

Georgia State University

ScholarWorks @ Georgia State University

Middle and Secondary Education Dissertations Department of Middle and Secondary Education

Spring 5-12-2023

Traces to Entrenchment: A Mixed Methods Study Examining the Use of Reading and Writing to Facilitate Lexical Sequence Acquisition Among Bilingual Learners and Factors Influencing Their Language Learning

Andrea J. Lofgren
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/mse_diss

Recommended Citation

Lofgren, Andrea J., "Traces to Entrenchment: A Mixed Methods Study Examining the Use of Reading and Writing to Facilitate Lexical Sequence Acquisition Among Bilingual Learners and Factors Influencing Their Language Learning." Dissertation, Georgia State University, 2023.
doi: <https://doi.org/10.57709/35325546>

This Dissertation is brought to you for free and open access by the Department of Middle and Secondary Education at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Middle and Secondary Education Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

ACCEPTANCE

This dissertation, TRACES TO ENTRENCHMENT: A MIXED METHODS STUDY EXAMINING THE USE OF READING AND WRITING TO FACILITATE LEXICAL SEQUENCE ACQUISITION AMONG BILINGUAL LEARNERS AND FACTORS INFLUENCING THEIR LANGUAGE LEARNING, by ANDREA J. LOFGREN, was prepared under the direction of the candidate's dissertation advisory committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education & Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

Ewa McGrail, Ph.D.
Committee Chair

Jennifer Esposito Norris, Ph.D.
Committee Member

Gertrude Tinker Sachs, Ph.D.
Committee Member

Diane Belcher, Ph.D.
Committee Member

Robert Clayton Hendrick, Ph.D.
Committee Member

Date

Gertrude Tinker Sachs, Ph.D.
Chairperson, Department of Middle and
Secondary Education

Paul A. Alberto, Ph.D.
Dean, College of Education &
Human Development

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education & Human Development's Director of Graduate Studies, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

Andrea J. Lofgren

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Andrea J. Lofgren
Department of Middle and Secondary Education
College of Education & Human Development
Georgia State University

The director of this dissertation is:

Dr. Ewa McGrail
Department of Middle and Secondary Education
College of Education & Human Development
Georgia State University
Atlanta, GA 30303

CURRICULM VITAE

Andrea J. Lofgren
1920 Sweetgum Trail
Cumming, GA 30041

Formal Education

- B.S.S. in Liberal Arts (Ohio University, 2002)
- M.Ed. in English Education (Georgia State University, 2005)
- Ph.D in Teaching and Learning (Georgia State University, 2023)

Selected Teaching Experience

- Primary Instructor, EDRD 2120 *Sociocultural Perspectives on Diversity* [hybrid], Georgia State University (Atlanta, GA, 2023)
- Instructor of Record, EDRD 3600/7210 *Family Literacy and Community Learning* [online], Georgia State University (Atlanta, GA, 2022)
- Instructor of Record, EDRD TSLE *Language and Discourse in Professional Contexts* [online], Georgia State University (Atlanta, GA, 2019)
- Primary Instructor, PERS 2001 *Teaching in Today's Society*, Georgia State University [in person], Georgia State University (Atlanta, GA, 2018)
- Part-time Instructor, Corners Academy (Dunwoody, GA, 2023)
- English and ESOL Teacher, Cross Keys High School (Atlanta, GA, 2008-2017)
- English and ESOL Teacher, Avondale High School (Decatur, GA 2005-2008)
- English Teacher, Alghanim Bilingual School (Salwa, Kuwait 2002-2003)

Selected Research Experience

- Doctoral Graduate Research Assistant for Dr. Omer Ari (2022-2023)
- Doctoral Graduate Research Assistant for Dr. Ewa McGrail (2018)

Selected Service to Department, College, and Profession

- Prepared and wrote program proposal, BA in Human Development for University System of Georgia (2022)
- Co-wrote, co-edited, conducted data analysis for MSE Self-study Report for SACSCOC, Georgia State University (2019)
- Journal Manager and Editorial Assistant, *Ubiquity* (2020-2023)

Selected Publications

- Ari, O., Hageman, K., & **Lofgren, A.** (under review). Effects of quizzes on consolidation of undergraduates' knowledge gains in an online Interdisciplinary Studies course. Submitted to *Active Learning in Higher Education* [12 pages.]
- **Lofgren, A.**, & Colquitt, E. (under review). Inequity by the numbers: Gender and tenure differentials and quantitative promotion policies. Submitted to *Gender and Education* [22 pages.]
- **Lofgren, A.** (revise/resubmit). College students and class attendance: How poverty and illbeing affect student success through attendance policies. Submitted to *College Teaching* [14 pages.]
- **Lofgren, A.** (2022). Unraveling contradictions: Which glosses facilitate reading comprehension among ELLs, and why? *Journal of Language Teaching and Research*, 13(1), 1-11. <https://doi.org/10.17507/jltr.1301.01>
- **Jeddi, A.** (2018). Anti-immigration sentiments and the true benefits of inclusion. *GATESOL in Action*, 28(1), 1-6. <https://doi.org/10.52242/giaj.v28i1.81>

Selected Presentations

- Ari, O., Hageman, K., & **Lofgren, A.** (2023, April 11-16). Using SPARKS to improve learning outcomes of undergraduate students in an online course [Poster presentation]. Annual Meeting of the American Educational Research Association, Chicago, IL, United States.
- Colquitt, E., & **Lofgren, A.** (2021, April 9). Gendered tenure differentials in academia [Paper presentation]. LSGSA Graduate Research Conference, Atlanta, GA, United States.
- **Jeddi, A.**, & Trinh, E. T. (2019, April 23). The Gadfly: Examining student realities and instructor expectations in an urban school [Paper presentation]. The 2019 Cultural Competency Summit, Atlanta, GA, United States.
- **Jeddi, A.** (2015, August). Using TIENET reports to improve student achievement. Dekalb County ESOL Conference, Atlanta, GA, United States.

Awards/Honors

- Golden Key Honor Society (2021)
- Certificate of Exemplary Achievement, Mastery of Online Teaching (2019)
- Teacher of the Year, Cross Keys High School (2016)
- Outstanding Contribution Award, Cross Keys High School (2015)
- 4H Rookie of the Year Award, 4H Club (2011)
- Outstanding Service Award, International Research Exchange Board (2011)
- Certificate of Excellence, MSIT Initial Teacher Preparation (2005)

**TRACES TO ENTRENCHMENT: A MIXED METHODS STUDY EXAMINING THE
USE OF READING AND WRITING TO FACILITATE LEXICAL SEQUENCE
ACQUISITION AMONG BILINGUAL LEARNERS AND FACTORS INFLUENCING
THEIR LANGUAGE LEARNING**

by

ANDREA J. LOFGREN

Under the direction of Dr. Ewa McGrail

ABSTRACT

Informed by entrenchment and usage-based theories, this mixed methods study, modeled after Riazi's mixed methods fully integrated design, examined how reading and writing may be used to facilitate lexical sequence acquisition for high school bilingual learners. The study compared the efficacy of two interventions on increasing the depth of knowledge of lexical sequences: a reading intervention requiring the reading of short informational texts embedded with eight target lexical sequences and a writing intervention requiring participants to follow the readings with the intentional use of target sequences in writing. A repeated measures factorial ANOVA found no impact of gender or first language on results; both groups made statistically significant gains on depth of knowledge during both interventions with large effect sizes, and they retained gains three to four weeks later. However, reading was found to be more impactful for one group, whereas reading and writing was found to be more impactful for the second group. Subsequent analyses of written responses by participants coupled with a repeated

measures ANOVA measuring the acquisition of individual sequences demonstrated that participants were more likely to acquire and use some lexical sequences than others. Data from semi-structured interviews from 12 participants were analyzed through the lens of Anthias' multilevel model of intersectionality to determine factors influencing intervention results as well as English learning generally. Results of this qualitative strand revealed that the requirement to read twice weekly led reading to become a habit for some participants, that text interest impacted some participants' motivation to comprehend the text, and that interlinear glossing (providing synonyms above target sequences) was helpful for reading comprehension. Gender and L1 were found to influence learning through the long-lasting impact of bullying, which often led females and those with perceived accents to refrain from speaking English even years after bullying occurred. Other salient findings suggest that the employment of high school students may be underreported and that implicit teacher bias may impact the schooling of some bilingual learners. One LGBTQ bilingual learner narrative provided a positive example of empowerment despite the bullying of others. Several recommendations for research, policy, and teaching practices are discussed.

INDEX WORDS: lexical sequence acquisition, vocabulary acquisition, depth of vocabulary knowledge, bilingual learners, bullying, multilevel model of intersectionality, entrenchment theory, usage-based theory, ELLs

TRACES TO ENTRENCHMENT: A MIXED METHODS STUDY EXAMINING THE USE
OF READING AND WRITING TO FACILITATE LEXICAL SEQUENCE ACQUISITION
AMONG BILINGUAL LEARNERS AND FACTORS INFLUENCING THEIR LANGUAGE
LEARNING

by

ANDREA J. LOFGREN

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

Teaching and Learning (Language and Literacy Education)

in

Department of Middle and Secondary Education

in

the College of Education & Human Development

Georgia State University

Atlanta, GA

2023

Copyright by
Andrea Lofgren
2023

ACKNOWLEDGEMENTS

I would like to thank those who have served on my committee, including Dr. Chris Oshima, Dr. Jennifer Esposito Norris, Dr. Diane Belcher, and Dr. Robert Clayton Hendrick, for the invaluable guidance they have provided to me during every step of this project, and I would like to thank Dr. Omer Ari for his invaluable mentorship during my GRA work with him. I especially owe a great deal of gratitude to my committee chairperson, Dr. Ewa McGrail, and to Dr. Gertrude Tinker Sachs, Chair of the Department of Middle and Secondary Education. Thank you, Dr. McGrail, for setting high standards for my work and for your cheerful and kind encouragement along the way. Dr. Tinker Sachs, I will forever be grateful to you for the trust you placed in me to assist our department with high stakes accreditation and assessment tasks. Your trust inspired confidence in me that has reshaped my life, and my newfound skills, abilities, and knowledge has led me to develop a clear vision for my future work life. So have the gentle lessons you have provided about what social justice means and how to fight for it effectively.

I would also like to thank my family for their encouragement and help along the way. Thank you, Yousef, for your revision suggestions and your panache with words, which has so greatly improved my work. Thank you to Haleh, Reese, and Riley for expanding my understanding of gender and for teaching me about your generation. Thank you, Doug, for your boundless love and encouragement. Thank you, Mom and Dad, for instilling the values in me that form the foundation of this project. Thank you to both of my sisters, Dr. Teresa Dean and Dr. Jennifer Dean, for showing me possibilities.

Table of Contents

LIST OF TABLES	vi
LIST OF FIGURES	vii
1 THE PROBLEM.....	1
Vocabulary as a Predictor of Reading and Writing Proficiencies	1
Vocabulary as a Predictor of Educational Success.....	3
Vocabulary as a Predictor of Further Vocabulary Acquisition	4
Definitions.....	6
The Impact of L1 on Second Language Vocabulary Acquisition.....	11
The Impact of Gender on Vocabulary Acquisition.....	15
Purpose of the Study	18
<i>Research Questions</i>	<i>19</i>
Overview of the Study	19
2 REVIEW OF THE LITERATURE	21
Section 1: Scope of Literature Review	21
Section 2: Theoretical Framework.....	25
<i>Entrenchment</i>	<i>26</i>
<i>Usage-based Theory</i>	<i>29</i>
Section 3: Empirical Research.....	31
<i>Vocabulary Development</i>	<i>31</i>
<i>Vocabulary Acquisition Through Reading.....</i>	<i>45</i>
<i>Vocabulary Acquisition Through Writing.....</i>	<i>61</i>
<i>Summary of Vocabulary Studies.....</i>	<i>77</i>
3: METHODOLOGY	81
Researcher Subjectivities	81

Research Design	84
<i>Defining Mixed Methods Methodology</i>	<i>86</i>
<i>A Rationale for Mixed Methods Methodology</i>	<i>87</i>
<i>Mixed Methods as Fully Integrated.....</i>	<i>88</i>
Pragmatism as a Guiding Paradigm	89
Research Challenges During COVID-19	92
Study Procedures	93
Quantitative Strand	95
<i>Participants.....</i>	<i>95</i>
<i>Instruments and Materials</i>	<i>101</i>
<i>Data Collection</i>	<i>106</i>
<i>Quantitative Data Analysis.....</i>	<i>110</i>
Qualitative Strand.....	112
<i>Participants</i>	<i>113</i>
<i>Data Collection</i>	<i>119</i>
<i>Qualitative Data Analysis.....</i>	<i>120</i>
Mixed Methods as Intergrative	134
4 FINDINGS	136
Quantitative Findings.....	136
<i>Within-Subjects Effects</i>	<i>138</i>
<i>Between-Subject Effects</i>	<i>139</i>
<i>Interaction Effects</i>	<i>140</i>
<i>Subsequent Repeated Measures Analysis.....</i>	<i>146</i>
<i>Summary of Quantitative Findings</i>	<i>149</i>
Qualitative Findings	149
<i>Theme 2: Claimed and Unclaimed Responsibilities.....</i>	<i>154</i>
<i>Theme 4: The Life-Altering Impact of Bullying</i>	<i>161</i>
<i>Summary of Qualitative Findings.....</i>	<i>164</i>

Mixed Methods Integration	166
5 DISCUSSION	170
Factors Influencing Language Learning	177
<i>Bullying</i>	177
<i>Claimed and Unclaimed Responsibilities</i>	180
Study Limitations.....	184
Recommendations for Policy and Practice.....	185
<i>Vocabulary Teaching</i>	186
<i>Combating Bullying</i>	189
Recommendations for Further Research	192
References.....	195
APPENDICES.....	2622

LIST OF TABLES

Table 1 Lexical Sequence Targets	102
Table 2 Passage Readability and Schedule of Lexical Sequence Dispersion	104
Table 3 Testing Schedule	108
Table 4 Comparison of Demographics, Participants of Both Strands	114
Table 5 Descriptive Statistics—Reading Intervention	137
Table 6 Descriptive Statistics—Writing Intervention	138
Table 7 Between-Subject Effects: Group, L1, and Gender	139
Table 8 Usage of Sequences During Writing Intervention by Group	144
Table 9 Usage Scores by Lexical Sequence	145
Table 10 Lexical Sequence Gains, Highest to Lowest, by Intervention	148

LIST OF FIGURES

Figure 1. Literature Review Study Focus	45
Figure 2. Literature Review Participants: Learning Contexts	46
Figure 3. Literature Review Participants: Gender	46
Figure 4. Literature Review Participants: Proficiency Level	79
Figure 5. Mixed Methods Fulfilled Integrated Design Model	86
Figure 6. Evolution of the Mixed Methods Study	94
Figure 7. Participant Self-Reported L1s	99
Figure 8. Study Participants: Ethnicity as Described by Participants	100
Figure 9. Study Participants: Race as Described by Participants	101
Figure 10. Testing of Lexical Sequences by Group	107
Figure 11. Race as Described by Participants	115
Figure 12. Ethnicity as Described by Participants	115
Figure 13. Interview Matrix	117
Figure 14: Initial Codes and Descriptions	121
Figure 15. Concept Map Linking Initial Codes	123
Figure 16. Multilevel Model of Intersectionality	124
Figure 17: Modified Codes and Descriptions	129
Figure 18. Alignment of Codes, Categories, and Themes	132
Figure 19. Group and Intervention Interactions	140
Figure 20. Group A Scores by Intervention and Test	141
Figure 21. Group B Scores by Intervention and Test	142
Figure 22. Test Scores by Lexical Sequence	146

Figure 23. Intervention and Test Results, Outlier Sequence Removed	147
Figure 24. Gendered Aspect of Teacher Critique	159
Figure 25. Mixed Methods Findings	167

1 THE PROBLEM

As the atoms of language, vocabulary forms the core of language meaning, making its acquisition an indispensable part, not only of language learning (Nation, 2001; Schmitt, 2010) but also of the “construction of human reality” (Dakhi & Nur Firtria, 2019, p. 17). A review by Alderson and Banajee (2001) found academic vocabulary correlated to success in all academic language abilities. It is a predictor of language testing success (Alderson, 2005). It has been correlated to university grade point averages (Dakhi & Nur Firtria, 2019) as well as “school achievement in general” (Beck et al., 2002, p. 1; Roche & Harrington, 2013). It may therefore come as no surprise that beginning language learners rely more on vocabulary knowledge than grammar for text comprehension (Grabe & Stoller, 1997) or that in both first and second language learning contexts, “By far, the greatest lexical obstacle to good reading is insufficient numbers of words in the learner’s lexicon” (Laufer, 1997, p. 31). Bilingual learners consistently state that a dearth in vocabulary is also the single greatest reason for writing difficulties (Kwon, 2009; Pilar & Llach, 2007). It is therefore difficult to overstate the importance of vocabulary to academic success, for both native-speaking and bilingual learner students.

