Ellis, N. C. (1994). Implicit and explicit learning – An overview. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-32). San Diego, CA: Academic Press.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143-188. https://doi.org/10.1017/S0272263102002024

Ellis, N., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 75-93). New York, N.Y.: Taylor Francis.

Eskildsen, S. W. (2012). L2 negation constructions at work. *Language Learning, 62*(2), 335-372. https://doi.org/10.1111/j.1467-9922.2012.00698.x

Espin, C. A., & Foegen, A. (1996). Validity of general outcome measures for predicting secondary students' performance on content-area tasks. *Exceptional Children, 62*(6), 497-514. https://doi.org/10.1177/001440299606200602

Fang, Z., & Schleppegrell, M. J. (2008). *Reading in secondary content areas: A language based pedagogy.* Ann Arbor, MI: University of Michigan Press.

Fang, Z., & Schleppegrell, M. J. (2010). Disciplinary literacies across content areas: Supporting secondary reading through functional language analysis. *Journal of Adolescent & Adult Literacy, 53*(7), 587-597. https://doi.org/10.1598/JAAL.53.7.6

Fang, Z., Schleppegrell, M. J., & Cox, B. E. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research, 38*(3), 247-273. https://doi.org/10.1207/s15548430jlr3803_1

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first and second-language reading of fluent bilinguals. *Memory & Cognition, 11*(6), 565-574. https://doi.org/10.3758/BF03198281

Fender, M. (2001). A review of L1 and L2/ESL word integration skills and the nature of L2/ESL word integration development involved in lower-level text processing. *Language Learning, 51*(2)*,* 319-396. https://doi.org/10.1111/0023-8333.00157

Fillmore, L. W. (2014). English language learners at the crossroads of educational reform. *TESOL Quarterly, 48*(3), 624-632. https://doi.org/10.1002/tesq.174

Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review, 21*(1), 45-58.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58*(5), 436-450. https://doi.org/10.1177/001440299205800507

Fujita, K., & Yamashita, J. (2014). The relations and comparisons between reading comprehension and reading rate of Japanese high school EFL learners. *The Reading Matrix, 14*(2), 34-49. Retrieved from http:// www.readingmatrix.com/files/11-864mw1b5.pdf

Fukkink, R. G., Hulstijn, J., & Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal, 89*(1), 54-75. https://doi.org/10.1111/j.0026-7902.2005.00265.x

García, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly, 26(*4), 371-392. https://doi.org/10.2307/747894

Garnett B. (2010). Toward understanding the academic trajectories of ESL youth. *The Canadian Modern Language Review, 66*(5), 677-710. https://doi.org/10.3138/cmlr.66.5.677

Gaux, C., & Gombert, J. (1999). Implicit and explicit syntactic knowledge and reading in preadolescents. *British Journal of Developmental Psychology, 17*(2), 169-188. https://doi.org/10.1348/026151099165212

Gebhard, M. (2010). Teacher education in changing times: A systemic functional (SFL) perspective. *TESOL Quarterly, 44*(4), 797-803. https://doi.org/10.5054/tq.2010.237335

Gebhard, M., Chen, I., Graham, H., & Gunawan, W. (2013). Teaching to mean, writing to mean: SFL, L2 literacy, and teacher education. *Journal of Second Language Writing, 22*(2), 107-124. https://doi.org/10.1016/j.jslw.2013.03.005

Gee, J. P. (2008). What is academic language? In A. S. Rosebery & B. Warren (Eds.), *Teaching science to English language learners: Building on students' strengths* (pp. 57-70). Arlington, VA: National Science Teachers Association.

Gellert, A. S., & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment, 31*(1), 16-28. https://doi.org/10.1177/0734282912451971

Genesee, F., Geva, E., Dressler, C., & Kamil, M. (2006). Synthesis: Cross-linguistic relationships. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth* (pp. 153-183). Mahwah, NJ: Lawrence Erlbaum.

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Lawrence Erlbaum.

Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 217-263). New York, NY: Academic Press.

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes, 23*(3), 265-304. https://doi.org/10.1080/01638539709544994

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430-445. https://doi.org/10.1037//0278-7393.16.3.430

Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading and Writing, 25*(8), 1819-1845. https://doi.org/10.1007/s11145-011-9333-8

Geva, E., & Wang, M. (2001). The development of basic reading skills in children: A cross-language perspective. *Annual Review of Applied Linguistics, 21,* 182-204. https://doi.org/10.1017/S0267190501000113

Goldenberg, C. (2013). Unlocking the research on English learners. What we know – and don't yet know – about effective instruction. *American Educator, 37*(2), 4-11. Retrieved from https://eric.ed.gov/?id=EJ1014021

Gorsuch, G., & Taguchi, E. (2008). Repeated reading for developing reading fluency and reading comprehension: The case of EFL learners in Vietnam. *System, 36*(2), 253-278. https://doi.org/10.1016/j.system.2007.09.009

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*(2), 330-344. https://doi.org/10.1037/a0014320

Gottardo, A., Yan, B., Siegel, L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing. *Journal of Educational Psychology, 93*(3), 530-542. https://doi.org/10.1037/0022-0663.93.3.530

Government of Canada (2017). #WelcomeRefugees: Canada resettled Syrian refugees [website]. Retrieved from https://www.canada.ca/en/immigration-refugees-citizenship/services/refugees/welcome-syrian-refugees.html

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* New York, NY: Cambridge University Press.

Grabe, W. (2010). Fluency in reading – thirty-five years later. *Reading in a Foreign Language, 22*(1), 71-83. Retrieved from http://nflrc.hawaii.edu/rfl/April2010/articles/grabe.pdf

Graesser, A. C., Cai, Z., Louwerse, M. M., & Daniel, F. (2006). Question understanding aid (QUAID): A web facility that tests question comprehensibility. *Public Opinion Quarterly, 70*(1), 3-22. https://doi.org/10.1093/poq/nfj012

Grant, A., Gottardo, A., & Geva, E. (2011). Reading in English as a first or second language: The case of grade 3 Spanish, Portuguese, and English speakers. *Learning Disabilities Research & Practice, 26*(2), 67-83. https://doi.org/10.1111/j.1540-5826.2011.00327.x

Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL, 25*(2), 165-198. https://doi.org/10.1017/S0958344013000013

Gue, D., Hutton, G., Jeans, S. L., Leong, E., Lunn, D., Mason, A.,…Webb, M. (2004). *ScienceFocus 10.* Toronto, ON: McGraw-Hill Ryerson.

Gunderson, L. (1986). Content reading and ESL students. *TESL Canada Journal, 4*(1), 49-53. https://doi.org/10.18806/tesl.v4i1.496

Gunderson, L., D'Silva, R. A., & Odo, D. M. (2012). Immigrant students navigating Canadian schools: A longitudinal view. *TESL Canada Journal, 29*(6), 142-156. Retrieved from http://www.teslcanadajournal.ca/index.php/tesl/article/view/1114/933

Guo, Y., & Roehrig, A. D. (2011). Roles of general versus second language (L2) knowledge in L2 reading comprehension. *Reading in a Foreign Language, 23*(1), 42-64. Retrieved from http://nflrc.hawaii.edu/rfl/April2011/articles/guo.pdf

Hacquebord, H. (1994). L2-reading in the content areas: Text comprehension in secondary education in the Netherlands. *Journal of Research in Reading, 17*(2), 83-98. https://doi.org/10.1111/j.1467-9817.1994.tb00056.x

Halliday, M. A. K. (1989). *Spoken and written language.* New York, NY: Oxford University Press.

Halliday, M. A. K. (2004a). Some grammatical problems in scientific English. In J. J. Webster (Ed.), *The language of science. The collected works of M. A. K. Halliday* (Vol. 5) (pp. 159-180). London, England: Continuum. (Original work published 1989)

Halliday, M. A. K. (2004b). On the grammar of scientific English. In J. J. Webster (Ed.), *The language of science. The collected works of M. A. K. Halliday* (Vol. 5) (pp. 181-198). London, England: Continuum. (Original work published 1997)

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power.* London, England: Falmer Press.

Halliday, M. A. K., & Matthiessen, C. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). New York, NY: Hodder Education.