Vocabulary as a Predictor of Reading and Writing Proficiencies

Psychologist Edward Thorndike (1921) was one of first individuals to describe the importance of vocabulary for reading comprehension (Wright & Cervetii, 2016). Thorndike’s consideration of vocabulary’s role in text comprehension—something that would not appear again until Anderson and Freebody in 1979—seems remarkably early. Nearly 100 years later, a deluge of correlational studies regularly find vocabulary linked to success in reading comprehension (e.g., Karakoç & Köse, 2017; Laufer & Goldstein, 2004; Manihuruk, 2020; Milton, 2013; Nagy, 2007; Qian, 1999; Zhang & Zhang, 2022). Vocabulary knowledge has been

directly correlated with reading success at all educational levels (McKeown et al., 2017; Urdaniz & Skoufaki, 2019). Because vocabulary is a fundamental distinguishing feature of genre (Coxhead & Byrd, 2007; Olinghouse & Wilson, 2013) and of discourse type (Biber et al., 1999), it is equally important for writing success. Consequently, there are numerous correlational studies linking vocabulary to writing proficiency (e.g., Amini & Iravani, 2021; Crossley and McNamara, 2012; González, 2017; Karakoç & Köse, 2017; Kiliç, 2019; Qian & Lin, 2020; Won, 2019).

Some studies (e.g., Baba, 2009; Karakoç & Köse, 2017; Rashidi & Khosravi, 2010) have discovered that while the ability to understand vocabulary when encountered—or *receptive* knowledge—is important to reading and writing, the ability to use vocabulary correctly—or *productive* knowledge—may be a better predictor. Rashidi and Khosravi (2010), for instance, found that the ability to provide definitions of words predicted higher levels of reading comprehension than receptive knowledge, and by analyzing the lexical aspects associated with high proficiency scores in two summaries each written by 68 Japanese students, Baba (2009) demonstrated that higher productive vocabulary knowledge is needed for greater writing proficiency. Karakoç and Köse's (2017) multiple regression study of 178 English as a foreign language (EFL¹) students discovered that while both receptive and productive vocabulary knowledge predicted reading and writing proficiencies, the contribution of productive vocabulary knowledge was 48% greater than was receptive vocabulary knowledge. Thus, having a large productive vocabulary facilitates both reading and writing for bilingual learners. The implications of these studies' findings lie in the great advantages for students who can transfer

¹ English as a foreign language describes bilingual learners learning English in non-English dominant countries.

vocabulary understanding into active word usage, thereby improving both reading as well as writing proficiencies.

Vocabulary as a Predictor of Educational Success

The place of social and economic power behind successful vocabulary intervention outcomes looms large in the work of Hoff (2003) Fernald et al. (2012), whose work connects vocabulary knowledge to socioeconomic status, and in the work of Hoff and Naigles (2002), who found that early exposure to lexical richness, syntactic complexity, and language exposure grants later academic advantages among toddlers. It also appears in the work of Cunningham et al. (2011). Cunningham and his colleagues found support for the environmental opportunity hypothesis, which links vocabulary knowledge to parental educational level and early exposure to written text.

As demonstrated by Hyland (2011), knowledge of academic vocabulary bestows power to readers and writers. Hyland examined the language use in academic writing as an expression of authority. Reflecting on the revered place of texts labeled “scientific” and examining how language presentation achieves such authority, he explains that academic text exerts its rhetorical, persuasive nature through vocabulary presentation: the use or absence of hedges, boosters, attitude markers, self-mentioning, and engagement with the reader through directives, questions, references to knowledge, and asides. Hyland’s argument that the use of academic language to project authority is convincing. Research like Hyland’s proposes an understanding that with exposure to academic vocabulary, students may not only acquire the vocabulary of privilege, but they may also grow to understand the role of language in social stratification, and with it, the means to critically analyze their place within it. Consequently, vocabulary acquisition

may be a means toward both socioeconomic success and a critical reflection of one's place in society.

Vocabulary as a Predictor of Further Vocabulary Acquisition

Caro and Rosado Mendinueta's work (2017) accentuates the importance of common vocabulary items as the tools for engaging in basic communication to support interaction that will, in turn, promote higher levels of language acquisition. This observation is known as the Matthew effect. Adopting its name from the Gospel of Matthew's observation that the rich get richer and the poor get poorer, Robert Merton coined the term in 1968 to denote the "commonly observed tendency ... for initial advantages to accumulate through time" including, in educational contexts, vocabulary acquisition (Rigney, 2010, p. 4). Wright and Cervetti's (2016) review of vocabulary acquisition studies found evidence of the Matthew effect throughout:

Children who arrive at school with low levels of vocabulary knowledge are likely to continue to have relatively small vocabularies and are likely to struggle with text comprehension throughout their school lives (pp. 203-204).

Nuttall (1982) describes the Matthew effect in terms of a cycle during which a student's lack of understanding leads to slow reading, consequently leading to lack of enjoyment of reading, and with it, a reluctance to read. A more recent explanation credits the ability to make inferences that vocabulary knowledge provides (Albrechtsen et al., 2008). An early recognition of the Matthew effect in second language learning context may be seen in Coady's (1997) acknowledgement of a "beginner's paradox," which he describes as the need for a second language learner to have "fundamental competence" of basic vocabulary for the extensive reading needed to facilitate further vocabulary acquisition (p. 235). Cunningham and Stanovich's

(1997) longitudinal study found evidence for the Matthew effect in following the vocabulary development of a single group of 27 students over a period of 10 years, from grade 1 until grade 11. A consequence of the Matthew effect is the related finding that connects the vocabulary knowledge of early learners to “subsequent educational attainment” (Milton, 2022, p. 267). Effective vocabulary learning, then, may lead to an upward spiral of growth in reading and writing success, and with it, higher levels of education.

As Harklau and Pinnow (2009) note, bilingual learners are much less likely to be the focus of English literacy studies than native English speaking (NES) students, despite literature highlighting the great opportunity gaps bilingual learner students often face (Fry, 2008; Garcia et al., 2010; Gibson, 2016; Grantmakers for Education, 2013). Recently arrived bilingual learners in U.S. secondary schools face difficult demands in having to learn academic material through a language they must acquire concomitantly. Many of these students have had limited formal schooling in their own countries or have experienced periods of interruption to their schooling (DeCapua & Marshall, 2011). Some are refugees whose trauma has been shown to negatively impact learning (Kaplan et al., 2015). When calculated using the baseline budget poverty rate measure, 30% of immigrants from all but a handful of countries of origin live in poverty² (Hernandez et al., 2010). As Hernandez et al. explain, this reflects “greater negative consequences than either limited mother’s education or living in a one-parent family” (p. 14). Bilingual learner students born in the United States who retain bilingual learner status into high school are often aural learners who lack academic language skills in either language. Consequently, they often possess weaker literacy skills than their native counterparts (Moore & Klingner, 2014).

² The baseline budget poverty rate is a measure of poverty that accounts for cost of living adjustments, including food, shelter, educational resources, healthcare, and other necessities.

At the intersection of bilingual learner status and other social identities, there are complexities. These include race (Kubota, 2015; Liggett, 2014; Monzó & Rueda, 2009; Roxas & Roy, 2012; Roy, 2018), special education status (Blanchett et al., 2009; Hawley, et al., 2013; Kangas, 2017; Kayumova et al., 2015; Poon-McBrayer, 2016; Trainor et al., 2016; Waitoller & Kozleski, 2013, Núñez, 2014; Yu, 2016), refugee status (Pittaway & Bartolomei, 2001; Roy, 2015; Roy, 2018), sexual orientation (Paiz, 2019), socioeconomic status (Jiménez-Castellanos & Garcia, 2017), religion, (Aguilar et al., 2016), and country of origin (Lee, 2006). As Seirk (2019) contends, schools continue to be sites of institutional oppression for bilingual-student learners, propelled and reproduced by administrators, instructors, and other students.

It is therefore an understandable if sobering fact that within the backdrop of such educational difficulties, which must seem insurmountable at times, bilingual learner students are twice as likely as native speakers to drop out of school (Morse, 2005), although latent class variables (e.g., socioeconomic status) influence *which* bilingual learners are more likely than to drop out than others (Zaff et al., 2020). As Urdaniz and Skoufaki (2019) aptly state, “Academic vocabulary knowledge predicts students’ academic achievement across educational levels” (p. 1). For these reasons alone, researching effective vocabulary interventions to enhance academic vocabulary skills through reading and writing seems not only a worthwhile endeavor, but a critical one for social justice-oriented language teachers and researchers. Far from advancing deficit views of bilingual learners, my intention is to emphasize the importance of finding ways to improve their educational experiences and to examine the role of academic vocabulary in doing so.

Definitions

The following section defines important terms and operationalizes study constructs.

Bilingual learner. I define a bilingual learner student as one designated as an *English language learner* (ELL) by the State of Georgia. Originally referred to as limited English proficient (LEP), *English language learner* was legislatively defined by the English Language Acquisition, Language Enhancement, and Academic Achievement Act, a grant program within Title III of the Elementary and Secondary Education Act (2001). To receive Title III funding, states must report the number of LEP students and report progress in English learning for this student population (Every Student Succeeds Act, 2015). The State of Georgia uses WIDA Consortium's ACCESS test results to determine a student's status as a bilingual learner (Georgia Department of Education, 2021). I reject the use of the term *English language learner* because it has arisen from a deficit view of how these learners, who are proficient speakers of a language other than English, are perceived. Deficit perceptions inherent in terms such as *English language learner* continue to detrimentally impact their educational experiences through both policies and practices (Columbo et al., 2018).

First Language (L1). Cook (2016) explains that in attempting to move away from the paradigm of a learner of English as someone with a language deficit to be rectified, her message to students was simply, "Do not see yourselves as failures always trying to be like native speakers; see yourselves as successes, achieving things as L2³ users that are out of the reach of monolinguals" (pp. 187-188). Further, Hansen (2017) argues that speakers of colonized varieties of English—those of Kachru's (1982) outer circle—are native English speakers in their own right. In this spirit of this research, I define a first language with a focus on "language expertise, language inheritance, and language affiliation" (p. 543), a definition that is reflected in the participant questionnaire and therefore includes English.

³ second language

Gender. Although the terms sex and gender are often used interchangeably, as a culturally defined construct, I define gender as fluid, self-identified performance (Kimmel & Gordon, 2019) and therefore not necessarily cisgender for all people. Like first language, this definition is reflected in the participant questionnaire.

Vocabulary Acquisition. As a function of memory, vocabulary acquisition may not follow prescribed steps, and like memory, vocabulary knowledge may not be consistent from time to time even within the same person. Nevertheless, for second language learners, distinctive knowledge of a vocabulary item as strength of memory often occurs in graduated steps (Blumenthal-Dramé, 2012). Learners may have different degrees of familiarity of a vocabulary item, may recognize its form, may be able to provide a working definition of a vocabulary item, may be able to identify the word's meaning when presented with choices, may know the word when contextualized, or may be able to use the word effectively in speaking or writing. I define vocabulary acquisition as the obtainment of any measurable gain along a continuum ranging from unfamiliarity with a term or a lexical sequence to the demonstrated ability to use it contextually with correct intended meaning. Consequently, acquisition also includes the obtainment of memory traces of a term or lexical sequence. The vocabulary knowledge scale (VKS; Brown, 2008; Wesche & Paribakht, 1996) guided my understanding of vocabulary acquisition for the sake of determining whether any degree of vocabulary acquisition has occurred for my participants.

Lexical Sequences as Vocabulary. A traditional understanding of vocabulary seems to be that of single words acquired in a linear fashion, beginning with recognition and moving incrementally towards proficient production in speaking and writing. However, being heavily dependent upon contexts for meaning (Brusnighan, 2012; McKeown et al. 2017; Schmitt, 2010)

and therefore bound to the most fundamental of these contexts—the co-occurring surrounding words—for the sake of this study, I define vocabulary in terms of lexical sequences: commonly identified by corpus linguistic research and sometimes noted to be characteristic of the register and genre they are found within (Biber, 1999; Coxhead, 2000; Hyland, 2008). I agree with Swales' (2004) assertion that corpus linguistics studies therefore “erode traditional distinctions between grammar and vocabulary” (p. xi). I adopt Heid's (2008) broad definition of lexical sequences, that “any meaningful unit made up of more than one string of characters qualifies” (p. 340).

The topic of lexical sequences has garnered interest from multiple people and places of language research, resulting in variety in terminology and descriptions (Griers, 2008). Construction grammarians and logicians often refer to them as “*n*-grams,” where *n* represents the number of words comprising the sequence (Cappelle et al., 2019, p. 220; Cappell & Grabar, 2016, p. 271). Nesselhauf (2003) includes collocations among them. They have been called “semi-automatic word chunks” (Salazar, 2014, p. 2), “multi-word expressions” (Heid, 2008, p. 339), “pre-constructed phrases” (Davis & Morley, 2015), “lexical units” (Cruse, 2015, p. 81), “lexical bundles” (Biber & Barbieri, 2007), “phrasicons” (Swales, 2004, p. xi), “recurrent word combinations” (Ädel & Erman, 2012), “academic lexical phrases” (Cai, 2016), “phrasemes” (Griers, 2008), and “phraseological items” (Sinclair, 2008, p. 408). Sinclair asserts that lexical sequences are “the place where structures are engineered to allow meanings to take shape” (p. 408). The importance of lexical sequences is also acknowledged by the idiom principle having arisen from the field of lexical semantics. Underlying the idiom principle is an understanding of language as “strings of co-selected words that constitute single choices” (Granger & Paquot, 2008, p. 29)—in other words, lexical phrases that are stored and retrieved as single units. Among

others, Wray (2002) and Ellis and Wulff (2014) suggest that the presence of lexical sequences makes language “formulaic” (p. 1; p. 79). List of commonly found lexical sequences generally range from three-word to seven-word strings (Salazar, 2014).

Lexical sequences are more common in writing than in speaking (Biber & Barbieri, 2007), and research has demonstrated that in both first and second language contexts, common lexical sequences are processed more rapidly than single words of the same length (Conklin & Schmitt, 2008). Consequently, common advice to practitioners is to incorporate the most frequently occurring lexical sequences into teaching practice, especially in the teaching of writing (e.g., Cumming, 2017; Ebrahimi et al., 2021; Nation & Chung, 2011; Schmitt, 2010). Henriksen (2013) emphasizes the importance of teaching lexical sequences because when misused, they “may at least signal a lack of academic expertise” (p. 37), whereas Davis and Morley (2015) propose that learners’ use of lexical sequences in writing may assist with semantic prosody, or the connotations associated with phrases. Salazar (2014) notes the importance of writers’ understanding of lexical sequences most closely associated with academic vocabulary to convey a writer’s professional expertise in the writing of scientific articles. Finally, Flowerdew (2011) discusses lexical sequences’ affordances in learning grammar, vocabulary, and writing.

Lexicalization. The term *lexicalization* was coined in 1949 by Edward Sapir to describe a word having been defined in a language. In second language learning contexts, *lexicalization* refers to the presence of a targeted vocabulary item in a learner’s first language.

Nonlexicalizations are therefore those words for which a concept has no equivalency in a given language (Hopp et al., 2018). In other words, a nonlexicalized word is one that “could not be

translated (or replaced) with one single word in the learners' L1” (Heidari-Shahrea et al., 2014, p. 92). Contrastingly, lexicalized words are those having direct translations.

Lived Experience. I define *lived experience* in a dialogic sense. Guided and interpreted by language and thought, the interactions people have in the world may not be readily grasped as experiences unfold in real time (Van Manen, 2016), but whether acknowledged or not, such interactions necessarily influence future language and thought. I include in my definition not only these interactions, but also “the structural, social and emotional facilitators or barriers that accompany what (people) do” (Reichenbach & Wesolkowska, 2008, p. 164). In the context of my study, lived experiences include educational experiences, relationships, responsibilities, and the past experiences of my participants as well as how these experiences are framed by participants’ first languages and by the learning of English.