Hamada, M., & Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Language Learning, 58*(1), 1-31. https://doi.org/10.1111/j.1467-9922.2007.00433.x

Harper, C., & de Jong, E. (2004). Misconceptions about teaching English-language learners. *Journal of Adolescent and Adult Literacy, 48*(2), 152-162. https://doi.org/10.1598/JAAL.48.2.6

Harrison, J., & Lakin, J. (2018). Mainstream teachers' implicit beliefs about English language learners: An implicit association test study of teacher beliefs. *Journal of Language, Identity & Education, 17*(2), 85-102. https://doi.org/10.1080/15348458.2017.1397520

Himmel, J., Short, D., Richards, C., & Echevarria, J. (2009, May). Using the SIOP model to improve middle school science instruction. *Create Brief*. Retrieved from http://www.cal.org/create/publications/briefs/pdfs/using-the-siop-model-to-improve-middle-school-science-instruction.pdf

Hines, P. J., Wible, B., & McCartney, M. (2010). Learning to read, reading to learn. *Science, 328*(5977), 477. https://doi.org/10.1126/science.328.5977.447

Hirano, E. (2015). 'I read, I don't understand': Refugees coping with academic reading. *ELT Journal, 69*(2), 178-187. https://doi.org/10.1093/elt/ccu068

Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*(2), 127-160. https://doi.org/10.1007/BF00401799

Houle, R., & Corbeil, J. P. (2017). *Language projections for Canada, 2011 to 2036.* Statistics Canada Catalogue no. 89-657-X2017001. Ottawa, ON: Statistics Canada. Retrieved from https://www.statcan.gc.ca

Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430. Retrieved from http://nflrc.hawaii.edu/rfl/PastIssues/rfl131hsuehchao.pdf

Huang, J. (2004). Socialising ESL students into the discourse of school science through academic writing. *Language and Education, 18*(2), 97-123. https://doi.org/10.1080/09500780408666870

Hudson, T. (2007). *Teaching second language reading.* Oxford, England: Oxford University Press.

Hudson, T., & Lorena, L. (2015). Design issues and inference in experimental L2 research. *Language Learning, 65*(s1), 76-96. https://doi.org/ 10.1111/lang.12113

Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly, 8*(3), 229-249. https://doi.org/10.1080/15434303.2011.565844

Hulstijn, J. H. (2015). *Language proficiency in native and non-native speakers: Theory and research.* Amsterdam, Netherlands: John Benjamins.

Immigration, Refugees, and Citizenship Canada. (2017a). *Growing Canada's economic future* [Press release]. Retrieved from https://www.canada.ca/en/immigration-refugees-citizenship/news/2017/11/news_release_growingcanadaseconomicfuture.html

Immigration, Refugees, and Citizenship Canada. (2017b). *Supplementary information 2018-2020 immigration levels plan* [Press release]. Retrieved from https://www.canada.ca/en/immigration-refugees-citizenship/news/notices/supplementary-immigration-levels-2018.html?_ga=2.221685072.990218399.1525819899-2039875686.1444261230

Indrarathne, B., & Komos, J. (2018). The role of working memory in processing L2 input: Insights from eye-tracking. *Bilingualism: Language and Cognition, 21*(2), 355-374. https://doi.org/10.1017/S1366728917000098

Jensen, E. D., & Vinther, T. (2003). Exact repetition in input enhancement in second language acquisition. *Language Learning, 53*(3), 373-428. https://doi.org/10.1111/1467-9922.00230

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*(1), 160-212. https://doi.org/10.1111/lang.12034

Johnson, E. S., Semmelroth, C., Allsion, J., & Fritsch, T. (2013). The technical properties of science content maze passages for middle school students. *Assessment for Effective Intervention, 38*(4), 214-223. https://doi.org/10.1177/1534508413489337

Jones, R. H., & Lock, G. (2011). *Function grammar in the ESL classroom: Noticing, exploring and practicing.* New York, NY: Palgrave Macmillan.

Kanno, Y., & Kangas, S. E. (2014). "I'm not going to be, like, for the AP": English language learners' limited access to advanced college-preparatory courses in high school. *American Educational Research Journal, 51*(5), 848-878. https://doi.org/10.3102/0002831214544716

Kim, Y. S., Petscher, Y., & Foorman, B. (2015). The unique relation of silent reading fluency to end-of-year reading comprehension: Understanding individual differences at the student, classroom, school, and district levels. *Reading and Writing, 28*(1), 131-150. https://doi.org/10.1007/s11145-013-9455-2

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*(2), 163-182.

Kintsch, W. (1998). *Comprehension: A framework for cognition.* New York, NY: Cambridge University Press.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach.* New York, NY: Cambridge University Press.

Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning, 57*(Suppl. 1), 1-44. https://doi.org/10.1111/0023-8333.101997010-i1

Koda, K. (2010). The role of reading in fostering transcultural competence. *Reading in a Foreign Language, 22*(Suppl. 1), 5-10. Retrieved from http://nflrc.hawaii.edu/rfl/January2010/articles/koda.pdf

Krahn, H., & Taylor, A. (2005). Resilient teenagers: Explaining the high educational aspirations of visible-minority youth in Canada. *Journal of International Migration and Integration, 6*(3-4), 405-434. https://doi.org/10.1007/s12134-005-1020-7

Krajcik, J. S., & Sutherland, L. M. (2010). Supporting students in developing literacy in science. *Science, 328*(5877), 456-459. https://doi.org/10.1126/science.1182593

Kremmel, B., Brunfaut, T., & Alderson, J. C. (2017). Exploring the role of phraseological knowledge in foreign language reading. *Applied Linguistics, 38*(6), 848-870. https://doi.org/10.1093/applin/amv070

Kuhn, M., & Stahl, S. (1998). Teaching children to learn word meanings from context: A synthesis and some questions. *Journal of Literacy Research, 30*(1), 119-138. https://doi.org/10.1080/10862969809547983

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology,* 95(1), 3-21. https://doi.org/10.1037/0022-0663.95.1.3

Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly, 45*(2), 230-251. https://doi.org/10.1598/RRQ.45.2.4

Kurvers, J. (2015). Emerging literacy in adult second-language learners: A synthesis of research findings in the Netherlands. *Writing Systems Research, 7*(1), 58-78. https://doi.org/10.1080/17586801.2014.943149

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* (Doctoral dissertation). Retrieved from http://scholarworks.gsu.edu/alesl_diss/35

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R* (2nd ed.). New York, NY: Routledge.

Larson-Hall, J. (2017). Moving beyond the bar plot and the line graph to create informative and attractive graphics. *The Modern Language Journal, 101*(1), 244-270. https://doi.org/10.1111/modl.12386

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(s1), 127-159. https://doi.org/ 10.1111/lang.12115

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp.126-132). London, England: Macmilian Academic & Professional.

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review, 59*(4), 567-587. https://doi.org/10.3138/cmlr.59.4.567

Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307-322. https://doi.org/10.1093/applin/16.3.307

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15-30. Retrieved from http://nflrc.hawaii.edu/rfl/April2010/articles/laufer.pdf

Lee, O., & Fradd, S. H. (1998). Science for all, including students from non-English-language backgrounds. *Educational Researcher, 27*(4), 12-21. https://doi.org/10.3102/0013189X027004012

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher, 42*(4), 223-233. https://doi.org/10.3102/0013189X13480524

Lesaux, N., & Geva, E. (2006). Synthesis: Development of literacy in language-minority students. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth* (pp. 53-74). Mahwah, NJ: Lawrence Erlbaum.

Lesaux, N., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal, 51*(6), 1159-1194. https://doi.org/10.3102/0002831214532165

Lesaux, N., Koda, K., Siegel, L., & Shanahan, T. (2006). Development of literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth* (pp. 75-122). Mahwah, NJ: Lawrence Erlbaum.

Lipka, O., & Siegel, L. S. (2012). The development of reading comprehension skills in children learning English as a second language. *Reading and Writing, 25*(8), 1873-1898. https://doi.org/10.1007/s11145-011-9309-8

Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*(1), 33-42. https://doi.org/10.1177/003368828501600103

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496.

Macaro, E., Handley, Z., & Walter, C. (2012). A systematic review of CALL in English as a second language: Focus on primary and secondary education. *Language Teaching, 45*(1), 1-43. https://doi.org/10.1017/S0261444811000395

Mackay, T., & Tavares, T. (2005). *Building hope: Appropriate programming for adolescent and young adult newcomers of war-affected backgrounds and Manitoba schools: A preliminary report for consultation and discussion.* Winnipeg, MB: Manitoba Education, Training and Citizenship.

MacSweeney, M., Waters, D., Brammer, M. J., Woll, B., & Goswami, U. (2008). Phonological processing in deaf signers and the impact of age of first language acquisition. *Neuroimage, 40*(3), 1369-1379. https://doi.org/10.1016/j.neuroimage.2007.12.047

MacWhinney, B. (2017). A shared platform for studying second language acquisition. *Language Learning, 67*(S1), 254-275. https://doi.org/10.1111/lang.12220

Martin, J. R. (1993). Literacy in science: Learning to handle text as technology. In M.A.K. Halliday & J. R. Martin (Eds.), *Writing science: Literacy and discursive power* (pp. 184-222). London, England: Falmer Press.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*(1), 299-320. https://doi.org/10.1093/applin/ams010

McMaster, K. L., Wayman, M. M., & Cao, M. (2006). Monitoring the reading progress of secondary-level English learners: Technical features of oral reading and maze tasks. *Assessment for Effective Intervention, 31*(4), 17-31. https://doi.org/10.1177/073724770603100402

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. New York, NY: Cambridge University Press.