The Impact of L1 on Second Language Vocabulary Acquisition

For decades, researchers argued that an individual has a mental space from where all vocabulary knowledge is drawn (e.g., Cook, 1997; Dijkstra, 2005; Singleton & Little, 1999). For instance, Dijkstra’s (2005) review of psycholinguistic studies led him to suggest that a speaker accesses isolated words from a single lexicon, regardless of whether L1 or L2. However, Singleton’s (2006) review, coupled with the last 15 years of lexical research (e.g., Ellis, 2006; Paquot, 2013), seems to question the inferences that may be drawn from earlier findings. Singleton states that “in some sense, there is separation between the lexicons associated with the different languages known to an individual” (p. 13). Consequently, Singleton maintains that cross-lexical transfer occurs during language learning, and he proposes that the issue of degree depends in part on how different an individual’s native language is from the target language.

Vocabulary scholars more recently have concluded that a learner's native language decidedly impacts the acquisition of second language vocabulary (Ellis, 2006; Liu & Chen, 2019; Masrai & Milton, 2015; Paquot, 2013). As Hopp et al. (2018) explain, "more closely related languages at the lexical form level lend themselves more to transfer than less related languages" (p. 309). Hopp et al. (2018) identify three types of lexical knowledge that bilingual and multilingual individuals possess: translations of words with similar meanings, cognate words, which they define as "lexical entries that share word form and meaning representations across more than one language" (p. 307), and unique words—including those for which form, but not meaning, have been acquired. Within an individual speaker, Hopp et al. propose, the lexicon of languages acquired differ in number, in interconnectivity, and importantly, in "how lexical knowledge from one language can influence lexical performance in the other language(s)" (p. 307). The cross-linguistic influence of a learners' first language on bilingual learner vocabulary acquisition has recently been examined primarily through explorations of loanwords, cognates, and lexicalization.

Lexical inferencing, or the process of using background information and context clues to determine the meaning of an unknown word or phrase, is more difficult during second language learning when a word or phrase does not exist in a speaker's first language (Paribakht & Wesche, 2016). Paribakht's (2005) lexicalization hypothesis portrays lexical inferencing as a process during which a learner exposed to a new word attaches it to already acquired semantic and syntactical information. Therefore, when a concept does not already exist in the learner's lexicon, second language learners are less able to infer the meaning of the word in context (Chen & Truscott, 2010; Golaghaei & Yamini, 2014). In a study of 20 Persian undergraduate EFL students, Paribakht (2005) found that Persian speakers were less likely to know English terms

nonlexicalized in Persian, and Chen and Truscott (2010) discovered that this truism holds even when learners have been exposed to the term contextually up to seven times; learners possessed less knowledge of nonlexicalized vocabulary at every stage of knowledge when tested using the vocabulary knowledge scale assessment. Heidari-Shahreza et al. (2014) report that among L1 Persian undergraduate bilingual learners, nonlexicalized words required more exposure for vocabulary acquisition than did lexicalized words, leading them to suggest, like Chen and Truscott (2010), that such words be taught to language students directly. These findings suggest that nonlexicalized vocabulary items are more difficult to learn incidentally. Lexicalization may also change the nature of second language vocabulary use. For instance, in an early examination of the role of lexicalization in bilingual learner vocabulary learning, Blum and Levenston (1979) discovered that Hebrew learners of English avoid using English terms that are not lexicalized in Hebrew.

According to Corson (1997), “Control of Graeco-Latin academic vocabulary of English is essential to academic success” (p. 671). Corson also asserted that cultures and life histories greatly impact the propensity for students to engage in academic vocabularies, whether in or out of school. Corson’s finding demonstrates that speakers of some first languages, such as Proto Indo-European languages spoken in Western Europe in the Italic branch (i.e., Spanish, French, Italian, Portuguese, Catalan, and Romanian) have advantages for learning English vocabulary over speakers of other languages. As noted by Daulton (2010), the exploration of English loanwords and cognates are promising ways to access English vocabulary, an advantage that speakers of some first languages naturally have over other first languages.

Van Benthuisen (2010) identified semantic change, phonological change, and orthographic opacity as three confounding factors involved in the ease of English loanword

acquisition among L1 Japanese bilingual learners, and indeed, some studies have suggested that easy access through shared cognates is not automatic. For instance, Bayat and Oveidi (2015) demonstrated that Iranian speakers may not gain an advantage from English cognates unless learners are taught cognate-seeking strategies, and in investigating whether Serbian English learners were more likely to know English vocabulary sharing cognates with Serbian, Danilović (2010) found that cognates did not provide advantages. Like Bayat and Oveidi (2015), Danilović (2010) concludes that learners need explicit instruction in seeking cognates. Agustín Llach (2016) compared receptive English vocabulary knowledge of 82 Spanish and German 4th grade students to examine whether Spanish students gained an advantage in having a larger number of cognates, but the researcher found no significant differences in receptive knowledge between speakers of the two groups. Thus, even when cognates between an L1 and English exist, they may not necessarily provide an advantage to bilingual learners for whom the connection has gone unnoticed.

Among the ways that cross linguistic influence can affect vocabulary acquisition also includes lexical intrusion. So called “lexical intrusion” is often seen during L3 (third language) production, causing such things as unintentional use of L2 during L3 use, even when L2 learners have high proficiency (Hopp et al., 2018, p. 308). Research on lexical intrusion has led many researchers to believe that dynamic interplay occurs between a learner’s L1 and L2 mental lexicon (Singleton, 2006). For instance, Hopp et al. (2018) avers that during L3 acquisition, language distance between an L2 and L3 may impact the degree of lexical intrusion that may occur while learning a third language. These studies make clear that second language vocabulary acquisition may sometimes be a function of a learner’s first language.

The Impact of Gender on Vocabulary Acquisition

Gender difference has long been an area of interest to sundry fields of research (Na, 2016), including educational research. For instance, gender has been shown to impact educational outcomes as a factor of self-regulation of behavior and subject matter (Weis et al., 2013), and the gendered use of English is regularly examined (Newman, et al., 2008), often finding that men and women use language differently (Newman, et al., 2008; Singer et al., 1977). Differences rear themselves through such things as the unbalanced use of apologies by women (Schumann & Ross, 2010) and interruptions by men (Plug, et al., 2021). In a foundational study, Lakoff (1972) found that women are more likely to use hedges—language that expresses indirectness or lack of certainty (Hyland, 1998; Lakoff, 1972). Such differences may be considered a reflection of social and cultural gender stereotypes (Lakoff, 2003), including such things as the threshold for what individuals should find offensive (Schumann & Ross, 2010). For instance, comparing the gendered use of language in the 1970s with the late 19th century, Singer et al. (1977) found that “vocabulary changes appear to reflect changes in the social structure of the society” (p. 267). The gendered use of language seems to appear true in bilingual learner contexts as well. For example, Ishikawaa’s (2014) analysis of Spanish 6th grade writing found that girls used more intensive adverbs when writing in English. Gender has also been shown to correlate to bilingual learner perceptions of language learning (Fatemi & Asghari, 2012). Finally, gender disparities are often found in writing development studies (Ahmed et al., 2014) as well as productive vocabulary knowledge (Montero-SaizaAja, 2021). Most studies examining the relationship between bilingual learner vocabulary acquisition and gender fall under the umbrellas of vocabulary learning strategy or differences in vocabulary knowledge.

Although gender seems to inform the preferred choices of vocabulary learning strategies (Kobayahsi & Little, 2020; Shadikah et al. 2017; Ahour & Abdi, 2015; Na, 2016) when a specific strategy is investigated, gender often has little to no effect on how well vocabulary items are acquired by participants (Elhami & Mahmood 2015, Kobayashi & Little; 2020; Purwanti Heri et al., 2020; Tabrizi & Abbasi, 2016). Exceptions can be seen in Manuel (2017), who found no differences in strategy use among 60 EFL students in Angola, and in Roohani and Akbarpou (2016), who found that while strategies incorporating music were effective for both genders, the impact was greater for females.

Research measuring L2 English vocabulary knowledge by gender shows mixed results. Fernández Fontecha (2015) found no gender differences in English vocabulary knowledge among 66 middle school students in Spain, and Elhami and Mahmood (2015) found no difference in English vocabulary knowledge among young Iranian children. On the other hand, Fardad et al. (2015) discovered that males outperformed females in receptive knowledge of English vocabulary among 65 female and 24 male undergraduate Iranian EFL students, and Maghsoudi and Shamaei (2016) likewise found that young male Iranians had greater knowledge of English than their female counterparts. Thus, as a corollary, the global nature of language research seems to suggest that language learning data may also carry data about cultural attitudes and gender educational differences, even if left unexamined by the researchers.

Research finding L2 English vocabulary learning advantages for female children have suggested both biological and cultural explanations for gender differences when found. Agustín Llach and Terrazas (2012) conducted a 6-year study tracking the English receptive vocabulary repertoire of EFL Spanish speaking adolescents. The researchers report that into early adolescence, English vocabulary development was greater for females than for males, but after

the 7th grade, males began to catch up. Agustín Llach and Terrazas propose that “the special characteristics of this period of life, teenage years, may also account for this difference in favour of males” (p. 67) and note that the early maturation of girls may lie behind this finding.

However, as Kobayashi and Little (2020) explain, social and cultural expectations necessarily play a role in gendered intervention outcomes. For instance, Gu (2002) explains that Chinese females are expected to succeed in language learning, and women may be more apt to experience the anxiety that Krashen (1982) hypothesized may affect language acquisition.

As Lee (2020) explains, the contradictory reports of vocabulary gender studies might be explained by gender’s interaction effects with other factors. Lee’s study of 492 Korean EFL undergraduate students investigated the interrelatedness of aptitude, motivation, and gender to English vocabulary knowledge and found gender unrelated. Wang’s (2016) true experimental design study examined the role of gender in learning emotionally neutral English vocabulary among 126 Chinese middle school students who were presented with English vocabulary, exposed to either neutral or humorous videos of three or nine minutes in length, and then tested on the vocabulary memory consolidation. Wang found that girls learned more vocabulary overall but found that when the humorous video was nine minutes in length, vocabulary acquisition dropped even further among males. Studies conducted by Kaushanskaya et al. (2011) and Kaushanskaya et al. (2013) demonstrate that among both adults and young children, females are better able to acquire phonologically familiar words than males, but this difference disappears for terms that are non-phonologically familiar. The findings of these two studies imply that interaction effects may occur for gender and first language during language acquisition, whereas the findings of Wang (2016) and Lee (2020) imply that social, cultural, and cisgender biological factors may also play a role.

Purpose of the Study

The purpose of my study was to examine factors that affect English language learning, particularly vocabulary acquisition through reading and writing, among high school bilingual learner students. As can be seen from the studies laid out in this chapter, the problem is a multi-dimensional one. Consequently, I investigated it as a QUAN → QUAL mixed methods study following Riazi's (2017) fully integrated design model for mixed methods language research. Using this model, both qualitative and quantitative data were collected and analyzed independently, although they informed one another, at times leading to additional analyses and investigation. During a third phase, I made meta-inferences from the findings that arose from both the qualitative and quantitative analyses.

As can be seen in the literature review presented in Chapter 2, in both NES and second language learning contexts, most vocabulary acquisition occurs incidentally through reading, which offers affordances in presenting form as well as meaning (Webb et al., 2020), and when perceived in terms of lexical sequences, incidental vocabulary learning offers grammatical advantages as well. Readers can reread to clarify comprehension can consult dictionaries as needed, yet the intentional transfer of newly encountered words after reading may offer cognitive advantages that assist in acquisition.

Through an intervention comparing vocabulary exposure through reading to exposure through reading coupled with writing, I used quantitative methods of data collection and analysis to examine the role of L1 and gender on intervention outcomes. Through in-depth interviews after the nine-week intervention phase of the study, I examined additional factors that have influenced the language learning of the 12 participants. Finally, I considered how the findings of both study strands inform a rounded understanding, not only of the factors impacting my

participants' English learning experiences and vocabulary acquisition through the interventions, but also how factors interacted to do so.

Research Questions

My research questions are the following:

1. (QUAN) Over a 4-week period, can exposure to targeted lexical sequences through reading and writing of academic texts lead to acquisition among secondary school bilingual learners?
 - a. Can exposure through reading lead to targeted lexical sequence acquisition?
 - b. Can exposure through reading followed by intentional use of the sequences lead to acquisition?
 - c. Is one intervention more effective than the other?
 - d. Are any initial gains retained in delayed vocabulary testing?
 - e. Are there any interaction effects due to participants' first language (L1) or gender?
2. (QUAL) What other factors affect intervention results?
3. (QUAL) What factors affect the educational experiences of the participants?
4. (MIXED METHODS) What meta-inferences can be drawn from the quantitative and qualitative strands of the study about factors affecting bilingual learners' learning?

Overview of the Study

The three-month study randomly assigned 49 high school bilingual learner students into two groups. A simplified version of the vocabulary knowledge scale (VKS) assessment (Brown, 2008; Wesche & Paribakht, 1996) was administered to participants as a pretest measuring the breadth and depth of the eight target lexical sequences. For four weeks, Group A read short

information texts twice weekly. Embedded within the readings were four targeted lexical sequences characteristic of academic writing⁴, evenly dispersed over the four-week period. After the four-week period, the participants took a vocabulary posttest of the four targeted lexical sequences. For the next four weeks, Group A read passages with four different lexical sequences but followed each biweekly reading with short, written responses to writing prompts, incorporating the four targeted lexical sequences into their written responses. Group B performed identical tasks but completed the tasks in the opposite order for the sake of counterbalancing.

At the end of the eight-week period, posttesting of vocabulary measured again. One month later, the same test was administered as a delayed posttest to determine if any initial gains were lost and to uncover changes in vocabulary depth of knowledge of the targeted vocabulary items. At the end of the eight-week period, a qualitative study strand took place with 12 participants. The participants were chosen based on their representation of the larger sample in terms of demographic data and intervention outcomes; participants who identified Spanish or English as an L1 were chosen based on their test scores, with males and females having equal representation. That is, the participants with the highest scores, the lowest scores, and those closest to the median scores were asked to take part. These semi-structured, in-depth interviews elicited information about participants' daily lives to explore factors that impact their language learning experiences. After the delayed posttest, a factorial ANOVA with repeated measures was conducted with VKS scores as a dependent variable and with group number, L1, and gender as independent variables.

⁴ The academic vocabulary sequences chosen for this study, the pattern of distribution, and the instruments used for the study are discussed in depth in chapter 3.

2 REVIEW OF THE LITERATURE

The planning and organization of my literature review has been informed by Cooper and Hedges' (1994) taxonomy and may best be described as an *integrative research review*. Cooper and Hedges explain that such a review provides criticality, suggests areas of new research, and includes theory⁵ for the sake of generalizing and identifying needed research. The overall organization of Chapter 2 is based on the foci of literature reviewed and includes three major sections followed by a conclusion. In the first section, I describe the scope of the literature review and provide a rationale for inclusion criteria. In the next section, I describe the theoretical framework that has arisen from an analysis of the literature, and in Section 3, I analyze empirical studies of vocabulary acquisition through reading, writing, and the two in tandem. Finally, I briefly conclude with an overview of findings and provide a study rationale based on these findings.

Section 1: Scope of Literature Review

My review explores literature about English vocabulary in both native English speaker (NES) as well as bilingual learner contexts. The two populations have similar cognitive processes (Carroll, 2001; van den Bergh et al., 2016), similar reading processes (Grabe, 2011), similar semantic mapping processes (Jiang, 2002; Zareva, 2007) and similar use of copy and paraphrasing of source text (Keck, 2014). In fact, Feak and Swales (2009) as well as Canagarajah (2013) have suggested that the distinction between L1 and L2 writing in the realm of research is eroding. Consequently, a thorough review of literature examining empirical studies, including

⁵ I draw distinctions among ways the term *theory* is commonly used in educational research: theory as discipline-specific hypotheses supporting my argument (learning and language acquisition theories), theory as epistemology guiding my research design choice and methodology (pragmatism), and theory as the framework through which I examined my results. Learning and language acquisition theories are examined as part of this literature review.

interventions that may impact bilingual learner vocabulary acquisition, requires not only an investigation of second language acquisition (SLA) vocabulary scholarship but a consideration of studies undertaken with NES participants as well.