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-metrix: Capturing linguistic features of cohesion. *Discourse Processes, 47*(4), 292-330. https://doi.org/10.1080/01638530902959943

Menken, K. (2008). *English learners left behind: Standardized testing as language policy.* Clevedon, England: Multilingual Matters.

Miller, D. (2011). ESL reading textbooks vs. university textbooks: Are we giving our students the input they may need? *Journal of English for Academic Purposes, 10*(1)*,* 32-46. https://doi.org/10.1016/j.jeap.2010.12.002

Miller, J. (2009). Teaching refugee learners with interrupted education in science: Vocabulary, literacy and pedagogy. *International Journal of Science Education, 31*(4), 571-592. https://doi.org/10.1080/09500690701744611

Min, H. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning, 58*(1), 73-115. https://doi.org/10.1111/j.1467-9922.2007.00435.x

Mohan, B. (1986). *Language and content.* Reading, MA: Addison-Wesley.

Mohan, B., & Slater, T. (2006). Examining the theory/practice relation in a high school science register: A functional linguistic perspective. *Journal of English for Academic Purposes, 5*(4), 302-316. https://doi.org/10.1016/j.jeap.2006.08.004

Morency, J. D., Malenfant, É, C., & MacIsaac, S. (2017). *Immigration and diversity: Population projections for Canada and its regions, 2011 to 2036.* Statistics Canada Catalogue no. 91-551-X. Ottawa, ON: Statistics Canada. Retrieved from https://www.statcan.gc.ca

Muijselaar, M. M. L., Kendeou, P., de Jong, P. F., & van den Broek, P. W. (2017). What does the CBM-Maze test measure? *Scientific Studies of Reading, 21*(2), 120-132. https://doi.org/10.1080/10888438.2016.1263994

Munro, J., Abbott, M. L., & Rossiter, M. J. (2013). Getting to the science: Helping English language learners show what they know. *Alberta Science Education Journal, 43*(1), 34-43.

Musgrave, J., & Parkinson, J. (2014). Getting to grips with noun groups. *ELT Journal, 68*(2), 145-154. https://doi.org/10.1093/elt/cct078

Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 52-77). New York, NY: Guildford.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91-108. https://doi.org/10.1002/RRQ.011

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233-253. https://doi.org/10.2307/747758

Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *Modern Language Journal, 87*(2), 261-276. https://doi.org/10.1111/1540-4781.00189

Nassaji, H. (2014). The role and importance of lower-level processes in second language reading. *Language Teaching, 47*(1), 1-37. https://doi.org/10.1017/S0261444813000396

Nassaji, H., & Geva, E. (1999). The contribution of phonological and orthographic processing skills to adult ESL reading: Evidence from native speakers of Farsi. *Applied Psycholinguistics, 20*(2), 241-267. https://doi.org/10.1017/S0142716499002040

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York, NY: Heinle Cengage Learning.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge, England: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63*(1), 59-82. https://doi.org/10.1353/cml.2006.0049

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics, 21*(2), 229-241. https://doi.org/10.1017/S0142716400002046

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of the subgroups.* Washington, DC: The National Institute of Child Health and Human Development. Retrieved from http://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Ngo, H. V., & Schleifer, B. (2007). Immigrant children and youth in focus. *Canadian Issues, Spring*, 29-33. Retrieved from http://canada.metropolis.net/pdfs/CITC_Spring_05_EN.pdf

Norouzian, R., & Plonsky, L. (2018). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research, 34*(2), 257-271. https://doi.org/10.1177/0267658316684904

Norris, J. M. (2015). Statistical significance testing in second language research: Basic problems and suggestions for reform. *Language Learning, 65*(S1), 97-126. https://doi.org/10.1111/lang.12114

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417-528. https://doi.org/10.1111/0023-8333.00136

Norris, J. M., Plonsky, L., Ross, S. J., & Schoonen, R. (2015). Guideline for reporting quantitative methods and results in primary research. *Language Learning, 65*(2), 470-476. https://doi.org/ 10.1111/lang.12104

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*(2), 224-240. https://doi.org/10.1002/sce.10066

Norton, B. (2013). *Identity and language learning: Extending the conversation* (2nd ed.). Bristol, England: Multilingual Matters.

Norton, B., & De Costa, P. (2018). Research tasks on identify in language learning and teaching. *Language Teaching, 51*(1), 90-112. https://doi.org/10.1017/S0261444817000325

Odo, D. M. (2012). The impact of high school exit exams on ESL learners in British Columbia. *English Language Teaching, 5*(9), 1-8. https://doi.org/10.5539/elt.v5n9p1

OECD & Statistics Canada. (2011). *Literacy for life: Further results from the adult literacy and life skills survey.* Ottawa, ON: OECD Publishing.

Palincsar, A. S., & Schleppegrell, M. J. (2014). Focusing on language and meaning while learning with text. *TESOL Quarterly, 48*(3), 616-623. https://doi.org/10.1002/tesq.178

Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for academic purposes students. *Journal of English for Academic Purposes, 14,* 48-59. https://doi.org/10.1016/j.jeap.2013.12.001

Pearson, P. D., Moje, E., & Greenleaf, C. (2010). Literacy and science: Each in the service of the other. *Science, 328*(5977), 459-463. https://doi.org/10.1126/science.1182595

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in Foreign Language, 22*(1), 31-55. Retrieved from http://nflrc.hawaii.edu/rfl/April2010/articles/pellicersanchez.pdf

Perfetti, C. (1985). Reading ability. In B. Hutson (Ed.), *Advances in reading/language research* (pp. 231-256). London, England: Oxford University Press.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language processing* (pp. 167-208). London, England: Oxford University Press.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading, 11*(4), 357-383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Britt, M. A. (1995). Where do propositions come from? In C. A. Weaver, S. Mannes & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 11-34). Hillsdale, NJ: Lawrence Erlbaum.

Perfetti, C., & Hart, L. (2001). The lexical basis of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67-86). Washington, DC: American Psychological Association.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227-247). Oxford, England: Blackwell.

Pica, T. (2002). Subject-matter content: How does it assist the interactional and linguistic needs of classroom language learners? *The Modern Language Journal, 86*(1), 1-19. https://doi.org/10.1111/1540-4781.00133

Pierce, R. L., McMaster, K. L., & Deno, S. L. (2010). The effects of using different procedures to score maze measures. *Learning Disabilities Research & Practice, 25*(3), 151-160. https://doi.org/10.1111/j.1540-5826.2010.00313.x

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language, 18*(1), 1-28. Retrieved from http://nflrc.hawaii.edu/rfl/April2006/pigada/pigada.pdf

Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition, 35*(4), 655-687. https://doi.org/10.1017/S0272263113000399

Plonsky, L. (2014). Study quality in quantitative L2 research (1990-2010): A methodological synthesis and call for reform. *The Modern Language Journal, 98*(1), 450-470. https://doi.org/10.1111/j.1540-4781.2014.12058.x

Plonsky, L. (2015). Quantitative considerations for improving replicability in CALL and applied linguistics. *CALICO Journal, 32*(5), 232-244. https://doi.org/10.1558/cj.v32i2.26857

Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912. https://doi.org/10.1111/lang.12079

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513-536. https://doi.org/10.1111/1467-9922.00193

Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing, 21*(1), 28-52. https://doi.org/10.1191/0265532204lt273oa

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*(1), 77-89. https://doi.org/ 10.2307/3585941

Roessingh, H., & Douglas, S. (2012a). English language learners' transitional needs from high school to university: An exploratory study. *Journal of International Migration and Integration, 13*(3), 285-301. https://doi.org/10.1007/s12134-011-0202-8

Roessingh, H., & Douglas, S. R. (2012b). Educational outcomes of English language learners at university. *Canadian Journal of Higher Education, 42*(1), 80-97. Retrieved from http://ojs.library.ubc.ca/index.php/cjhe/article/view/182449/182509

Römer, U. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*, *7*, 141-163. https://doi.org/10.1075/arcl.7.06rom

Rossiter, M. J., & Rossiter, K. R. (2009). Diamonds in the rough: Bridging gaps in supports for at-risk immigrant and refugee youth. *Journal of International Migration and Integration, 10*(4), 409-429. https://doi.org/10.1007/s12134-009-0110-3

Rossiter, M. J., Hatami, S., Ripley, D., & Rossiter, K. R. (2015). Immigrant and refugee youth settlement experiences: "A new kind of war". *International Journal of Child, Youth, and*

Family Studies, 6(4.1), 746-770. Retrieved from http://journals.uvic.ca/index.php/ijcyfs/article/view/15056/5996

Rubinstein-Avila, E., & Lee, E. H. (2014). Secondary teachers and English language learners (ELLs): Attitudes, preparation, and implications. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 87*(5), 187-191. https://doi.org/10.1080/00098655.2014.910162

Sandner, L., Lacy, D., Schaeffer, H., Sosnowski, C., McDougall, M., & Martha, J. (2004). *Addison Wesley science 10*. Toronto, ON: Pearson Education Canada.

Schall-Leckrone, L., & McQuillan, P. J. (2012). Preparing history teachers to work with English learners through a focus on the academic language of historical analysis. *Journal of English for Academic Purposes, 11*, 246-266. https://doi.org/10.1016/j.jeap.2012.05.001

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* London, England: Palgrave MacMillan.

Schmitt, N. (2014). Size and depth of vocabulary knowledge. *Language Learning, 64*(4), 913-951. https://doi.org/10.1111/lang.12077

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26-43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Segalowitz, N. S., & Segalowitz, S. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14*(3), 369-385. https://doi.org/10.1017/S0142716400010845

Segalowitz, N. S., Poulsen, C., & Komoda, M. (1991). Lower level components or reading skill in higher level bilinguals: Implications for reading instruction. *AILA Review, 8,* 15-30. Retrieved from http://www.aila.info/download/publications/review/AILA08.pdf

Shakya, Y. B., Guruge, S., Hynie, M., Akbari, A., Malik, M., Htoo, S.,…Alley, S. (2010). Aspirations for higher education among newcomer refugee youth in Toronto: Expectations, challenges, and strategies. *Refuge, 27*(2), 65-78. Retrieved from http://pi.library.yorku.ca/ojs/index.php/refuge/article/viewFile/34723/31553

Shanahan, T., & Beck, I. (2006). Effective literacy teaching. In D. August, & T. Shanahan (Eds.), *Developing literacy in second language learners: Report of the national literacy panel on language-minority children and youth* (pp. 415-488). Mahwah, NJ: Lawrence Erlbaum.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34*(3), 164-172. https://doi.org/10.1177/002246690003400305

Shinn, M. R., & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of reading maze for use in general outcome measurement*. Retrieved from http://www.aimsweb.com/wp-content/uploads/maze-manual.pdf

Shiotsu, T. (2009). Reading ability and components of word recognition speed: The case of L1-Japanese EFL learners. In Z. Han & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 15-39). Ann Arbor, MI: University of Michigan Press.

Shiotsu, T. (2010). *Components of L2 reading: Linguistic and processing factors in the reading test performances of Japanese EFL learners*. New York, NY: Cambridge University Press.

Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing, 24*(1), 99-128. https://doi.org/10.1177/0265532207071513

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328*(5977), 450-452. https://doi.org/10.1126/science.1182597

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). https://doi.org/10.1017/CBO9780511609664.008

Sonbul, S., & Schmitt, N. (2010). Direct teaching of vocabulary after reading: Is it worth the effort? *ELT Journal, 64*(3), 253-260. https://doi.org/10.1093/elt/ccp059

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning, 59*(1), 203-243. https://doi.org/10.1111/j.1467-9922.2009.00504.x

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Do L1 reading achievement and L1 print exposure contribute to the prediction of L2 proficiency? *Language Learning, 62*(2), 473-505. https://doi.org/10.1111/j.1467-9922.2012.00694.x

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*(1), 32-71. https://doi.org/10.2307/747348

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407. Retrieved from http://www.jstor.org/stable/747612

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* New York, NY: Guilford.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Upper Saddle River, NJ: Pearson.

Taber, K. S. (2017). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education.* Advanced online publication. https://doi.org/10.1007/S11165-016-9602-2

Taguchi, E., & Gorsuch, G. J. (2002). Transfer effects of repeated EFL reading on reading new passages: A preliminary investigation. *Reading in a Foreign Language, 14*(1), 43-65. Retrieved from http://nflrc.hawaii.edu/rfl/April2002/taguchi/taguchi.pdf

Taguchi, E., Takaysu-Maass, M., & Gorsuch, G. J. (2004). Developing reading fluency in EFL: How assisted repeated reading and extensive reading affect fluency development. *Reading in a Foreign Language, 16*(2), 70-96. Retrieved from http://nflrc.hawaii.edu/rfl/October2004/taguchi/taguchi.pdf

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381-398. doi:10.1207/s1532799xssr1004_3

Taverniers, M. (2006). Grammatical metaphor and lexical metaphor: Different perspectives on semantic variation. *Neophilologus, 90*(2), 321-332. https://doi.org/10.1007/s11061-005-0531-y

Taylor, R. P. (1980). Introduction. In R. P. Taylor (Ed.), *The computer in school: Tutor, tool, tutee* (pp. 1-10). New York, NY: Teacher College Press.

Thiessen, V. (2009). The pursuit of postsecondary education: A comparison of first nations, African, Asian, and European Canadian youth. *Canadian Review of Sociology, 46*(1), 5-37. https://doi.org/10.1111/j.1755-618X.2009.01201.x

Tolar, T. D., Barth, A. E., Francis, D. J., Fletcher, J. M., Stuebing, K. K., & Vaughn, S. (2012). Psychometric properties of maze tasks in middle school students. *Assessments for Effective Intervention, 37*(3), 131-146. https://doi.org/10.1177/1534508411413913

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Toohey, K. (2007). Are the lights coming on? How can we tell? English language learners and literacy assessment. *The Canadian Modern Language Review, 64*(2), 253-272. Retrieved from https://muse.jhu.edu/article/233951

Toohey, K., & Derwing, T. M. (2008). Hidden losses: How demographics can encourage incorrect assumptions about ESL High School students. *Alberta Journal of Educational Research, 54*(2), 178-193. Retrieved from http://ajer.synergiesprairies.ca/index.php/ajer/article/viewFile/628/611

Toohey, K., Dagenais, D., Fodor, A., Hof, L., Nuñez, O., Singh, A, & Schulze, L. (2015). "That sounds so cooool": Entanglements of children, digital tools, and literacy practices. *TESOL Quarterly, 49*(3), 461-485. https://doi.org/10.1002/tesq.236

Toronto District School Board. (2018). *English as a second language/English literacy development.* Retrieved from Toronto District School Board website: http://www.tdsb.on.ca/HighSchool/YourSchoolDay/EnglishasaSecondLanguage.aspx

Townsend, D., Filippini, A., Collins, P., & Biancarosa, G. (2012). Evidence for the importance of academic word knowledge for the academic achievement of diverse middle school students. *The Elementary School Journal, 112*(3), 497-518. https://doi.org/10.1086/663301

Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning. *Annual Review of Applied Linguistics, 30*, 270-291. https://doi.org/10.1017/S0267190510000140

Ure, J. (1971). Lexical density and register differentiation. In G. E. Perren & J. L. M. Trimm (Eds.), *Applications of linguistics: Selected papers of the 2nd International Congress of Applied Linguistics, Cambridge, 1969* (pp. 443-452). New York, NY: Cambridge University Press.

Valdés, G. (2004). Between support and marginalization: The development of academic language in linguistic minority children. *International Journal of Bilingual Education and Bilingualism, 7*(2), 102-132. https://doi.org/10.1080/13670050408667804

Van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology, 22*(3), 335-351. https://doi.org/10.1002/acp.1418

Van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A “landscape” view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165-187). Hillsdale, NJ: Lawrence Erlbaum.