It must be acknowledged that differences between the language development of bilingual learners and NESs certainly exist. Proponent of universal grammar Meisel (2011), for instance, argues that the language acquisition device⁶ operates differently for L1 than for L2 language acquisition and that for L2 learners, language acquisition is dependent upon the age of onset of acquisition (AOA). Differences between L1 and L2 acquisition of grammar explain why ultimate attainment seems to become more difficult as a learner's L1 differs more greatly from the learner's target L2. However, Meisel also explains,

It is not language which is affected by such changes (occurring during the critical period) but certain domains of grammar. Lexical knowledge, for example, is predicted not to be concerned at all. (p. 204).

For the sake of vocabulary learning, Meisel believes that differences between L1 and L2 language learning may, therefore, not apply to vocabulary learning.

My review integrates studies from both NES and bilingual learner secondary school students, or students attending grades 6-12, yet I included post-secondary and adult studies as well, not only because of the preponderance of these studies and the importance of their findings, but also because secondary school age students are more akin to young adults than to young children. Although notable exceptions make deterministic understandings of it problematic (Birdsong, 2006), puberty seems to mark a turning point for decline in L2 learners' ability to reach nativelike language acquisition, or ultimate attainment (Abrahamsson & Hyltenstam, 2009;

⁶ Universal grammar proponents define the language acquisition device as an innately human mental apparatus that allows language acquisition to take place.

Birdsong, 1992; Zhai, 2020), something that studies in neuroscience seem to support (Ortega, 2014; Zhai, 2020). Ortega (2014) states,

For decades now, the new field of cognitive neuroscience has contributed interesting evidence on the issue ... (and) make it possible to measure in milliseconds the activation patterns of neural networks involved in different cognitive operations while the brain is processing language stimuli. The converging findings favour a critical period interpretation for L2 morphosyntax. (pp. 20-21)

Consequently, my review is confined to studies of secondary school students and young adults: participants whose ages approximate those my study seeks to recruit and participants who have cognitive similarity in the areas of interest to this study.

Literature selection process. From January 2016 until February 2020, I conducted regular searches of academic literature about English vocabulary in both NES and bilingual learner contexts. I garnered works discussing reading, writing, and vocabulary relationships, including theoretical literature supporting these relationships as well as intervention studies testing these theories. Once I became familiar with language learning theories that have guided vocabulary research, I examined articles, books, book chapters, and dissertations describing the process of vocabulary acquisition, types of vocabulary acquisition, measures of vocabulary knowledge and acquisition, and models of vocabulary acquisition. I explored discussions of vocabulary thresholds for reading comprehension, literature defining degrees and depth of vocabulary knowledge, and literature describing the nature of English vocabulary as it appears in the English language generally as well as how as it appears in specific texts and genres.

To locate this literature, I periodically collected abstracts of peer-reviewed academic journals through discipline-based searches in education, language and linguistics, and literature and writing databases from EBSCO⁷. I also searched Google Books and Google Scholar, and I examined book chapters from second language learning and teaching as well as vocabulary, reading, and writing handbooks, including those describing composition theory and those providing historical overviews of language learning theories and literacy teaching practices. I explored author indices from handbooks to familiarize myself with important scholarship, located these works, and perused their findings.

To understand the findings about vocabulary acquisition in English learning contexts, I sought works commonly cited within literature I read, and I scanned articles, presentation proceedings, book chapters, and books that specifically address the process of language learning with an eye towards the process of vocabulary acquisition. Throughout this process, I set aside empirical studies of English vocabulary acquisition through reading, writing, or the two in tandem. These studies are described and critiqued in Section 3. I set aside empirical research studies during the literature review collection process to examine more closely. I also set aside studies tracking the development of English vocabulary, and finding them scant, I returned to the literature to include development studies of primary school participants as well. Next, I examined empirical studies of secondary, tertiary, or young adult participants investigating any of the following:

1. English vocabulary acquisition through reading
2. English vocabulary acquisition through writing

⁷ These searches have included hundreds of databases from each of these three fields.

Throughout my research journey, I reached a point of saturation, after which very little research emerged that provided new findings or insights about English vocabulary acquisition, whether NES or bilingual learning in nature. I discovered that most vocabulary research has taken place in second language learning scholarship, and I uncovered numerous prominent areas of English vocabulary research interest that has emerged over the past 30 years in vocabulary acquisition through reading and writing. I have categorized these areas of research under the umbrellas of *vocabulary development*, *vocabulary through reading*, or *vocabulary through writing*. I begin with the theoretical framework. Then, after describing development studies, I discuss the general characteristics of reading and writing studies as they are outlined in Appendix A. Next is a detailed analysis of these studies and their findings. A conclusion then follows.

Section 2: Theoretical Framework

Two major linguistic theories inform my study, and their spattering appears in different ways throughout the literature I reviewed⁸. These two theories—entrenchment and usage-based theory (Blumenthal-Dramé, 2012; Ellis, 2003)—are closely related. In general, both suggest that language exposure coupled with attention and regular usage are the fundamental ways in which vocabulary items, particularly lexical sequences, are acquired by a language learner in both L1 and L2 contexts. Usage-based theory proposes that it is through such a process that grammar and usage conventions are also acquired. Entrenchment theory adds that high frequency of exposure and regular usage help language to become more permanently engrained in a language learner's repertoire. Because the goal of language acquisition is to retain language learned, usage-based theory and entrenchment theory are fundamental to understanding how to facilitate this process.

⁸ Núñez's (2014) multilevel model of intersectionality informed the interpretation of findings of the qualitative strand of the study. This model is described in depth in the methodology chapter.

Entrenchment. Entrenchment is defined as the memory strength of linguistic concepts within an individual or within a language (Schmid, 2017; Stefanowitsch & Flach, 2017), particularly words and phrases (Blumenthal-Dramé, 2012; De Smet, 2017). For the sake of vocabulary learning, it may be thought of as the degree of permanence to which vocabulary has been acquired, with strength measured by automatization (De Smet, 2017; Schmid, 2017). Eye-tracking studies, for instance, have demonstrated that words with high frequency in a language are more rapidly processed than are rare words, and the same is true of common lexical sequences (Underwood, et al., 2004). Usage-based models of entrenchment hold that because every communicative act changes language organization within the mind, such representations are continually in flux (Schmid, 2017), leaving a memory trace each time a person is exposed to a new word or phrase (Blumenthal-Dramé, 2017). There are two underlying assumptions implicit in the concept of entrenchment. First, words and phrases become entrenched in the mind through memory consolidation. Second, frequency of exposure and use leads to processing of common lexical sequences as a single unit (Schmid, 2017). High frequency words are also associated with specific sentence patterns; when words are encountered outside of the common patterns, they are processed more slowly. This phenomenon is known as “frequency effects” (p. 12). The process of entrenchment seems to require multiple exposure to words or phrases and presentation of them in varied contexts (Schmid, 2017).

Entrenchment research (e.g., Blumenthal-Dramé, 2012; Schmid, 2017) considers in part how words and phrases become the most frequent within a language and the role of individual entrenchment in that process, the role of social discourse in the process, and how different types of lexical sequences become entrenched differently within the mind. Social discourse becomes apparent when salience of information is considered. Gobet (2017) explains that the attention

people pay to language they perceive as important helps facilitate entrenchment. In general, the lexical sequence umbrella contains two major sequence types. The first type, those that are *fixed*, may be thought of as fixed groups of words with a single meaning processed as a single unit. Irreversible binomials, such as “law and order,” idioms, phrasal verbs, and collocations provide some examples of fixed lexical sequences (Schmid, 2017). Alternatively, sequences can be thought of as *frames*, or groups of fixed words with slots that may be filled with alternative words of a fixed part of speech (Blumenthal-Dramé, 2012): “V the hell out of NP” provides an example of this (Hilpert & Deissel, 2017, p. 58). Some slots are open, whereas others may be closed, filled only with a finite, specific group of words. Wood (2020) refers to unfixed sequences as “congrams” (p. 36). Phrases held in the mind as single, fixed units are more automatically processed than are frames (Schmid, 2017). Numerous variables have been proposed as influential in the process of entrenchment (Herdina, 2017). These variables may influence either the entrenchment process itself, or they may impact the influence of frequency and dispersion on entrenchment (Schmid, 2017).

One important factor is the language itself, specifically the degree to which the word order in the language is fixed; because the nature of lexical sequence chunking is linear, languages that are less dependent upon linear phrases for meaning may not be processed in the same manner that entrenchment implies (Schmid, 2017). Behrens (2009), for instance, proposes that because English is a fixed word order language with reliable form patterns, “English is more susceptible to the entrenchment of lexically specific patterns because identical strings of words are more likely to occur” (p. 395). Perhaps the largest factor determining entrenchment, however, is the individual language learner. Entrenchment differs from person to person; individuals differ by so many variables—some of which unmeasurable—that it is difficult to

control for all, or even most, factors. Nevertheless, there are several generalizations researchers have made about how and why entrenchment occurs. For example, Schmid (2017) states for L1 learners, phrases are learned as holistic units and are later “decomposed” into first simple, and then more complex, frames (p. 14). Additionally, the degree of strength of a word or phrase may depend on how complex a phrase is, the length of the phrase (Blumenthal-Dramé, 2012), or its degree of abstraction (Hilpert & Diessel, 2017). It is also understood that social interaction affects entrenchment, yet within this category, numerous subfactors arise. The degree of solidarity shared with interlocuter, for instance, or the source of language as well as the perceived prestige of the language source, whether person or media, effects entrenchment. Schmid (2017) states that research has not yet determined whether the type of language processing—oral or written—better supports entrenchment. Nevertheless, Ädel and Erman (2012) have demonstrated that NESs use common lexical sequences more often and with more variety than bilingual learners.

Entrenchment research also investigates the role of frequency of exposure and dispersion on entrenchment within an individual, or in effect, how traces lead to entrenchment. It seems clear that while frequency of exposure certainly helps facilitate the process (Stefanowitsch & Flach, 2017), frequency alone cannot fully account for why an individual may remember some words or phrases better than others. For example, multiple exposures do not necessarily lead to entrenchment, yet in some cases, a person need only be exposed once to a word or phrase in order for entrenchment to occur (Blumenthal-Dramé, 2017). Nevertheless, acquisition of a first language means that entrenchment of numerous words or phrases must occur for every human being who uses language. Usage-based theorists have suggested that together, dispersion coupled within context are a significant source of entrenchment. Baayen (2010) has proposed that

contextual diversity assists in the process. That is, the reappearance of a word or phrases over a long period of time and in multiple contexts facilitates the ability to remember it.

Usage-based Theory. Although Davis and Morley (2015) assert that nearly 50% of all discourse is composed of some type of phrase or word sequence, usage-based theorist Nick Ellis (2003) as well as Cappelle and Grabar (2016) treat language acquisition as schema-building and provide compelling evidence that *all* language may be acquired in naturally occurring chunks of language, with *syntax* simply being the patterns having emerged from the brain's averaging of constructions: in effect, an individual's collective life experiences with language input. Semantics scholar Barsalou (1992) defines the development of an individual's semantic prototype remarkably similarly, as the storing of multiple exemplars averaged out, from which the mind creates a single prototype. Because each person's cultural milieu and exposure to exemplars vary, individuals have idiosyncratic prototypes in mind for a given concept. Usage-based theory hinges on the hypotheses that language acquisition is highly dependent upon exposure to the language and that usage rules are induced by the learner through such input. Within usage-based theory, language is coded by the individual as *constructions*, which may vary in string size from morpheme to idiom. Through associative learning, an individual connects phonological, orthographic, and semantic information about the construction into the brain (Ellis & Wulff, 2014).

Once a significant number of language examples are stored in the long-term memory, Ellis (2003) suggests, a schema emerges to make way for "slot-and-frame" patterning that facilitates creative language. Ellis explains that patterns emerge, creating ever-stronger schemata with each exposure. The Five Graces Group (2009) describes language acquisition as well-worn patterns of usage, and within usage-based theory, repeated exposure to strings of

words eventually creates a well-trodden path leading to entrenchment (Blumenthal-Drame, 2012; Ellis, 2003; Granger & Paquot, 2008). Blumenthal-Drame describes this process:

Higher token frequencies will correlate with a gradual increase in processing ease, more precisely enhanced fluidity in composition or parsing. At some point, this process will lead to a new, holistic representation (p. 104).

As Alharthi (2015) explains, students recognize that attrition results from lack of language use. Blumenthal-Drame (2012) explains that it is from such repetition of patterns that lexical sequences—and their semantic representations—may become entrenched, or deeply acquired, through reading and writing. Entrenchment may lie behind Lantolf's (1994) finding that recall tasks devised to assess reading comprehension “may not consistently assess reading comprehension but may instead, enhance comprehension” (p. 419) and behind Blom et al.'s (2012) finding that vocabulary acquisition leads to greater gains in acquiring grammatical structures. The need for multiple exposure and semantic context is supported by research like Grabe and Stoller's (1997) case study of an untutored learner of Portuguese, for whom multiple dictionary references for the same terms were required for vocabulary acquisition through reading. In L1 literacy research, Hilden and Pressley (2011) emphasize the role of noticing for expert reading, calling such a reader “massively active” (p. 428).

Usage-based theory may hold the key to opening a century-old philosophical solution regarding language learning as well. In introducing the semantic analysis theory, Landauer and Dumais (1997) provide an answer to the “poverty of the stimulus” that seems to have plagued thinkers from multiple places and times, beginning with Plato (Chomsky, 1986). The poverty of the stimulus, as an extension of Plato's problem, expresses the observation that individuals seem to know more language than can be explained through exposure alone. Landauer and

Dumais' answer relies on the semantics concept of frames and fields, and perhaps a prototypical version of a usage-based theory. They describe the process thus:

Over short time spans, contiguities among output words would reflect closeness in the sender's semantic space. A receiver could make first-order estimates of the distance between pairs by their relative frequency of occurrence in the same temporal contexts (e.g., paragraph). If the receiver then sets out to represent the results of its statistical knowledge as points in a space of the same or nearly the same dimensionality as that from which it was generated, it may be able to be better, especially perhaps in estimating the similarity of words that never or rarely occur together (p. 215).

While I adhere strongly to a usage-based understanding of language acquisition, I hypothesize that the weight of individual input events towards Ellis's (2003) and Barsalou's (1992) averages is highly variable, and not simply because of differences students have in exposure. Perhaps Barsalou's (1992) prototype is better thought of as a multilayered fog of *remnants* that the triadic brain has encoded and individualized through a cacophony of notes, chords, and even the banging of pots and pans that our travels, our social and cultural orientations, our idiosyncrasies, our mothers' love—in effect, our life histories.

Section 3: Empirical Research

Within the empirical research examined below, empirical studies—including both qualitative and quantitative studies—are divided into the categories of vocabulary development, vocabulary acquisition through reading, and vocabulary acquisition through writing. Each subsection begins with an overview of theories supporting the empirical studies.

Vocabulary Development

Development studies may trace the development trajectory of bilingual learners or NES students through examinations of vocabulary use over multiple years or multiple drafts. Such studies may compare bilingual learner vocabulary development to that of NES speakers (e.g., Arnaud & Savignon, 1997). They may compare equivalent groups of different age groups or grade levels (e.g., Levitzky-Avid & Laufer, 2013; Roessingh et al., 2005, White, 2015) or they may track the use of vocabulary by single participants through multiple drafts of the same assignment, longitudinally (e.g., Lessard-Clouston, 2012; Muncie, 2002). They may also report the findings of case studies of vocabulary learners through reading or writing, whether of individuals (e.g., Alsaif & Masrai, 2019; Ma, 2013; Parry, 1997; Severino & Deifell, 2011), groups (e.g., Schmitt et al., 2004), or cohorts (e.g., Ahmed et al., 2014; Duff et al., 2015; Schoonen et al., 2011).

Theories of vocabulary development and their accompanying studies provide important considerations to English vocabulary teaching and learning for at least two reasons. First, longitudinal studies often provide information about how learner vocabularies grow vis-à-vis reading and writing development, including the underpinning elements that may hamper or help such growth. Second, an understanding of what word mastery means will support educator efforts to facilitate such mastery among their students.

Theories of Vocabulary Development

Theories of vocabulary development are beholden to understandings of word knowledge, which has been perceived in terms ranging from raw numbers to word associations and have been measured in terms ranging from retrieval speech to free production (Bardel et al., 2013). Examinations of word knowledge have been undertaken in the disciplines of literacy education (e.g., Cronbach, 1942; McKeown et al., 2017; Richards, 1976) as well as applied linguistics (e.g., Laufer & Goldstein, 2004; Webb, 1997). Research examining the development of vocabulary

knowledge is less common in literacy research than in the field of applied linguistics, but when discussed by scholars of literacy education, vocabulary knowledge is often characterized as categorical.