Van Gelderen, A., Schoonen, R., Stoel, R. D., & de Glopper, K. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology, 99*(3), 477-491. https://doi.org/ 10.1037/0022-0663.99.3.477

Van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology, 96*(1), 19-30. https://doi.org/10.1037/0022-0663.96.1.19

Van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology, 106*(4), 1036-1048. https://doi.org/10.1037/a0036293

Van den Bosch, L. J., Segers, E., & Verhoeven, L. (2018). Online processing of causal relationships in beginning first and second language readers. *Learning and Individual Differences, 61*, 59-67. https://doi.org/10.1016/j.lindif.2017.11.007

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4*(4), 313-330. https://doi.org/10.1207/S1532799XSSR0404_4

Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing, 25*(8), 1805-1818.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*(2), 217-234. https://doi.org/10.1017/S0142716401002041

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*(2), 130-163. Retrieved from http://nflrc.hawaii.edu/rfl/October2003/waring/waring.pdf

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching, 31*(2), 57-71. https://doi.org/10.1017/S0261444800012970

Wayman, M. M., McMaster, K. L., Sáenz, L. M., & Watson, J. A. (2010). Using curriculum-based measurement to monitor secondary English language learners' responsiveness to peer-mediated reading instruction. *Reading & Writing Quarterly, 26*(4), 308-332. https://doi.org/10.1080/10573569.2010.500260

Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*(2), 85-120. https://doi.org/10.1177/00224669070410020401

Wexler, J., Reed, D. K., Mitchell, M., Doyle, B., & Clancy, E. (2015). Implementing an evidence-based instructional routine to enhance comprehension of expository text. *Intervention in School and Clinic,* 50(3), 1-8. https://doi.org/10.1177/1053451214542042

Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26*(4), 207-214. https://doi.org/10.1177/07419325050260040301

Willis, M., & Ohashi, Y. (2012). A model of L2 vocabulary learning and retention. *The Language Learning Journal, 40*(1), 125-137. https://doi.org/10.1080/09571736.2012.658232

Wong Fillmore, L. (2014). English language learners at the crossroads of educational reform. *TESOL Quarterly, 48*(3), 624-632. https://doi.org/10.1002/tesq.174

Yamashita, J. (2013). Word recognition subcomponents and passage level reading in a foreign language. *Reading in a Foreign Language, 25*(1), 52-71. Retrieved from http://nflrc.hawaii.edu/rfl/April2013/articles/yamashita.pdf

Yamshita, J., & Shiotsu, T. (2017). Comprehension and knowledge components that predict L2 reading: A latent-trait approach. *Applied Linguistics, 38*(1), 43-67. https://doi.org/10.1093/applin/amu079

Zabjek, A. (2015, October 9). As many as one-quarter of Edmonton students learning English as second language. *The Edmonton Journal.* Retrieved from http://edmontonjournal.com/news/local-news/as-many-as-one-quarter-of-all-edmonton-students-learning-english-as-second-language

Zhang, D. (2012). Vocabulary and grammar knowledge in second language reading comprehension: A structural equation modeling study. *The Modern Language Journal, 96*(4), 558-575. https://doi.org/10.1111/j.1540-4781.2012.01398.x

Appendix A

Background Questionnaire

Name: _____________

Please provide the following background information.

1. How old are you? ______ years
2. Are you: ❑ Male ❑ Female ❑ Prefer not to say
3. Where were you born?

   City/town: ____________________________

   Country: _____________________________

4. What is your first language (mother tongue)? _____________
5. How many years of schooling have you had in your first language? ________ years
6. How long have you studied ESL in Canada? Since ______(month) ______ (year)

Appendix B

Post-intervention Questionnaire

Read the following questions, then circle the answer that is the closest to how you feel.

1. How interested were you in using my website, *Reading Scientific Texts*?

Not interested at all	Somewhat interested	Interested	Very interested

2. How interested are you in science in school?

Not interested at all	Somewhat interested	Interested	Very interested

3. How much has using the website help you become better at reading in English?

Not at all	A little bit	Somewhat	A lot

4. How difficult were the readings for you?

Very difficult	Slightly difficult	Easy	Very easy

a. Please explain why they were easy or difficult for you.

____

____

____

____

5. How difficult were the activities for you?

Very difficult	Slightly difficult	Easy	Very easy

a. Please explain why they were easy or difficult for you.

____

____

6. How difficult was the website to use for you?

Very difficult	Slightly difficult	Easy	Very easy

a. Please explain why it was difficult or easy to use for you.

7. Would you continue using the website on your own to practice reading in English?

Yes	No

a. Please explain.

8. What suggestions do you have for improving the website?

9. What suggestions do you have for improving the activities?

10. Do you think a similar set of online activities could help your reading in your other subjects?

Yes	No

a. Please explain.

11. How enjoyable were the activities? Please explain

12. Do you have other comments about the website? If so, please share them here.

Thank you!!

Appendix C

Knowledge of Nominal Groups Test

Directions: Read the following sentences. Focus on the underlined portion. Then read the three choices. For each choice, decide if it has a similar meaning or different meaning to the underlined portion. Place a checkmark beside your decision[1].

***Example**. <u>The popularity of cellphones with students</u> remains high.*

a) Cellphones are popular. *Similar* ☑ *Different* ☐

b) Students are popular. *Similar* ☐ *Different* ☑

c) Cellphones are popular with students. *Similar* ☑ *Different* ☐

1. <u>A ball at rest on a billiard table</u> will remain at rest.

a) A ball is on a table. The ball is not moving. Similar ☑ Different ☐

b) A ball is not moving. The table is not moving. Similar ☐ Different ☑

c) A ball is resting on a billiard table. Similar ☑ Different ☐

2. <u>Work done on the system by its surroundings</u> is considered positive work because the energy of the system increases.

a) The surroundings do work. This work is on the system. Similar ☑ Different ☐

b) The system does work. This work is on the surroundings. * Similar ☐ Different ☑

c) The surroundings do work on the system. Similar ☑ Different ☐

---

[1] Correct answers have been provided for review purposes.

3. In <u>the absence of any external unbalanced forces</u>, such as resistive forces, all objects tend to maintain uniform motion or stay at rest.

a) External unbalanced forces are not there. Similar ☑ Different ☐

b) Forces are not external, not balanced, and not there. Similar ☑ Different ☐

c) Unbalanced are external, forces are external, but are not there.

Similar ☐ Different ☑

4. By accident, he passed <u>a metal wire that had a current passing through it over a compass.</u>

a) The metal wire was over a compass. Similar ☑ Different ☐

b) The current passed through a wire and over a compass. Similar ☑ Different ☐

c) The current passed through a compass and a wire. Similar ☐ Different ☑

5. <u>The difference in temperature between the junctions</u> caused the electrons inside the metal to move, producing an electric current.

a) The temperature is between the junctions. Similar ☐ Different ☑

b) The junctions have different temperatures. Similar ☑ Different ☐

c) One junction has one temperature. The other junction has another temperature.

Similar ☑ Different ☐

6. The angle of inclination also causes <u>variation in the number of hours of daylight at different latitudes.</u>

a) Different latitudes have different number of hours of daylight. Similar ☑ Different ☐

b) As latitudes change, the number of hours of daylight changes. Similar ☑ Different ☐

c) The number of hours of daylight changes. The latitudes also change.

Similar ☐ Different ☑

7. Figure D2.15 shows the relative contribution of different aspects of Earth's average net <u>radiation budget</u>, based on data collected by NASA for the Earth Radiation Budget Experiment project.

a) The budget is created from radiation. Similar ☐ Different ☑

b) The budget measures radiation. Similar ☑ Different ☐

c) The budget is about radiation. Similar ☑ Different ☐

8. The **Coriolis effect** is <u>the deflection of any object from a straight line path by the rotation of Earth</u>.

a) The Earth rotates and deflects objects. Similar ☑ Different ☐

b) Objects deflect the Earth from its line and the Earth rotates. Similar ☐ Different ☑

c) Objects move in a straight line. The Earth deflects them. Similar ☑ Different ☐

9. <u>Biomes function as a system, or a set of interconnected parts</u>.

a) Interconnected parts is called a system. Similar ☑ Different ☐

b) Biomes function in two ways: 1) a system and 2) a set of interconnected parts.

Similar ☐ Different ☑

c) Biomes have interconnected parts. Similar ☑ Different ☐

10. Deserts support <u>a number of unique animal species</u> adapted to the hot, dry climate, such as camels and running birds.

a) There are many different animals from the same species. Similar ☐ Different ☑

b) There are many different animals from different species. Similar ☑ Different ☐

c) There are many different species of animals. Similar ☑ Different ☐

Appendix D

Silent Reading Fluency Test – Form A[2]

Directions: Below is a passage about energy demand in society. *When the teacher tells you to begin, silently read the following passage. Some of the words are replaced by three words. As you read, circle the one that makes the most sense. Only one word is correct. You will have one minute to read as much as you can.*[3]

Example: Learning about science (about, is, play) fun.

Several factors have placed dangerous demands on our energy supplies. First, the amount of energy consumed **(in, *per*, were)** person has been increasing exponentially. Our **(*ancestors*, slept, energy)** consumed about 10 000 000 J or 10 MJ of **(direction, dangerous, *energy*)** per person per day. Today, the **(*global*, exactly, generating)** average for energy consumption per person **(*per*, near, fossil)** day is 180 MJ. This figure rises **(coal, on, *for*)** people in industrialized countries, who consume, **(fuel, *on*, over)** average, 1000 MJ per person per day!