Categorical word knowledge. Descriptions of categories of word knowledge may begin with Cronbach (1942), who identifies five types of word knowledge. Cronbach describes these as the ability to provide a definition, to recognize correct form, to recall contextual meaning—which he refers to as “breadth of meaning.” Cronbach also includes “precision of meaning,” or knowledge of unfamiliar applications of the term, and finally, the ability to apply the term to “thinking and discourse” (p. 207). Another early examination of word knowledge is Richards (1976). Richards discusses eight assumptions about vocabulary knowledge among NESs, including derivations, homonyms, connotations, semantic associations, and pragmatic knowledge. McKeown et al.’s (2017) list of 11 types of word knowledge mentions knowledge of a word’s part of speech, its collocational use, its register, its semantic constraints, and the connotations associated with it. With the exception of Cronbach’s (1942) distinction between receptive and productive vocabulary knowledge, these descriptions of vocabulary knowledge seem to reflect a perception of vocabulary as single words acquired, as part of information packets, in a linear fashion.

Fitzgerald and Shanahan (2000) categorize four types of knowledge as a type of tree map that includes pragmatic knowledge (metacomprehension, purpose, and word “identification or production”) domain knowledge (prior knowledge, content knowledge, vocabulary, and how text develops meaning), and knowledge of “universal text attributes” (p. 41), which include phonology, syntax, punctuation, and genre. Fitzgerald and Shanahan introduce a model proposing critical knowledge at various stages of reading and writing based on the

commonalities between the two activities. Within their model, stages occurring within secondary school ages include, among other things, vocabulary meaning and knowledge of genre, which they describe as “syntax of chunks of ... text structure” (p. 45).

Degrees of word knowledge. A relatively large body of research has examined vocabulary knowledge through distinctions between depth and breadth of vocabulary knowledge and receptive and productive (or active and passive) vocabulary knowledge (e.g., Laufer & Goldstein, 2004; Webb, 1997). The receptive-productive/active-passive vocabulary constructs were originally described as two separate lexicons, “qualitatively different” (Meara, p. 150, 1990); Liu (2018) also describes vocabulary knowledge as having both “quantitative and qualitative aspects” (p. 705). Nevertheless, the distinction is now more often discussed as a hierarchical continuum, along which individuals move from understanding a word encountered to being able to quickly recall and produce the word accurately in either speaking or writing, especially in terms of Webb’s (1997) depth of knowledge theory (Nation, 2001; Read, 2000, Schmitt, 2014, Wesche & Paribakht, 1996). Laufer and Goldstein (2004) categorize four positions along the continuum to account for such differences as the ability to provide a translation and the ability to use an item in writing or speaking. Accordingly, the researchers distinguish between knowledge of form (active) and of meaning (passive). Thus, *passive recognition* refers to the ability to select the equivalent word in one’s first language given multiple choices, while *active recognition* refers to the ability to choose the correct word given its definition in one’s first language. *Passive recall* is the ability to provide a translation given the target word, while *active recall* refers to the ability to provide the target word when provided with the translation. However, Laufer and Goldstein’s typology seems to ignore a learner’s

fluency in response, and more importantly, it does not accommodate a learner's ability to produce a word contextually.

Lexical quality hypothesis. Within cognitive psychology, words are “frontier” (Trembly, 1966) when form is familiar but for which semantic knowledge remains incomplete (Durso & Shore, 1991; Frishkoff et al., 2008; Kennison, 2013). The incremental process of vocabulary acquisition is sometimes represented through the lexical quality hypothesis, which holds that because of the close memory association between word form and meaning, partial knowledge of a word enhances the ease of its acquisition (Adlof & Patten, 2016) and is supported by neuroscience research, which has established that vocabulary meaning acquisition is a complex, multifaceted process (Frishkoff et al., 2008). Durso and Shore (1991) identify word knowledge as a range between an inability to determine whether a word exists to what the researchers call “expressive” vocabulary (p. 191), which involves vocabulary recall and the ability to produce it in speech or writing. Providing early evidence of the nuanced nature of vocabulary learning, among seven true experimental studies reported by Durso and Shore includes one asking participants to define or describe word definitions through free response when presented with the word. Students were simply asked to define words presented to them through interview. The researchers' findings reveal that despite “repeated denials of knowledge of a word,” participants sometimes exhibited “the kernel of some understanding of the word” (p. 197). For instance, participants were sometimes able to respond with common lexical sequences containing the word even while explaining they didn't know the word. These results imply that a language user's trace knowledge of a word may sometimes be unrecognized, even to the user. Alcón (2007) presents similar findings in a study demonstrating that vocabulary learning may be enhanced by contextual exposure and may lead a learner to use vocabulary in writing with long

term retention even when the learner may not be able to define the word's meaning. Finally, Landauer and Dumais (1997) demonstrate that students exposed to a text may acquire words absent from but related to words present in the text. These studies support an understanding of vocabulary acquisition as a holistic endeavor, with the addition of one concept or word changing all parts of an individual's conceptual frame (Nagy & Scott, 2000). Findings such as these strongly suggest that even terms like "degree" and "continuum" may fail to capture the intricate complexity of process underlying vocabulary acquisition.

Empirical studies. I located 14 studies in NES and bilingual learner contexts that trace the development of vocabulary through reading and writing. These studies examine the development of vocabulary usage over multiple drafts or multiple years. Multiyear studies may measure general change in vocabulary knowledge of study participants, but most often, they compare groups of varied ages and grade levels that researchers consider otherwise equivalent. For instance, Huang (2015) compared the frequency of lexical sequences to the accuracy of their use using a colossal corpus of 5990 EFL argumentative writing samples collected from 30 universities throughout China. Comparing university seniors to juniors, Huang found that seniors used 1½ times more lexical sequences with greater diversity, leading the seniors to higher writing proficiency.

Multiple draft studies. Two development studies, Muncie (2002) and Lessard-Clouston (2012), examined the lexical development of learner writing over multiple drafts. Believing that "writing is an excellent opportunity for improving and consolidating vocabulary" (p. 227), Muncie (2002) analyzed the use of vocabulary sophistication over a month among the three drafts of 25 intermediate to upper intermediate level bilingual learners to find that sophistication of vocabulary generally improved from the first to the last draft. However, Muncie's study

compared drafts of different genres; the first assignment required students to write about themselves, whereas the final composition required students to discuss friendship. Depending upon how participants interpreted these topics, the difference in topic may have invited different types of vocabulary and lexical sequences usage. Furthermore, the baseline assignment was a timed assignment, although the remaining compositions were not. Using VocabProfile to identify 10 discipline-specific targets that included both phrases as well as words, Lessard-Clouston's (2012) multiple draft study tracked the vocabulary use of 12 bilingual learners and NES graduate students over multiple drafts of a writing assignment during a three-month period. They found similar development among both groups of students in the acquisition of targeted items, which included both words and phrases. However, having undertaken no quantitative analysis, it is unknown whether statistical significance was found between the two groups.

Multiyear studies of different cohorts. Roessingh et al. (2005) compared the academic language acquisition outcomes at graduation for 47 upper-level secondary bilingual learners who by grade 12 were identified as having intermediate proficiency. The research intended to measure the effects of age of arrival as a measurement of L1 proficiency and the effects of English for speakers of other languages (ESOL) instruction. Two of the four groups were composed of six and eight members respectively who had "little or no" ESL instruction (p. 7), while the remaining two groups were upper-level high school students who had received ESL support. The researchers identified a plateau appearing during middle school for most students, regardless of instruction received or of age on arrival. However, the study may be problematic in that control for internal validity of such varied groups, who had generational differences and may have experienced curricular differences as well, would be difficult to control. Another potential problem lies in the use of age on arrival as a stand-in for L1 proficiency. A learner's age of

arrival does not always show a negative correlation to language acquisition (Birdsong, 2006), and according to Laufer and Paribakht (1998), when age of arrival does correlate to language proficiency, it only becomes relevant two years after arrival.

White's (2015) NES study examined the three-year vocabulary development as evidenced in student writing from age 15-18 years old. White found that the use of academic vocabulary did not appear with statistical significance until 11th grade, but in direct contrast to Roessingh et al.'s (2005) findings, White (2015) appears to have found an increase in vocabulary use through the middle grades, having statistical difference from both elementary and high school vocabulary growth. In a similar study, Levitzky-Avid and Laufer (2013) tracked the cross-sectional, changing nature of vocabulary and collocation use in Hebrew EFL writing over seven years, from grade 6 until freshman college year, to examine how the language use of bilingual learners evolve over time. Although the researchers do not discuss lexical sequences as a stated aspect of analysis, Levitzky-Avid and Laufer define *collocations* as "word combinations" (p. 129), use examples that include three-word phrases, and invoke research discussing the formulaic aspect of language use. Using VocabProfile, they sought to uncover any correlations between specific years of study and several lexical dimensions, including variation in vocabulary use. They find inequitable development along the trajectory, with a significant increase in the use of lexical sequences occurring only during university studies. The study also found that lexical variation appears with significance in grade 12 and that statistical significance in increased use of collocation and rare words began only at the university level. However, like Roessingh's (2005) and White's (2015) studies, comparing different cohorts may be problematic for controlling internal validity.

Multiyear studies of the same cohort. Schoonen et al.'s (2011) two-year study of the writing development among 389 secondary students in the Netherlands included annual measurements of the receptive vocabulary acquisition of 35 English words. To seek strong validity in results, Schoonen et al. used multiple writing samples for each measurement. The study compared writing outcomes in Dutch and English, finding remarkably similar trajectories, although bilingual learner writing had a stronger dependence upon vocabulary knowledge on writing quality than did NESs. Arnaud and Savignon (1997) compared advanced proficiency groups of bilingual learners with varied years of experience to a single group of undergraduate NESs to track the development of advanced bilingual learners towards nativelike proficiency over a five-year period. Arnaud and Savignon's participants comprised three groups of French speakers in English teaching preparatory programs (years 1, 3, and 5 respectively) led by French-speaking English instructors who had at least 10 years of experience teaching English. The authors found significance among all groups, with the greatest difference occurring for experienced English teachers, especially in the acquisition of rare English words. When compared to native English-speaking undergraduate students, experienced English teachers demonstrated greater knowledge of rare English words than NES but knew fewer lexical sequences. However, all 60 lexical sequences were idioms, a group of phrases known to be problematic for L2 (Türker, 2019), which may explain this difference.

Other longitudinal development studies. As Schmitt et al. (2004) explained, "large phrasal vocabulary has seldom been put to empirical test" (p. 63). Accordingly, they provide a unique example of such a study. The researchers identified 20 target lexical sequences, painstakingly chosen through the process of taking common academic terms from several academic word lists and, with the use of a large corpus database, finding the lexical sequences

most commonly couching these terms. The list was then condensed to those lexical sequences found within ESL teaching materials. Schmitt et al. asked 94 participants of upper-intermediate proficiency to take pre- and post-receptive and productive tests to measure vocabulary development over the duration of the course, in some cases two months and for others a duration of three months. The researchers examined vocabulary development of these phrases against gender, L1, motivation, age, and language aptitude. They found evidence that overall, learners acquired the targeted items with an average movement from 85% to 95% of receptive knowledge. Schmitt et al. explain that productive scores showed higher gains. Those participants having beginning knowledge of the most frequent 2000 words in English were more likely to develop further knowledge of lexical sequences, yet Schmitt et al. found no correlations between gains and learner factors, which included gender, age, language aptitude, and motivation. They explain, “although learner characteristics might very well affect formulaic language development, their impact may be modified by other factors” (p. 69). The researchers did not undertake a factorial ANOVA study to examine these differences, instead comparing test findings to individual factors one by one. Perhaps a factorial ANOVA would have better illuminated the findings and uncovered potential interaction effects.

Kelley et al.’s (2010) study of 476 sixth-grade students compared two sets of classrooms teaching vocabulary through reading. The control group followed a district-mandated vocabulary program, while the experimental classrooms made use of three guiding principles borne from their review of extant vocabulary research:

1. Teachers should focus on deep understanding of only a small number of words,
2. teachers should choose the most common words in the English language, and
3. direct teaching should be accompanied by the teaching of word-learning strategies.

The researchers simply state that the experimental group of students gained “significantly better results” of receptive knowledge, with no further data provided to the reader. Although Kelley et al. (2010) include 346 bilingual learners and report that 58% of the students in the school setting were low income, the researchers did not control for either SES or bilingual learner status in their study.

Only one vocabulary development study I examined arguably attempted to uncover how traces lead to entrenchment by connecting to words exposed in reading to that of writing. Choi’s (2017) structural equation modeling study of 178 participants sought to understand the pathway leading learners from receptive to productive vocabulary with a focus on how writing may mediate this path by seeking to understand how reading may lead readers to incorporate vocabulary into writing. Although Choi did not control for learner proficiency levels, the findings led Choi to suggest that receptive vocabulary knowledge occurring through initial reading may in turn spur a learner’s productive use in writing assignments. Therefore, Choi maintains that the productive use of words encountered through reading should be a critical part of vocabulary teaching and learning exercises.

Case studies of single participants. Two case studies tracked the vocabulary development of one or two young adult English learners. In a case study of a 21-year-old Chinese speaker’s English vocabulary development, Severino and Deifell (2011) envision vocabulary as both single terms as well as lexical sequences and richly describe a bilingual learner student’s vocabulary growth as it occurred through tutorials, comparing the physical presence in a writing center to online tutoring. The researchers sought to understand strategies that lead to the student’s uptake, taking care to note that uptake does not necessarily correlate to learning, and the reverse is equally true; Severino and Deifell describe the need for multiple exposures and note the

complexities involved in vocabulary knowledge. In comparing online with face-to-face tutoring, they find little difference, however. Alsaif and Masrai (2019) traced both the receptive and productive vocabulary development of a single Arabic EFL learner over eight weeks. The researchers provided the participant with five graded readers, both fiction and nonfiction, containing words within the most common 4,000-6,000 English words, as recommended by the Common European Framework of Reference for the participants' levels (C1). Their participant gained 700 words. First, Alsaif and Masrai identified the number of words acquired by Arabic learners of English in classroom settings. Next, the researchers deducted this number from their participant's gains and found that that their participant learned three times as many words as those acquired through classroom learning alone.

Case studies comparing two participants. Ma (2013) examined the longitudinal learning of two undergraduate students in Hong Kong, finding that the students engaged in similar strategies. Both students read newspapers and kept vocabulary notebooks to record newly encountered words. Both intentionally used newly encountered words in writing. However, the approaches taken by participants to intentionally learn vocabulary also differed in many ways. Although both also incorporated memorization strategies, the specific types of memorization techniques varied. Student A relied on mnemonic devices that associated new words with sounds. The other student, K, tried to use a variety of mnemonic devices but found them less useful, explaining to the researchers that the intentional use of words in writing was more fruitful. Student K used new vocabulary words in emails, while Student A intentionally used new words in more formal contexts. Whereas Student K arranged notebook vocabulary by theme, Student K arranged it alphabetically. Using RANGE software, Ma measured vocabulary differences in essay writing before and after exploring their strategies and presented the data in

comparison to NES essays. Ma reports that for both essays, Student A showed lexical profiles more akin NES to essays than did Student K, leading Ma to conclude that the intentional use of new vocabulary in formal writing is more effective than its use in informal writing. However, it is worth noting that the essays Ma compared were of different genres, and as a result, for both NESs and the participants, lexical profiles were less sophisticated in the second essay, written later. Ma's study, therefore, may demonstrate the power of genre to guide how writers use vocabulary.