(*Second*, Natural, Environment), the world population is also growing **(consuming, *exponentially*, confusingly)**. Up until the 1950s, the growth **(over, power, *of*)** the world's population was steady but **(*slow*, wild, supplies)**. In the 1950s, there were huge **(televisions, *innovations*, significant)** in medicine, such as the development **(under, learnt, *of*)** vaccines and antibiotics. Better public health **(used, *increased*, secondary)** the average life span. There were **(produced, sternly, *also*)** technological innovations which improved people's lifestyles. **(Next, *As*, People)** a result of all these developments, the (***world*, car, now)**

---

[2] Passage taken from *Addison Wesley Science 10* (Sandner et al., 2004, p. 223).

[3] Correct answers are italicized.

population has been increasing at an **(infinite, *exponential*, electric)** rate. Currently, there are over 6 billion **(water, span, *people*)** on the planet. So, there are (***more*****, hydro, huge)** people consuming more energy.

Third, many **(stations, *societies*, significant)** now use non-renewable energy sources rather **(*than*, hate, on)** renewable sources as the primary source **(*of*, over, life)** energy. Until the 1950s in Alberta, **(*most*, have, result)** power stations produced electrical energy from **(state, old, *hydro*)** power, or flowing water. When large **(*coal*, shell, place)** reserves were discovered near Edmonton, it **(*then*, strain, before)** became cheaper to build coal-fired generating **(environment, *stations*, hover)** than hydro-electric dams. Today, only 5% **(rough, *of*, above)** electricity generation in Alberta is from **(successfully, *hydro*, leaking)** power and 85% is from coal. The **(reflect, fossil, *other*)** 10% is from natural gas.

Our **(*dependence*, refusal, exemplify)** on fossil fuels has placed a **(plough, *strain*, burning)** on existing supplies. It has also **(*had*, danger, are)** significant effects on the environment. The **(permanent, reveal, *search*)** for and extraction of fossil fuels **(give, *have*, climate)** resulted in damage to ecosystems. For **(*example*, fashion, only)**, drilling for oil in the muskeg **(*of*, after, demand)** northern Alberta has left scars on the **(color, *landscape*, hungry)**. There is also the danger of **(overall, *oil*, bees)** leaking from wells, polluting the surrounding **(cheaper, *area*, candle)**. Even though fossil fuels are relatively **(panicky, supplies, *cheap*)** and convenient sources of energy, burning **(steady, her, *them*)** may damage our environment permanently. Emissions **(rain, under, *from*)** the combustion of fossil fuels contain **(overall, *greenhouse*, library)** gases and oxides. Greenhouse gases contribute **(from, a, *to*)** climate change and oxides contribute to **(water, *acid*, cheaper)** rain.

The overall effect of an **(*increasing*, running, danger)** population, an increasing demand for energy, **(so, has, *and*)** the shift toward the exploitation of **(effect, jagged, *non-renewable*)** energy sources is having far-reaching consequences.

Silent Reading Fluency Test – Form B[4]

Directions: Below is a passage about energy conversions. *When the teacher tells you to begin, silently read the following passage. Some of the words are replaced by three words. As you read, circle the one that makes the most sense. Only one word is correct. You will have one minute to read as much as you can.*[5]

Example: Learning about science (about, is, play) fun.

In general, one of several things might happen as a consequence of energy being converted from one form to another.

Motion is the most obvious evidence **(*that*, it, ball)** an energy conversion has happened. When a **(drawn, *pitcher*, silky)** throws a ball, her arm does **(*work*, gained, string)**, which becomes the ball's kinetic energy.

(*Climbing*, sleeping, pulls) stairs is a less obvious evidence **(*of*, and, by)** energy. A diver gains gravitational potential **(chemical, stones, *energy*)** by climbing the stairs of a **(singing, become, *diving*)** tower. When he's on the platform, **(they, *he*, change)** has raised his position relative to the **(again, arrow, *surface*)** of the water. Whenever something is **(stairs, *raised*, drunk)** above the surface of Earth, this ***(react, being,*** **change*)*** *in position* is evidence of gravitational **(back, electric, *potential*)** energy.

A *change in shape* is **(very, *also*, pot)** evidence that an object has undergone a **(pole, *change*, drive)** in energy. The drawn bow in Figure B2.15 **(*has*, evidence, crush)** gained elastic potential energy as the **(*archer*, snake, releases)** pulls the bow string back, changing **(my, *its*, mid)** shape. When she releases the bow, **(they, drawn, *it*)** will change its shape again as the

---

[4] Passage taken from *Addison Wesley Science 10* (Sandner et al., 2004, p. 190-191).

[5] Correct answers are italicized.

(*elastic*, basic, obvious) potential energy changes to kinetic energy **(*of*, are, as)** the released arrow. A stretched elastic **(*band*, occurs, glass)** or a pole vaulter's pole mid **(*jump*, stream, strikes)** has gained elastic potential energy.

For a **(raised, table, *pot*)** of water on the stove, a **(change*, kinetic, coin)*** *in temperature* is evidence of energy **(*transfer*, bus, arrow)**. Energy is being transferred from the **(*hot*, new, tower)** stove to the cooler pot and **(chair, *water*, above)**. The pot and water are gaining **(fat, sun, *heat*)**. Heat is the transfer of kinetic **(*energy*, another, climbing)** of the particles in one substance **(*to*, under, water)** another; in this case from the **(solar, bridge, *element*)** to the pot and water.

The hydrogen–hydrogen **(*nuclear*, physical, reflected)** fusion reaction that occurs at the **(piece, *centre*, react)** of the Sun releases tremendous amounts **(over, most, *of*)** solar energy that travels to Earth **(*as*, until, life)** electromagnetic waves. When this radiation strikes Earth, **(*it*, you, water)** is either absorbed by Earth or **(carbon, *reflected*, betrayed)** back into space. When light energy **(*from*, with, converts)** the Sun strikes the chlorophyll in **(*plants*, houses, provides)** a chemical reaction, photosynthesis, occurs that **(sweeps, functions, *converts*)** carbon dioxide and water into glucose **(but, case, *and*)** oxygen. The glucose contains chemical potential **(process, *energy*, releases)**. When animals eat plants, this chemical **(excellent, *potential*, dioxide)** energy in the glucose is released **(*through*, either, above)** the process of respiration in the **(*animals'*, water's, strikes)** bodies. Glucose (sugar) from food reacts **(*with*, upon, also)** oxygen in animal cells to produce **(travels, *carbon*, molecule)** dioxide and water. The energy released **(except, necessary, *during*)** respiration, in the form of adenosine triphosphate (ATP), **(*provides*, closes, tremendous)** the energy necessary for the animal **(*to*, off, less)** carry out life functions. It also **(endangers, *produces*, chemical)** heat.

Silent Reading Fluency Test – Form C[6]

Directions: Below is a passage about the laws of thermodynamics. *When the teacher tells you to begin, silently read the following passage. Some of the words are replaced by three words. As you read, circle the one that makes the most sense. Only one word is correct. You will have one minute to read as much as you can.*[7]

Example: Learning about science (about, is, play) fun.

It is important to distinguish heat from work. Work involves the movement of matter **(*from*, until, done)** one location to another, whereas heat **(make, *is*, two)** a transfer of thermal energy from **(total, six, *one*)** location to another. Both heat and **(*work*, graphs, never)** can affect systems.

The energy of a **(distance, *system*, whereas)** can be increased in two ways. **(*Either*, Both, Into)** heat can be added to a **(*system*, planet, done)** from the surroundings, or work can **(*be*, ways, try)** done on a system by its **(considered, location, *surroundings*)**. Work done on the system by the **(*surroundings*, decreased, weather)** is considered positive work because the **(math, *energy*, added)** of the system increases.

Similarly, the **(*energy*, from, atoms)** of a system can decrease in **(*two*, really, one)** ways. Either heat can flow out **(*of*, than, just)** a system to its surroundings, or **(terms, equal, *work*)** can be done by a system **(*on*, as, law)** its surroundings. Work done by a **(lottery, cannot, *system*)** on its surroundings is considered negative **(*work*, light, happens)** because the energy of the system **(*will*, other, need)** decrease.

The law of conservation of **(interview, similarly, *energy*)** can be stated in more general **(*terms*, magnets, added)**: Energy cannot be created or destroyed. **(*It*, We, On)** can only be

---

[6] Passage taken from *Addison Wesley Science 10* (Sandner et al., 2004, p. 199-200).

[7] Correct answers are italicized.

transformed from one **(desire, positive, *form*)** to another, and the total amount **(*of*, by, can)** energy never changes. The first law **(past, one, *of*)** thermodynamics is really just a restatement **(one, *of*, past)** the law of conservation of energy, **(*except*, before, systems)** one of the forms of energy **(cooperate, negative, *involved*)** is heat. This law states that the **(organic, *total*, flow)** energy, including heat, in a system **(*and*, most, nor)** its surroundings remains constant. Whenever heat **(stop, what, *is*)** added to a system, it transforms **(minus, *into*, rub)** an equal amount of some other **(*form*, pill, contact)** of energy. This law is supported **(for, part, *by*)** Joule's experiments.

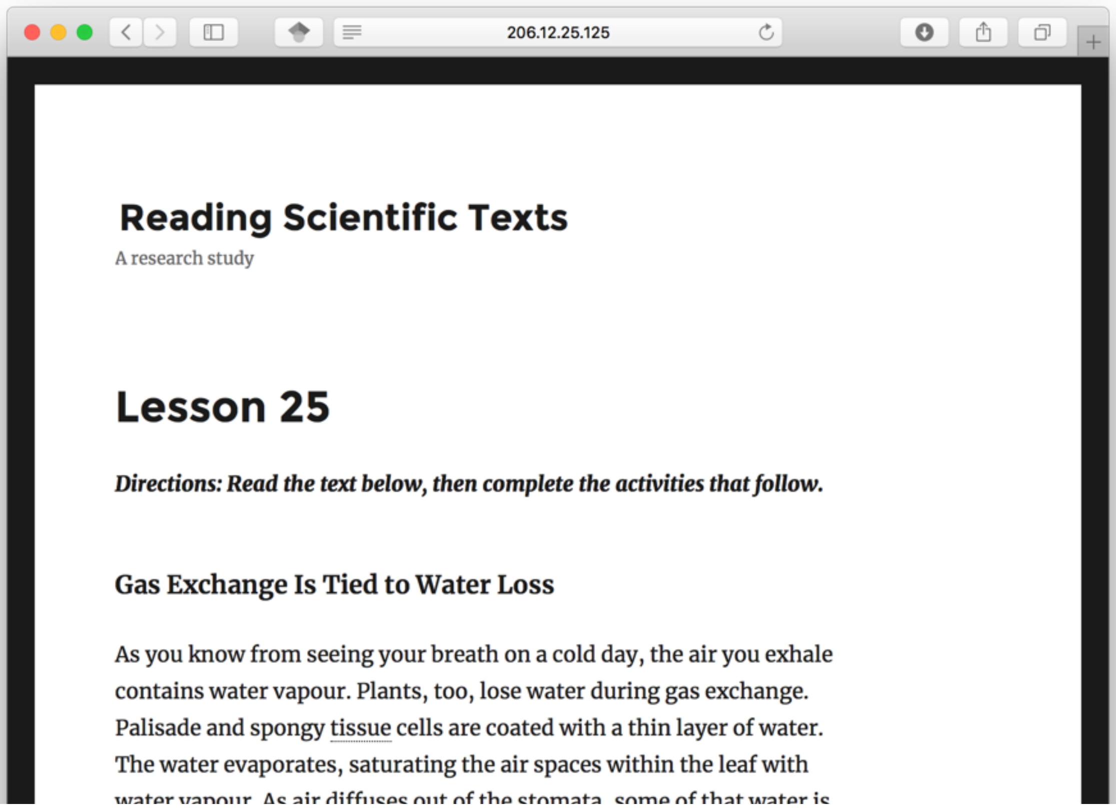
When heat is added **(*to*, since, amount)** a system, some of the energy **(stops, *goes*, that)** into increasing the internal energy of the **(involve, office, *system*)**. This increases the temperature. Some of the **(*energy*, wind, plus)** is used to move parts or **(*to*, off, this)** do work on the system. This **(though, *increases*, compares)** the mechanical energy.

This is what **(programs, *happens*, friction)** in most real-world situations. Most machines **(*involve*, decide, system)** many moving parts that are in **(effect, same, *contact*)** with one another. These parts rub **(*together*, apart, internal)** and this friction produces heat. Even **(for, *though*, close)** some energy is lost as heat, the **(*law*, same, light)** of conservation of energy or the **(*first*, tenth, amount)** law of thermodynamics still applies: The **(never, treat, *amount*)** of heat put into a system **(may, some, *must*)** equal the amount of mechanical energy **(among, reality, *plus*)** heat lost by the system. Theoretically, the **(beaker, *system*, transforms)** could gain the same amount of **(independent, destroyed, *mechanical*)** energy as the heat input energy. **(*But*, These, Because)** in reality, the mechanical energy gained **(later, *never*, many)** comes close to the theoretical maximum **(*because*, unless, situations)** most of the input energy is **(test, when, *lost*)** from the system as heat.

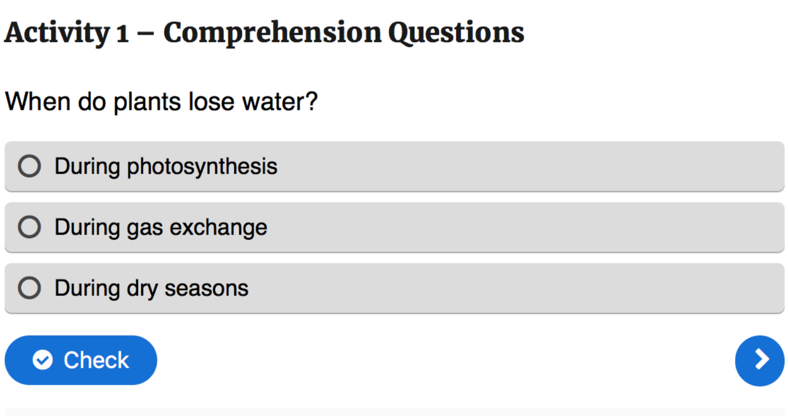
Appendix E

Online Modules – Sample Passage and Activities

Screenshot of a lesson (common across intervention conditions):



Sample Follow-up Activities – Traditional Group



Activity 2 – Science Vocabulary

Drag the words into the correct boxes

1) a cluster of similar cells that share the same structure and function (e.g., muscle tissue or vascular tissue) ______

2) a small opening in the epidermal layer of a leaf that allows gases in and out ______

3) the evaporation of water from leaves ______

4) cells located on either side of a stoma, that change shape to open or close the opening in order to allow gases in and out ______

5) the process by which plants use light energy to produce food in the form of carbohydrates ______

6) the movement of water molecules across a membrane from an area of higher concentration of water to an area of lower concentration of water ______

7) having the ability to return to its original form after its shape has been distorted ______

8) the boundary around a cell that separates the cell interior from the environment ______

9) the process by which cells obtain energy by breaking down glucose in the presence of oxygen ______

10) water pressure within plant cells that allows them to remain rigid ______

turgor pressure | cell membrane | osmosis | tissue | transpiration | guard cells | photosynthesis | elastic | cellular respiration | stoma

Check

Sample Follow-up Activities – Vocabulary Group

Activity – Academic Vocabulary 1

The academic words below are found in the text that you just read above. Try to match the words with their definitions. If you don't know a word, find it in the text, try to understand its meaning, then try matching.

Drag the words into the correct boxes

1) to happen
2) one of the parts that something such as an object or place is divided into
3) a thing, place, activity etc that you get something from
4) to use time, energy, goods, etc
5) to control an activity or process
6) an amount or piece of a material or substance that covers a surface
7) a short statement that gives the main information about something, without giving all the details
8) to gradually change your behaviour and attitudes in order to be successful in a new situation
9) to reduce something that is difficult, dangerous, or unpleasant to the smallest possible amount or degree
10) to make something return to its former state or condition
11) stiff and not moving or bending

summary | layer | rigid | restore | consume | regulate
occur | adapt | source | minimize | section

Sample Follow-up Activities – Grammar Group

Activity – The Grammar of Science

Match the "regular" English with academic English.

1) Water evaporates = ______

2) Water evaporates from leaves = ______

3) how hard water pushes against a surface = ______

4) where water comes from = ______

5) gas going in and out of something = ______

6) the water is saved by plants for use later = ______

7) a lot of something happening = ______ of something

high rates | water reserves | gas exchange | water source | water pressure | transpiration | evaporation

Check

Look at the bolded (dark) words in the first sentence. Which words in the second sentence mean the same thing as the bolded (dark) words?

As air diffuses out of the stomata, **some of that water is lost.** The evaporation of water from leaves is called transpiration.

Check

What is evaporating?

The evaporation of water from leaves is called transpiration.

Check

Which words tell you where evaporation is happening?