Parry's (1997) case study also compared two undergraduate students—one Greek L1 speaker and one Korean L1 speaker. For six weeks, the students were asked to create a list of unknown words they encountered during reading an anthropology textbook. Students were asked to guess the meaning of words and to provide dictionary definitions if they chose to look up word meanings. They were also asked to read a journal article and make a list of unknown words using a think-aloud protocol while being recorded. Two weeks later, the participants were asked to translate two paragraphs from the article into their first languages. The students were then given individualized tests that included all terms the student included on their lists as well as any list the students created on their own. The vocabulary tests were scored based on the ability to provide definitions within and without their original contexts. Parry found surprising results for the two participants. One student, whose L1 is Greek, read 90% more texts than did the Korean student, recorded fewer unknown words, and guessed a word's meaning more often. However, the other student, with a Korean L1, performed substantially higher on the vocabulary posttests. Parry determined that the Korean student had a better memory despite the Greek student's ability to learn more vocabulary. The think-aloud recordings revealed a difference in technique between the two: The Greek student took a holistic approach to interpreting words in context, whereas the

Korean student demonstrated a more analytic approach, attempting to learn the words in isolation. Parry also discovered that the holistic approach was faster, leading to more vocabulary acquisition through reading. During the posttest, students were permitted to provide L1 translations, or what Laufer and Goldstein (2004) refer to as passive recall. Parry (1997) does not differentiate between a student's ability to provide an English definition as opposed to an L1 translation, which Laufer and Goldstein (2004) would consider qualitatively different forms of word knowledge.

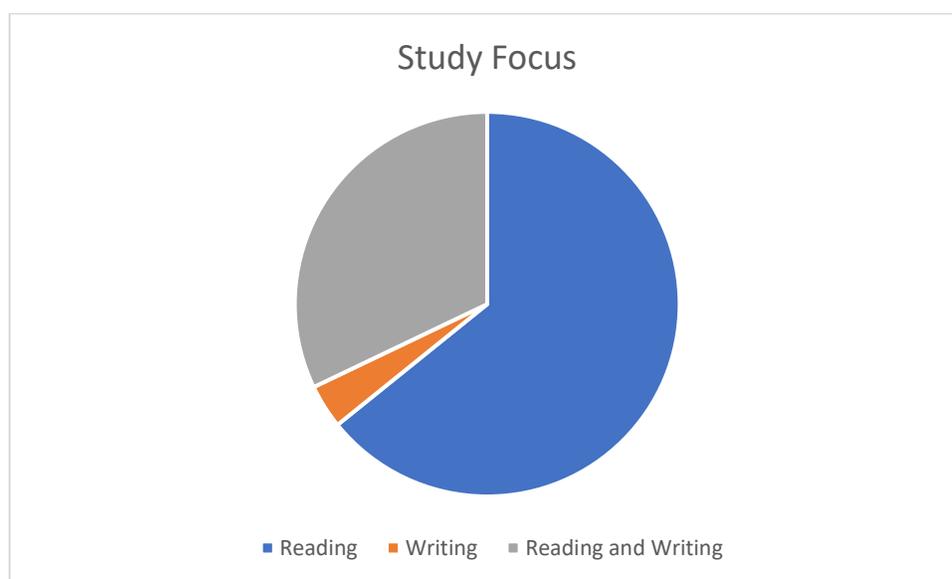
Conclusion. In both bilingual-learning and NES contexts, development studies provide insight into the number of words a learner acquires over time and in some cases, how learners' produce newly learned words in their writing. Only one development study, however, that of Choi (2017), connects the theoretical understanding of vocabulary development to entrenchment through investigating the purposeful production of newly taught words and phrases. Development studies provide a glimpse into differences that occur across age groups and grade levels, including areas of spiking growth or plateaus. On the other hand, longitudinally studies of vocabulary development through reading or writing often compare different cohorts of learners, a design rife with pitfalls. Within educational research, comparisons within even a single cohort of learners challenge scholars with multiple threats to internal validity, which include instructional differences that vary from classroom to classroom. In both bilingual learner and NES contexts, development studies tracing the vocabulary acquisition of a single group of students through reading and writing over multiple years are warranted. Such studies may provide better insight into the nature and impetus of spikes and lulls previous development studies have uncovered.

Vocabulary Acquisition Through Reading

Eighty empirical studies in this review examine the acquisition of vocabulary through reading (65% of these studies), writing (4% of these studies) or reading and writing in tandem (31% of these studies; See Fig. 1).

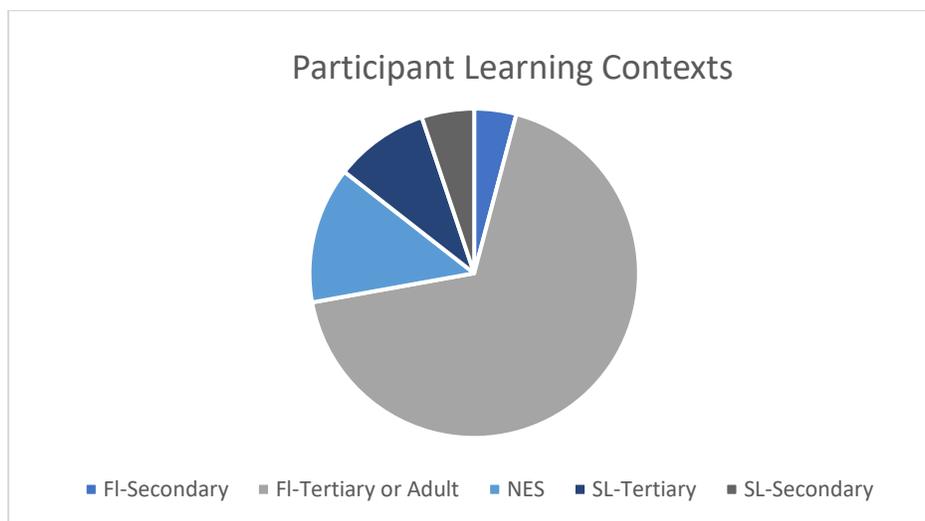
Figure 1

Literature Review Study Focus

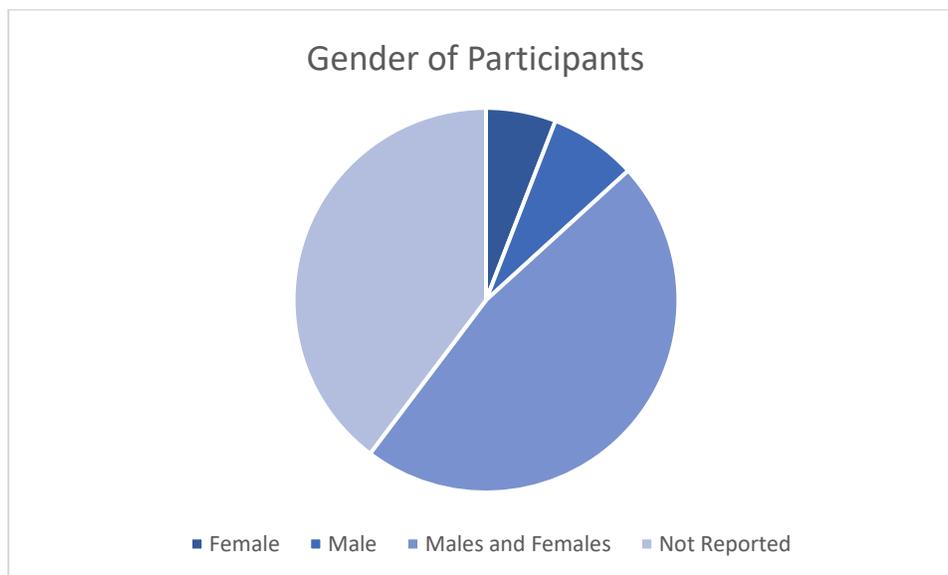


Gloss studies comprise the majority of reading studies ($n = 17$), which represents 21% of all non-developmental studies. Only six studies examined vocabulary acquisition through participant reading followed by writing (Amirian & Behshad, 2016; El-Dakhs et al., 2017; Larson, 2014; Lee & Muncie, 2006; Lee, 2003; Solati-Dekhordi & Salehi, 2016). Appendix A provides additional information about individual studies.

As can be seen in Figure 2, the participants of empirical studies are more often than not university students or adult students learning English as a foreign language (EFL), representing 72% of all studies I examined.

Figure 2*Literature Review Participants: Learning Contexts*

In 90% of the studies I examined, the gender of participants is either not reported or reported to include both females and males. (See Fig. 3.)

Figure 3*Literature Review Participants: Gender*

Theory of vocabulary acquisition through reading. Although incidental learning, or unintentional learning, has been explored since the beginning of the twentieth century (Laufer &

Hulstijn, 2001), its application to vocabulary acquisition seems to have gained ground with help from the incidental vocabulary learning hypothesis, introduced by Nagy and Herman in 1985 and popularized in L2 contexts by Laufer and Hulstijn in 2001 with their introduction of the involvement load hypothesis. The involvement load hypothesis proposes a way to calculate the impact of varied tasks on incidental vocabulary learning. Nagy and Herman (1985) hypothesize that native speakers learn nearly all vocabulary through incidental means, multiple exposures, and in varied contexts. The concept of incidental vocabulary acquisition seems to have held great sway over SLA reading research; studies of participants' incidental vocabulary acquisition through reading continue to abound (e.g., Ajideh et al., 2013; Coady, 1997; Dabaghi & Rafiee, 2012; Eckerth & Tavakoli, 2012; Joe, 1998; Nation & Coady, 1988; Restrepo Ramos, 2015; Reynolds, 2015; Teng, 2016; Varandi & Faezi, 2013, among others). The universality of its acceptance is also evidenced in *Theories in Second Language Acquisition*, a textbook by VanPatten and Williams (2015) describing 12 SLA theories in a chapter devoted to each. Each chapter contains a discussion of how the theory accounts for the universally recognized observation that “a good deal of SLA happens incidentally” (VanPatten & Williams, 2015, p. 9). Krashen (1989) proposes that the focus on meaning that comprehension requires is responsible for this phenomenon. Restrepo Ramos's (2015) literature review of incidental vocabulary acquisition through reading review finds further validity: Learners who are exposed to words through reading are more likely to produce words correctly than those who have learned them out of context.

Studies of vocabulary acquisition through reading. SLA researchers studying vocabulary acquisition through reading—discussed in detail in the following section—regularly undertake quantitative studies to measure gains in targeted vocabulary through

reading alone, often examining the mitigating effects of glossing target vocabulary items (e.g., Duan, 2018; Johnson, 1972), the effect of frequency of occurrence within the text (e.g., Horst et al., 1998; Teng, 2016), or may measure differences in receptive and productive vocabulary knowledge gains through reading (Alavi & Keyvanshekouh 2012; Vela, 2015). They may also examine the impact of motivation (e.g., Ajideh et al., 2013; Zhao et al., 2016) context (e.g., Kweon & Kim, 2008) or proficiency level (e.g., Tekman & Daloğlu, 2006; Zahar et al., 2001) on gains. Most studies, including those undertaken in NES contexts, generally find that reading leads to vocabulary acquisition, with longer and more sustained reading, greater frequency of word occurrence, contextual importance, and glossing all facilitating incidental vocabulary learning for both NESs and bilingual learners. Liu (2018) asserts that once bilingual learner students reach advanced levels, most vocabulary acquisition will take place through extensive reading. Accordingly, studies examining the role of reading alone tend to recommend extensive reading to enhance vocabulary acquisition, as described below.

Multiple independent variables. Nagy et al.'s (1987) foundational study of 212 participants in grades 3, 5, & 7 was one of the first to measure vocabulary acquisition through reading and provides an exemplar for the use of regression analysis to untangle the effect of other individual factors involved in the process. The participants of Nagy et al. read multiple texts of stratified difficulty and were given a battery of tests that included a checklist asking them to identify unknown words, which included target terms from all readings, words coined for the sake of the study, and 50 of the most common words in the English language. Nagy et al. includes genre, text difficulty, whether the term appeared in the readings presented to participants, and contextual support provided by the text as independent variables. Care was taken to ensure that participants were unaware that vocabulary was the focus of the study. Nagy

et al. found support that incidental learning of vocabulary occurs through reading, and they estimate that 1/3 of vocabulary growth occurs through reading for the age group of their participants. They further estimate that a single term offers a 5% chance of such acquisition with one exposure, yet they found no significant effect of the role of context except in expository texts. They conclude that “regular, wide reading must be seen as the major avenue of large-scale, long-term vocabulary growth” (p. 266). Nagy et al.’s study opened the gateway for the numerous incidental vocabulary studies that were to follow, although few would utilize multiple regression to include the numerous variables these scholars incorporated.

Zhao et al. (2016) also employed multiple regression to measure the effect of vocabulary acquisition through reading on numerous variables. Their study of bilingual learners included learner language proficiency, motivation, anxiety, and the use of strategy as independent variables. Zhao et al.’s found support that all variables except motivation predicted word uptake through reading. The researchers developed their own instruments to measure variables and interviewed participants through a focus group to elicit feedback about the instruments’ efficacy. It seems unclear why the researchers did not test their instruments against those with established validity or why they chose to forgo these readily available instruments. A study by Ajideh et al. (2013) also measuring motivation found partial support for Zhao et al.’s (2016) findings. Ajideh et al. (2013) reports that although the use of motivation strategies may initially ease vocabulary acquisition through reading, the gains are not retained long term.

Degree of word knowledge gained as an independent variable. Like Nagy et al. (1985), studies by Ehsanzadeh (2012) Alavi and Keyvanshekouh (2012) and Varandi and Faezi’s (2013) also measured the depth and breadth of vocabulary acquisition acquired through reading. Measuring the two variables independently, Ehsanzadeh (2012) reports that together, depth and

breadth of initial vocabulary knowledge accounts for 66% of success in learning vocabulary through reading. Varandi and Faezi's (2013) similar study takes an innovative approach to measure gains in depth of vocabulary knowledge. Each of five groups were given one of five sections of a depth and breadth test. Participants read an English language short story for pleasure with German target words seeded within it. The researchers found reading comprehension and vocabulary acquisition related to knowledge of both depth and breadth of vocabulary knowledge, and they also found evidence for the Matthew effect: Those who began with higher beginning vocabulary knowledge made greater gains. However, with a total of only 30 participants divided into five groups, comparison groups seem to have had only six participants per group, making any generalizations difficult.

Alavi and Keyvanshekouh (2012) also provide a strong argument for the benefits of prolonged, sustained reading that Nagy et al.'s (1987) study recommends. In Alavi and Keyvanshekouh's (2012) study, both experimental and control groups were required to engage in extended reading practices over a 3 ½ month period, and both groups showed significant gains in productive vocabulary despite having no identified target items. The scholars attribute the larger productive vocabulary gains over receptive knowledge to a ceiling effect, which implies that extensive reading helped to move receptive vocabulary knowledge along the receptive-productive continuum. The productive vocabulary gains were significantly greater than receptive knowledge with an effect size of .6, whereas gains made with the use of Moodle, which provides readers with texts of increasingly greater readability levels, had an effect size of .85. Sun's (2014) similar study used multiple regression to attempt to capture the development of word knowledge through the reading of fiction by using 10 vocabulary measures in a posttest of nonce (made-up) words, each word occurring six times within the text. Sun discovered that while

receptive knowledge gained was unrelated to aspects of the words, productive knowledge was. The productive knowledge gained by participants was confined to form, syntax, and grammar, with meaning much less likely to be acquired, however.

Time spent reading as an independent variable. In a four-month, quasi-mixed method study, Pereyra (2015) measured lexical sequence acquisition through extensive reading. Seven intermediate proficiency level, Spanish-speaking adults read three hours per week and met to discuss the readings, which included graded readers, journals, and novels in varied combinations as chosen by the participants. Pereyra included the time participants spent reading as an independent variable. All participants showed improvement. However, only descriptive quantitative data were presented, and the participant number was small (seven). Furthermore, it may be difficult to untangle how much learning resulted from group discussion and how much resulted from reading alone and what the effect of genre had on findings.

Vocabulary activities as independent variables. I located two studies that compared the effect of reading alone to other vocabulary activities. A study by Min (2008) compared narrow reading to reading enhanced with vocabulary exercises, to find the latter more effective, with some retention long term (defined as three months). Narrow reading, or the reading of thematically related texts, led to high frequency of exposure to key vocabulary items. To account for this presumed frequency in vocabulary exercises, Min took care to ensure that exposure of target terms occurred three or four times within texts. Vocabulary exercises included multiple choice, matching, cloze, and rearranging sentence parts to make syntactically-sound sentences. However, no attempt was made to ensure that thematically related articles participants read included target terms at the same frequency of occurrence found within the vocabulary exercises. Therefore, it is difficult to determine how much retention may have been related to the frequency

of exposure that the vocabulary exercises afforded. Finally, a study by Ponniah (2011) compared intentional vocabulary learning through dictionary learning to incidental learning through the reading of a short story seeded with the same 51 target items. Although both groups showed statistically significant gains from the pre to posttest, the reading group participants were better able to define the words as well as produce them accurately in the writing of sentences. Ponniah concludes that the reading group's focus on word meaning, rather than form, lies behind their success. However, Ponniah does not discuss the nature of the dictionary definitions that the comparison group received.