The evaporation of water from leaves is called transpiration.

Check

In one word, what is transpiration?

The evaporation of water from leaves is called transpiration.

Check

Transpiration is special. It is like transpiration, but a different a kind. Which words tell you what kind of evaporation?

The evaporation of water from leaves is called transpiration.

Check

What might seem to be in danger of drying out and dying? (1 word)

Plants that lose so much water through transpiration might seem to be in danger of drying out and dying

Check

What kind of plants might seem to be in danger of drying out and dying?

Plants that lose so much water through transpiration might seem to be in danger of drying out and dying

Check

In short, what controls what? (2 words)

_____ controls ______ .

The size of the stomata controls the amount of gas exchange and transpiration.

Check

What closes the stoma? (2 words)

As the amount of water in the guard cells decreases, the cells deflate and change shape again, as shown in Figure 9.8. This action closes the stoma.

Check

What does "this action" in the second sentence refer to? (5 words)

As the amount of water in the guard cells decreases, the cells deflate and change shape again, as shown in Figure 9.8. *This action* closes the stoma.

Check

Appendix F

Participant Information Letters and Consent Forms

Information Letter and Consent Form for Teachers

Title of Study: Improving Reading Comprehension of Scientific Texts: A Comparison of Traditional, Academic Vocabulary, and Grammar-based Online Modules for K-12 English Language Learners.

Research Investigator:
Kent Lee
Department of Educational Psychology
6-102 Education North
University of Alberta
Edmonton, AB, T6G 2G5
kent.lee@ualberta.ca

Supervisor:
Dr. Marilyn Abbott
Department of Educational Psychology
6-102 Education North
University of Alberta
Edmonton, AB, T6G 2G5
marilyn.abbott@ualberta.ca
780-492-9090

You are being asked to be in this study because you are a teacher of English language learners (ELLs) at a high school in Alberta. Your principal has recommended you as a potential participant. My goal is to investigate the effectiveness of a reading intervention that may improve ELLs' reading comprehension. The results of this study will be used in support of my thesis. This research is being funded by a Doctoral Fellowship from the Social Sciences and Humanities Research Council.

I would like to present your ELLs with the opportunity to use an internet-based reading tutorial that I have created. Students will be randomly assigned to one of three tutorials, all focusing on the readings from a Science 10 textbook. Because I want to compare the effectiveness of the tutorials, students who participate will complete a reading comprehension test, a grammar test, and a reading fluency test. This battery of tests will take 1.5 hours, and will be administered before and after the intervention. In addition, I will ask students to provide me with some personal information.

With your permission, I would like your ELLs to complete the online tutorial during regular school hours. They will use the tutorial for 30 minutes a day, five days a week for six weeks. All that I ask of you is to provide the time for the students to use the tutorial, ensure that students are on-task, and help the less computer-savvy ELLs with using it. Please refrain from providing students with any assistance with the content; it is the efficacy of the modules that I am testing. At the end of the six weeks, I will also ask the students to complete a short questionnaire about their experiences with the intervention. Students who do not wish to participate in the study may still use the online tutorial, engage in extensive reading, or self-study, but no data will be collected from them.

I hope that the results from this study will help to identify teaching techniques that will help ELLs improve in their ability to read and comprehend informational texts. Students will gain extra practice in reading science prose. Ultimately, I hope that the tutorial will play a role in helping improve the academic achievement of ELLs. At the very least, the information from this study will hopefully contribute to the design of effective reading interventions for ELLs.

You are under no obligation to participate in this study. Participation is completely voluntary. There is no financial cost associated with being involved. You can opt out at anytime

The information that I collect will be used in presentations and research articles. Participants' anonymity is guaranteed. The data will be kept confidential. Only the researcher, supervisor, and the Research Ethics Committee will be allowed access to it. Recorded comments will be anonymous; no names will be used in any data files or reports. No information will be provided that could identify individuals or schools.

The information collected will be kept in a secure location for 5 years after the project is completed. Anonymous electronic data will be password protected. Consent forms will then be shredded and electronic data will be destroyed. You may request a report of the research results at the end of this form.

If you have any further questions regarding this study, please do not hesitate to contact Kent Lee (kent.lee@ualberta.ca) or Marilyn Abbott (mabbott@ualberta.ca – 780-492-9090). The plan for this study has been reviewed for its adherence to ethical guidelines by the Research Ethics Board at the University of Alberta and your school district. If you have concerns about this study, you may contact the Research Ethics Board at 780-492-2615. This office has no direct involvement with this study.

I have read the information provided above, and I have been given a chance to ask questions.
My questions have been answered to my satisfaction and I agree to participate in this study.
I have been given a copy of this fully signed form.

Signature ____

Name ____

Date ____

Please provide your email address if you would like a summary of the research results:

____

INFORMATION LETTER and CONSENT FORM for STUDENTS

Study Title: Improving Reading Comprehension of Scientific Texts: A Comparison of Traditional, Academic Vocabulary, and Grammar-based Online Modules for K-12 English Language Learners.

Research Investigator:
Kent Lee
Department of Educational Psychology
6-102 Education North
University of Alberta
Edmonton, AB, T6G 2G5
kent.lee@ualberta.ca

Supervisor:
Dr. Marilyn Abbott
Department of Educational Psychology
6-102 Education North
University of Alberta
Edmonton, AB, T6G 2G5
marilyn.abbott@ualberta.ca
780-492-9090

Background
I am inviting you to join my research study because you are learning English as a second/additional language. I have created a website that may help you become better at reading in English. I want to see if the website can really help students to read better. The results of this study will be used in support of my thesis. This research is funded by the Government of Canada's Social Sciences and Humanities Research Council.

Purpose
Being a good reader is important to success in high school. The readings in classes such as science can be difficult. The language used in science writing is often different from everyday English. For example, there is a lot of difficult vocabulary, and the grammar is complex. So, I have created a website to give students practice in reading the language used in science. However, the website is new, so it needs to be tested to see if it works and if it leads to student improvements. I hope that the results from this study will help us learn about new ways to teach English language learners to read.

Study Procedures
If you agree to join this study, the following things will happen:

- I will ask you to tell me a bit about yourself by answering a questionnaire.
- I will also give you 3 tests. Two are on reading comprehension and reading fluency. One is on grammar.
- Your test results will be kept secret from your teachers. So the results will NOT affect your grades.
- You will practice reading from a website that I made. Your school will give you time to use the website [30 minutes a day, 5 days a week for 6 weeks].
- After six weeks, I will give you the reading and grammar tests again. I want to see how much my website helped you improve. I will also ask you to complete a questionnaire about my website.

Benefits
I cannot guarantee that being in this study will help, but I expect that using my website will help improve your reading by giving you more practice. We may learn something that will help other English language learners in the future. This study will help us learn more about ways teachers can help English language learners understand the language used in science.

Risk
There are no risks to participating in this study that I can see. If I learn anything during the study that might make you change your mind about participating, I will tell you right away.

Voluntary Participation
It is your choice to join this study. You can say yes now, and change your mind later. If you want to quit, just tell me that you want to quit. It's also okay to say no. No one will be angry if you do not want to be in the study. You can ask me questions at any time. If you want to quit, the information that I collected on you will be erased. You can quit before you write the last test. Once you have written the last test, I will use the information that I have collected from you.

Confidentiality & Anonymity
The information that I collected will be used for my dissertation, research articles, and presentations. However, your name and your school's will be kept a secret and will NOT be used. The information on you will be kept confidential – only my supervisor and I will have access to it. Everything you provide me will be anonymous.

The information collected will be kept in a secure place for 5 years after the project is completed. Electronic data will be password protected. Papers will be shredded and electronic data will be destroyed. You may request a copy of the research results at the end of this form.

Further Information
You can ask any questions that you may have about this research study. If you have a question later, you can email or call my supervisor, Marilyn Abbott (Marilyn.abbott@ualberta.ca 780-492-9090) and me (kent.lee@ualberta.ca).

The plan for this study has been reviewed for its adherence to ethical guidelines by a Research Ethics Board at the University of Alberta. For questions regarding participant rights and ethical conduct of research, contact the Research Ethics Office at (780) 492-2615."

Consent Statement
I have read this form and the research study has been explained to me. I have been given the opportunity to ask questions and my questions have been answered. If I have additional questions, I have been told whom to contact. I agree to participate in the research study described above and will receive a copy of this consent form. I will receive a copy of this consent form after I sign it.

____
Participant's (Student) Name (printed)

____ ________________
Participant's (Parent) Name (printed) and Signature Date

____ ________________
Name (printed) and Signature of Person Obtaining Consent Date

Please provide your email address if you would like a summary of the research results:

Your email address: ______________________________________