Cognitive style as an independent variable. Wu's (2018) unique approach to investigating the role of learner factors on vocabulary acquisition through reading measured participants' cognitive styles. Wu describes two cognitive styles, found on a continuum, as *field independent* and *field dependent*. Field independent learners are better able to extract parts from the whole and are considered analytical as opposed to global in thinking. In contrast, field dependent learners are more empathetic and more dependent upon contextual factors, or "external frames of reference" (p. 814). Wu found that field independent learners had significantly greater gains than field dependent learners, although Wu did not report effect sizes.

Context as an independent variable. Bai (2016) measured the impact of context on vocabulary through reading. Fifteen target words were taken from the most common 10,000 English words, four from the Universal Word List (Coxhead, 2000), and a single term was taken from a list of the most common 5000 English word list. Bai divided the 20 words between two groups, although Bai does not explain which group received the single word from the 5000-word list. The participants read an article of 1208 words, which were analyzed to ensure that readability matched the participants' homogeneous proficiency level. In one task, participants

replaced words in a sentence with target words and wrote sentences using the words. In another task, participants matched words to first language translations. Bai only provides descriptive results and does not explain how differences in contextual richness affected gains of specific terms but reports that with increases in contextual richness, vocabulary acquisition improves.

Glossing as an independent variable. Glossing refers to using definitions or translations to promote the noticing of target vocabulary in texts. Studies investigating the effects of glossing techniques seem to be an increasingly common area of research interest. I located 18 studies that examined the use of glosses to aid in bilingual learner vocabulary acquisition through reading. All have found glossing to have measurable benefits over reading without glosses when control groups were included or when the study compared glossing to no glossing conditions (e.g., Jung, 2016). Otherwise, results were mixed, often depending upon learner and text factors. For instance, among several studies comparing L1 to English glosses (e.g., Alharbi, 2018; Azari et al., 2012; Cheng & Good, 2009; Liu, 2017; Xu, 2010), some have found results dependent upon type of vocabulary knowledge gained (Dabaghi & Rafiee, 2012) or proficiency level of participants (Yusuf, 2014; Liu, 2017). Notwithstanding Schmitt's (2010) assertion that L2 glosses are preferable to L1 glosses, the studies examined in this review have generally found L1 glosses more effective, yet in cases where students are provided with choices (e.g., Alharbi, 2016; Azari et al., 2012; Fathi & Sarkhosh, 2019; Lenders, 2008), students performed better with both choices, and they expressed a preference for having both available. In the case of Fathi and Sarkhosh (2019), this included images. Other glossing studies have compared gloss locations (e.g., AbuSeileek, 2011; Zarei & Hasani, 2011), finding interlinear glosses best. When electronic gloss types are considered, multiple choice glosses are found to be most effective (e.g., Duan, 2018; Yoshii, 2013). One study, that of Far (2016) compared the impact of text length, genre,

and text difficulty on the efficacy of glossing. Far found that glossing was most effective in short easy texts, but within long texts, glossing provided a greater advantage when texts were expository. See Appendix A for a more detailed overview of these studies.

Frequency as an independent variable. Hilgard (1956) reveals an early understanding of learning as cognition by depicting learning as a change in “cognitive structure” (p. 274) and notes that repetition of content may sometimes be sufficient to achieve this change in structure. The understanding of repetition’s role in learning has influenced vocabulary studies of reading through a concern for the frequency of occurrence of words or phrases within a text. For example, in emphasizing the importance of reading for both grammar and vocabulary, Albay (2017) states, “Constant repetition of the words and structures allows learners to process them faster” (p. 177). Similarly, Stahl and Fairbanks’ (1986) review of vocabulary studies demonstrates that vocabulary instruction improves reading comprehension when learners have multiple exposures to vocabulary items, and Wright and Cervetti (2016) likewise propose that vocabulary teaching that focuses on frequency are the most successful of incidental vocabulary activities in studies they reviewed.

Accordingly, several studies have gauged the effect of multiple exposures of target items embedded within text to attempt to determine a precise number of exposures needed for a learner to acquire vocabulary through reading. Closely associated with incidental vocabulary acquisition through reading are, therefore, determinations of how frequency of occurrence impacts such acquisition. Although some L2 researchers strive to determine a statistically significant, magic number of exposures to a target word that will predict levels of its acquisition—suggested numbers include six, eight, over 10, over 15, 18, and 20 (Schmitt, 2010)—as Laufer and Rozovski-Roitblat (2011) emphasize, “there is not definite number of encounters that is thought

to ensure some kind of learning” (p. 294). Chen and Truscott (2010) similarly aver that “The goal of research should be not to identify a definitive number of exposures needed but rather to understand a complex process involving multiple, interacting variables” (p. 694). Therefore, vocabulary acquisition may be better thought of as a gradual, “cumulative process” (Nation, 2001, p. 6). Ellis (2011) employs a metaphor of sampling in statistics to explain acquisition through frequency, implying that attention to frequency is required in effective vocabulary instruction.

In a true experimental study conducted by Heidari-Shahreza et al. (2014), frequency of word exposure was found to correlate strongly to both receptive and productive knowledge. Three groups were exposed to one, three, or seven exposures to 20 target vocabulary words through the reading of 13 texts, and significant difference was found between the first and third group for all of seven vocabulary knowledge measures. In both receptive and productive knowledge of meaning, significance was found among all group comparisons. Among lower-intermediate to advanced participants reading a graded reader, Waring and Takaki (2003) examined six frequency bands to measure the frequency effect on receptive and productive knowledge, with productive knowledge measured through the ability to provide translations. They discover that 4% of terms introduced through reading were retained after three months—but only when encountered eight times. Laufer and Rozovski-Roitblat’s (2011) 13-week study compared the receptive vocabulary acquisition through reading, with and without vocabulary activities. Using a single group with exposure, the researchers found that the effect of frequency is more pronounced when reading is accompanied by vocabulary activities. However, it is unclear how the frequency of occurrence was dispersed over the thirteen weeks.

In a unique multiple regression study reported by Reynolds (2016), Reynolds (2014), and Reynolds (2015), the researcher compared NESs to bilingual learners over two weeks, finding that frequency of occurrence of target terms were more impactful for bilingual learners than for NESs. Reynolds includes, as a variable, what he terms “patternness,” or the word’s appearance as part of a common lexical sequence. Target words were chosen that had a frequency of two or three and were then replaced by nonce words in the text. In both groups, words with higher frequency of occurrence were acquired more readily. Reynolds also found interaction effects between frequency and word form, leading Reynolds to suggest that while frequency plays a key role in vocabulary acquisition through reading, the effect is more pronounced when the words appear in the same form or when they vary only inflectionally. He concludes that frequency through reading alone may not be an efficient means to learn vocabulary. Therefore, he recommends glossing to aid acquisition through reading. Other than an examination of interaction effects, Reynolds’ detailed descriptions of other variables, such as lexicalization, cognateness, and word length, are outside the scope of this review.

Frequency studies: Nonce, rare, or non-target language terms. Several studies have examined frequency with the use of either nonce words or foreign words embedded within a text. These words are chosen or created with an assumption that the terms would have no familiarity to the readers at the beginning of the study. For instance, Horst et al. (1998) notes that several studies have used *Clockwork Orange* by Anthony Burgess (e.g., Pitts et al., 1989; Saragi et al., 1978), to examine the propensity of readers to acquire vocabulary with high frequency because of the novel’s use of Russian vocabulary items. With these studies in mind, Horst et al. (1998) undertook a frequency study among low intermediate EFL learners reading graded readers over 10 days. The authors found that frequency of occurrence was a predictor of learning, whereas

word commonality was not a significant factor in learning 45 target words. On average, learners acquired five words. This is greater than the findings of most *Clockwork Orange* studies, perhaps because Horst et al.'s (1998) participants possessed latent knowledge of some terms, something not possible with nonce or innovative words such as those found in *Clockwork Orange*. Teng's (2016) three-month study incorporated nonce words into simplified texts to measure the impact of six frequency of occurrence bands, ranging from 1-22 occurrences, on depth of vocabulary knowledge among "relatively proficient" learners of English (p. 57). Teng found that over a three-month period, word retention dropped dramatically despite repeated readings of the text during each session. However, Teng's reliance on so many nonce words may have confounded results; I calculated that Teng's text would have included 417 nonce words, or 3.6% of the text.

Using the same principle as those of the *Clockwork Orange* studies, Pellicer-Sánchez and Schmitt (2010) examined how well participants may acquire Ibo terms through encountering them in Chinua Achebe's 1959 novel *Things Fall Apart*. With the smattering of Igbo terms throughout, Achebe's novel ensured that all participants read an authentic work of fiction, yet exposure to unknown vocabulary words without a pretest measurement. Depth of vocabulary knowledge gained was measured thorough one-on-one interviewing. The participants learned 28% of target items, with most gains in meaning recognition. Meaning recall was much lower, at only 14%, with words acquired after appearance within the 5-8- word exposure band.

Frequency studies: The role of learner proficiency. Three frequency studies measured the impact of bilingual learner proficiency levels. For example, Tekman and Daloğlu (2006) measured the effects of proficiency levels of three frequency of occurrence bands and word occurrence on single word acquisition: intermediate, upper intermediate, and advanced proficiency. All proficiency levels showed significant gains, with differences increasing with

level of proficiency, and after one week, no significant losses were found in any group. However, the same text was used for all students, regardless of proficiency level. Tekman and Daloğlu (2006) found 29% of variance among groups of differing proficiency levels dependent upon the frequency of word occurrence throughout the text. Similarly, Daskalovska (2014) measured the impact of frequency, initial word knowledge, and proficiency level on the acquisition of five target words through reading. The results did not find a significant correlation between frequency of occurrence and propensity for learning, nor was there a significance of participant proficiency level. No relationship was found between general word frequency and acquisition, although a multiple regression analysis was not conducted to tease out the effects of either learner proficiency or frequency in text.

Frequency studies: The role of context. Zahar et al.'s (2001) study of male secondary bilingual learner students in Canada measured context and frequency among participants of bilingual learners having five proficiency levels. However, regardless of proficiency levels, learners were asked to read the same text: a graded text designed for intermediate bilingual learners, and although Zahar et al. did examine items with this concern in mind, their concern was confined to the general frequency of target words in the English language. Therefore, it may be of little surprise that Zahar et al. found proficiency level a mediating factor for the impact of frequency, and they did not report their finding's effect size. Two additional studies also measured the impact of context joined with frequency. Kweon and Kim (2008) examined extensive reading of 4-6 hours daily for five weeks on vocabulary acquisition. Context was the prevailing factor for whether terms were acquired. Moreover, when word knowledge was required for understanding, Kweon and Kim found that words with lower—not higher—frequency of occurrence were more readily acquired. Overall, the words most common in the

English language were better acquired. However, students wrote in journals after each session, although the role of journal writing or the appearance of target times in such writing was left unexplored. If target terms were required for meaning construction, these terms may have been key for journal discussions as well. Like Kweon and Kim's study, Webb's (2007) true experimental study of intermediate learners measured the effect of four frequency of occurrence bands, with each presentation decreasing with contextual cues, through reading on both the receptive and productive knowledge of multiple aspects of word knowledge: orthographic, semantic association, meaning, syntax, form, and function. Webb found the greatest gains in receptive knowledge of orthography, while the lowest scores were found in the productive knowledge of semantic associations. In general, higher gains were found for receptive knowledge than for productive knowledge in all dimensions, but with increased repetition, the gains were also increased.

Frequency studies: The role of lexicalization. Like Webb, Heidari-Shahreza and Tavakoli (2016) examined frequency and learner L1 lexicalization on both the receptive and productive knowledge of 10 target words' semantic associations, meanings, parts of speech, and orthography, finding semantic information the most difficult to acquire. Each of three groups read 13 passages, with the number of passages seeded with target items (as frequency of occurrence) dependent upon the group number. The types of knowledge gained followed similar patterns, yet gains increased as frequency increased. However, productive knowledge of associations at seven occurrences had comparatively greater gains than three occurrences or one occurrence.

Frequency studies: The role of dispersion. No studies I have so far discussed examined word frequency in terms of dispersion. Sometimes referred to as "distributed learning" (Vlach &

Sandhofer, 2012), or “spaced learning” (Smolen et al., 2016), *dispersion* refers to patterns of vocabulary recurrence exposed to learners over time. As Alcaraz-Marmol, (2015) argues, word distribution is as equally important as word frequency for input, yet the researchers note that frequency studies often ignore this element. In a study of toddler word learning, Childers and Tomasello’s (2002) found that productive word knowledge among toddlers was more likely to occur when participants were exposed to a new word once a day for four days (spaced) than when they were exposed to the same word eight times in one day (massed). The study demonstrates that raw frequency of word encounters over a short amount of time is trumped by long-term, sustainable exposure to words over longer periods of time, and this finding holds true for second language learners as well.

Çekiç and Bakla (2019) and Sobel et al. (2011) are two rare studies examining the role of dispersion in tandem with word frequency for the sake of vocabulary acquisition. Çekiç and Bakla (2019) compared the acquisition of 20 targeted items in three experimental groups: one exposure per week for over nine weeks, three exposures per week for three weeks, and three exposures per week with increasing intervals: one week, three weeks, and four weeks later. Çekiç and Bakla found that as the exposures were increasingly dispersed over time, significance was found with increasingly greater effect sizes. Sobel et al.’s (2011) one-month study of 46 middle school students in Ontario exposed students to eight rare vocabulary words in both massed and spaced learning conditions. Students were provided with the word and its meaning and asked to write the definition and to use the word in a sentence. During massed learning conditions, students completed the activity a second time one minute later. During the spaced learning condition, students completed the activity a second time one week later. Five weeks later, researchers asked students to define all eight words and found that participants could recall

177% more definitions when exposed to the targets during the spaced condition. However, it is unclear whether the massed condition took place during the first or second week of the study. If the massed condition took place one week before the second exposure to the spaced words, it could be argued that the participants better remembered words that they had had more recent exposure to, whether massed or spaced.

Vocabulary Acquisition Through Writing

As Vidal's (2012) review of language maintenance reveals, even L1 speakers "show signs of loss in lexical richness after periods of disuse" (p. 55). She further asserts that when language is lost, "the problem is that words have not been consolidated" (p. 55). Words taught must be nurtured to lead to such consolidation, and for that to occur, usage is required (Swain, 1985). I located twenty-six studies examining the acquisition of vocabulary items through writing or entrenchment of words exposed to learners through reading followed by writing.

Theories of Vocabulary Acquisition Through Reading and Writing

Theories supporting vocabulary acquisition through writing, which are discussed below, therefore describe learner language output (e.g., Ellis, 2011; Swain, 1985), examine the role of cognition to enhance memory through noticing (Baddeley & Hitch, 1974; Godfroid, 2013; James, 1890), or engaging in tasks while using language (Craik & Lockhart, 1972). In 1985, Merrill Swain established the output hypothesis when she noticed that language learners who are not asked to produce language become proficient only in receptive language. Swain suggests that language acquisition requires noticing, hypothesizing and testing, metalinguistic reflection, and use. To Swain's output hypothesis, Ellis (2011) later added the concept of "learned attention" (p. 83) and the notion that language use results in the building of associations.

Noticing. The importance of attention in the learning process has been recognized for centuries, as William James (1890) demonstrates. James was an early theorist who connected noticing, in terms of attention, to memory. Studies of noticing have found it is needed for vocabulary acquisition to occur, perhaps because of the processing involved. In an eye-tracking study by Godfroid et al. (2013), for instance, the longer advanced bilingual learner students gazed at an unknown word while reading, the more likely they were to gain receptive knowledge of the word. Moreover, Choo's (2012) review of four studies found that regardless of whether explicit or implicit, learning was mainly "a matter of selective attention and elaborated processing" (p. 857), with intention less important. Schmidt (1990) suggests that noticing, defined as attention, is necessary for implicit language learning to occur and in some cases may be sufficient, given the facilitative role of language. It should be noted that although such noticing is required for language learning to occur (Alcón, 2007; Baddeley & Hitch, 1974; Carroll, 2001; Ellis, 1991; Ellis & Wulff, 2014), this does not preclude the possibility of incidental learning. In L1 literacy research, Hilden and Pressley (2011) emphasize the role of noticing for expert reading, calling such a reader "massively active" (p. 428), for example.

Working Memory and the Phonological Loop. Baddeley and Hitch (1974) first proposed the complexity involved in working memory as a challenge to the accepted multi-store memory model of their time. With Baddeley and Hitch's model emerged a more complex working memory—that part of memory engaged in real-time language use and acquisition as well as its components, the phonological loop and its "subvocal rehearsal system" (Baddeley et al., 2019, p. 578). The working memory model allows for both visual and auditory input to occur concurrently with no cognitive disadvantage. In fact, Baddeley et al. (1998) assert that vocabulary *acquisition*—not use—is the primary purpose of the

phonological loop. Consequently, Hummel and French (2010) suggest that learners relying solely on oral input are disadvantaged in the missed opportunity to enhance form-meaning connections through both oral and visual processing that language use requires for most people. The subvocal rehearsal system provides space for vocabulary acquisition to occur (Baddeley et al., 1998) through a similar, albeit internal, repetition that usage-based theorists believe results in entrenchment. The phonological loop system “mediates the acquisition of syntactic knowledge (through) a storehouse of multiword language patterns” (p. 161). These patterns, Baddeley et al. argue, provide L1 learners with models of rules governing usage. Baddeley et al. also assert that “the phonological loop appears to provide a critical input to the construction of the more permanent phonological structures that are stored in the mental lexicon” (p. 163). The phonological loop relies upon knowledge of phonology. This implies that English learning may be more difficult in non-English speaking countries, where phonological patterns may be less familiar to the learner.

As Grabe (2004) notes, an understanding of working memory is needed for any examination of language processing involving reading and writing. In the context of vocabulary acquisition, this requirement is readily met by the research of Craik and Lockhart (1972) and Baddeley and Hitch (1974). In 1972, Craik and Lockhart presented the levels of processing model to become “one of the most highly cited in the history of cognitive psychology” (Baddeley & Hitch, 2017). Craik and Lockhart’s model, linking the degree of cognitive processing to strength of retention in long term memory seems to have spawned multiple avenues of research in L2 vocabulary acquisition by supporting the negotiation of meaning underlying Long’s (1981) interaction hypothesis, Swain’s (1985) output hypothesis,

Schmidt's (1990) noticing hypothesis, Robinson's (2001) cognition hypothesis, and Laufer and Hulstijn's (2001) involvement load hypothesis.

Robinson's Cognition Hypothesis. Robinson's (2001) cognition hypothesis holds that the more demanding or complex a task is, the greater the propensity for a learner to acquire the language used in the task, but his cognition hypothesis grants that the "quantity and quality" are important variables (Robinson, 2003, p. 3). Task-based approaches to language teaching seek to engage learners in language while focusing on an otherwise unrelated activity, making it both meaning-focused and incidental. The place of the cognition hypothesis in task-based learning is supported by Baddeley and Hitch (1974), whose model allows for auditory language use to occur alongside other cognitive activity, particularly visual.

Interaction Hypothesis. Michael Long's (1981) interaction hypothesis holds that language is acquired through interaction, specifically through opportunities for negotiation of meaning to occur during language-related episodes. In Long's understanding, negotiation of meaning occurs through feedback in the form of confirmation checks, clarification requests, and recasts. Interaction theory therefore highlights noticing (Gass & Mackey, 2014). Because meaning-focused activities are thought to better facilitate vocabulary acquisition than form-focused activities (Khonamri & Roostae, 2014; Nation & Chung, 2011), Long's (1981) interaction hypothesis supports incidental vocabulary acquisition through reading and writing by highlighting the focus on meaning as key.

The Involvement Load Hypothesis. Laufer and Hulstijn's (2001) foundational study introduces the "task-induced involvement load" through the regularly tested involvement load hypothesis, which positions noticing as degrees of elaboration and hypothesizes a correlation between degree of processing required of an activity to vocabulary retention. Laufer and Hulstijn

characterize three elements of a vocabulary activity, “need, search, and evaluation,” with need described as in terms of motivation and with forced output through writing considered highly involved. The involvement load hypothesis assigns an index to an individual vocabulary activity based on how in-depth need, search, and evaluation an activity requires from the learner. In brief, the involvement load hypothesis supports the encouragement of a student’s use of new vocabulary through writing because this exercise requires depth in processing required for entrenchment. Involvement load studies are incorporated within my review only when measured tasks include a writing component.

The first significant support of the involvement load hypothesis to aid incidental vocabulary acquisition comes from Hulstijn and Laufer (2001) through a study concurrently undertaken in Israel and the Netherlands. Their study tests both short and long-term vocabulary gains subsequent to intervention and finds incidental vocabulary acquisition most likely when students were asked to write using the targeted vocabulary. Since then, a stream of experimental involvement load hypothesis research studies has been steadily maintained, with nearly all supporting incidental vocabulary acquisition enhanced through reading and writing (Eckerth & Tavakoli, 2012; Laufer & Rozovski-Roitblat, 2011; Mancilla-Martinez & Lesaux, 2010; Ong & Zhang, 2010; Paribakht & Wesche, 1997; Pichette et al. 2012; Varrick, 2016), sometimes finding writing more impactful, especially long-term (Bao, 2015; Cao, 2013; Ghorbani & Rahmandoust, 2012; Mancilla-Martinez & Lesaux, 2010; Pichette et al., 2012). Other studies have found that regardless of involvement load, writing supports long-term retention of vocabulary acquisition (Kim & Taguchi, 2015; Lee & Muncie, 2006; Pourakhari & Biria, 2015). Lee (2003) finds that reading and writing lead to 13% higher use of incidentally acquired vocabulary in their writing, with little decay, than either alone.

Technique Feature Analysis. Nation and Webb (2011) reconceptualized the involvement load hypothesis, extending it from three to five elements examined in a given vocabulary task, with each of the five components—motivation, noticing, retrieval, generation, and retention—rated on a scale of 0-4 and resulting in an index score for each activity. Nation and Webb recommend vocabulary activities that engage the learner such that the learner is motivated to learn, and they suggest the use of glossing or boldening of target items to promote noticing. Hence, in Nation and Webb's reckoning, glossed reading passages are accorded higher points than are unglossed target items. Retention refers to the linking of form with meaning, so the linking of form and meaning adds a one-point value to an activity's index. Perhaps another significant improvement made by Nation and Webb lies in the consideration of word meaning retrieval in terms of single or multiple, spaced or massed. Within Nation and Webb's operationalization of retrieval, issues of frequency of occurrence, dispersion, and the advantage of glossing are therefore acknowledged.

Williams (2012) argues that writing can enhance L2 learning because it affords needed time for L2 language processing and because the planning and writing encourages writers to seek explicit information. Accordingly, nearly all studies incorporating writing for the purpose of English vocabulary acquisition find writing to be effective. One rare exception is that of Barcroft (2006), whose investigation found that writing may thwart acquisition when students are asked to write while a teacher orally presents meaning. However, when understood in terms of Robinson's (2001) cognition hypothesis, this outcome seems natural; the processing required for encoding form and meaning are consumed through such split attention. The studies I critiqued do not require students to split attention between writing and listening by taking notes of vocabulary

meaning presented orally but instead ask participants to independently and creatively use vocabulary items in writing, whether by constructing sentences, paragraphs, or essays.

Empirical Studies

Notebooks. Two studies, Larson (2014) and Walters and Bozkurt (2009), incorporate the use of notebooks to facilitate vocabulary acquisition through writing. Larson's (2014) mixed method study incorporates the experience sampling method, a method of motivation data collection acknowledging that motivation is dynamic. The target terms selected were from three frequency bands, which the researcher discusses as tiers, based on the researcher's instinct. The activities were developed to motivate students through engagement. Students wrote in "inquiry notebooks" (p. 299), read informational texts, and wrote responses. Teachers asked students to try to incorporate newly taught vocabulary in their writing. Larson's study incorporated critical literacy practices, but Larson defined critical literacy as "engaging activities to allow students to use credible evidence, knowledge, and creativity to generate and demonstrate understanding of authentic ways to address a real-world problem" (p. 303). Larson reports descriptive data about the gender, ethnicity, socioeconomic status, bilingual learner status, and disability status of the participants in both groups but does not measure the impact of these factors on results except to say that because gender was unbalanced among the groups, the study controlled for gender. Therefore, data are not reported on whether any interaction effects between bilingual learner status or gender occurred. Larson found that the writing of experimental group participants was infused with greater use of academic vocabulary, and these students demonstrated greater knowledge of content material.

Walters and Bozkurt (2009) measured the acquisition of 72 target terms through a vocabulary notebook requiring lower intermediate EFL participants to define words and to write

weekly essays. Walters and Bozkurt found that students who kept vocabulary notebooks in addition to normal vocabulary classroom activities were more likely to use the target words, and they had significant gains in both receptive and productive vocabulary, with higher productive gains. The researchers used two control groups to mitigate teacher effect: each group was a single class with its own teacher. All target items were the focus of attention for all groups. An informal intervention study by Khan (2019) conveys the process that the researcher used to encourage students to use common lexical sequences in their writing. After introducing the Corpus of American Contemporary English, Khan encouraged participants to use the cluster feature to examine how the word “conclusion” is used. Khan states that the understanding of the formulaic nature of language led students to incorporate these phrases into their writing, thereby improving their writing proficiency.

Reading and Writing. One consequence of understanding how vocabulary acquisition occurs appears in research examining the effect of reading exposure followed by writing to enhance vocabulary acquisition. As Jesson et al. (2011) argue, deep understanding occurs through the intentional transfer of knowledge. Therefore, Jesson et al. and Ahmed et al. (2014) both assert that vocabulary acquired through reading must be intentionally transferred to productive knowledge through writing to enhance vocabulary acquisition. Graham and Herbert (2010) recommend asking students to write regularly to improve reading and to respond to text with writing, and entrenchment and usage-based theory support these suggestions. Finally, Corson (1997) states that “words are only fully learned when they are available for active use,” adding that “something more is needed to promote wider and greater active vocabulary development” (p. 699). Corson’s call may be answered through writing coupled with reading. In one such study, Khatib and Faruji (2012) measured the effects of reading an authentic short story followed

by writing in the form of a story map, finding that the group who created story maps demonstrated significantly greater knowledge. It is unclear whether the vocabulary test measured receptive or productive knowledge.

Joe (1998) may provide the first example of a vocabulary acquisition study applying the depth of processing model to examine vocabulary acquisition through reading and generation, although in the case of Joe's study, productive knowledge occurs through the retelling of a story. Undertaken before the introduction of the involvement load hypothesis by Laufer and Hulstijn (2001), Joe based her study on the generative model, which holds that elaboration of concepts contributes to learning. Joe's study is significant in that it accounts for the impact of background knowledge on vocabulary acquisition; participants who had greater background knowledge of the text subject had greater vocabulary gains through reading.

Amirian and Behshad (2016) examined the effects of using models of NES narratives to improve vocabulary among native Persian speakers' writing proficiency. Two groups of participants, intermediate and advanced level students, were asked to write in response to a picture prompt, followed by a self-reflection of their difficulties in using both English phrases and words. Next, students were asked to revise their narratives after first reading the model narratives. Two months later, participants were unexpectedly asked once again to write a narrative using the same picture prompt. Amirian and Behshad's study demonstrates the importance of noticing as well as the potential for students to improve writing through reading exposure. However, it is unclear which aspects of student writing improved; student essays were given holistic scores only.

Reading and Writing: Lexical Sequence Studies. The participants of a study by Lee and Muncie (2006) read an informational text as part of a unit about the Titanic. During the reading,

investigators stopped to define targeted items and distractors throughout, using “interaction negotiation” as a class reading (p. 300). Two days later, students wrote a response to a prompt asking them to pretend to be a victim of the sinking and to write about their experience. Post writing was judged by raters to be better than the pretest, although the researchers note that “attempts affected grammaticality” (p. 304). Lee and Muncie found significant improvement between drafts for below the 1000 most common English words and found significant improvement for words above the 2000 most common words. However, in the 1000-2000 most common word band, no significance was found. The researchers conclude that because students voluntarily used only 20% of target items, writing does not result in automaticity of newly taught words in writing. In immediate posttesting, participants did better on the use of single items than for lexical sequences, but using Lee and Muncie’s data, I calculated only a 2% difference. Further, for single items, entire word families were calculated, while for lexical sequences this was not possible. Accordingly, participants would have had more exposure to the single word items than for lexical sequences. Nevertheless, in delayed vocabulary recall, lexical sequences were more commonly retained than were single term items. Arguably, the activities may have included too many learning variables to untangle the effect of writing itself on vocabulary acquisition in Lee and Muncie’s study. Finally, the researchers do not explain how similar the NES activities were to the bilingual learners except that NES were not pre-taught vocabulary items.

El-Dakhs et al.’s (2017) study of female students enrolled in one university’s intensive English program employed a quasi-experimental design to compare the effects of teaching lexical sequences to teaching individual words on the participants’ writing quality. Although they find evidence of writing improvement, El-Dakhs et al. finds no evidence of the use of

targeted lexical sequences in their participants' writing. However, it may be important to note that the treatment involved the exposure of phrases through news reports, while the assessments measured the use of lexical sequences in narrative writing. Several scholars have noted that the use of lexical sequences is tethered to genre (Biber et al. 1999; Breeze, 2013; Gil & Caro, 2019;), and as Granger (2014) states, "quantity and quality of (lexical sequences) are highly sensitive to genre" (p. 63). Studies such as El-Dakhs et al. therefore may not accurately capture word knowledge gained through one genre when asking participants to respond in a genre much different. Cao (2013) also examined the effect of involvement load on lexical sequences acquisition, finding that sentence writing led to the greatest gains than two tasks with lower involvement loads, with differences greatest in delayed posttesting defined as one week later.

Reading and Writing: The Role of Word Concreteness. Measuring the effect of word concreteness on productive vocabulary learning, Pichette et al. (2012) compared the efficacy of reading to the efficacy of writing to enhance vocabulary acquisition of 16 rare words.

Participants either wrote three sentences for each target word or read three sentences for each term, with a target item appearing as a different sentence part or placement each time. The researchers found that in delayed recall, initial advantages of writing over reading of abstract terms were lost, although concrete words remained. However, the list of terms used by Pichette et al. among French speaking participants included some terms with clear similarities to French terms (e.g., cognates) while others did not. Such relationships were not investigated.

Reading and Writing: The Role of Working Memory. Yang et al. (2017) compared productive vocabulary activities on the acquisition of target vocabulary after reading and calculated participants' working memory to measure its impact. In one group, participants wrote sentences of more than seven words using the target words after being provided with definitions.

In another group, participants engaged in a cloze activity, choosing the correct item from a list of choices. However, the remaining groups may not have contributed to understanding the role of productive use towards vocabulary acquisition after exposure through reading. For instance, the control group did not read the text, and the essay-writing group did not require or encourage participants to use the target words in the writing of the essay. Yang et al. found that although sentence writing led to greater gains than the cloze activity in the short term ($d = .72$), gains from the cloze activity were better retained long term. Gains were reported to have large effect sizes over the comprehension and control groups ($d = 1.76$; $d = 2.68$). Working memory predicted 20% of gains on comprehension and cloze exercises. This finding led Yang et al. to suggest that completing those two activities required participants to comprehend meaning within context, while sentence-writing required participants to produce new terms in a new context in addition to the original context, which they believe overrode working memory effects.

Reading and Writing: The Involvement Load Hypothesis. Several writing studies of vocabulary acquisition test the involvement load hypothesis. The involvement load hypothesis is generally confirmed by such studies, which are nearly all studies of bilingual learners. Many involvement load hypothesis studies, however, take innovative approaches or test mitigating factors on the involvement load hypothesis. One example is that of Bao (2015), who measured the effect of involvement loads on receptive and productive vocabulary knowledge among 158 participants. All participants read 18 target terms within 18 sentences. Participants were then asked to perform one of four tasks or were part of a control group asked to complete matching exercises. The four required tasks included a cloze exercise with definitions provided, combining of phrases to create native-like sentences, translating of sentences into participants' first language (Chinese), or using target words to write sentences. Notwithstanding the combining

task, other tasks requiring vocabulary use outperformed the control group, with medium effect sizes, on the immediate productive vocabulary posttest. Bao explains that providing definitions led to greater increases in receptive vocabulary knowledge. In other words, tasks requiring usage led to greater productive knowledge, while tasks requiring understanding led to greater receptive knowledge. A more recent study by Alavinia and Rahimi (2019) followed Bao's procedures, but unlike Bao, Alavinia and Rahimi found that for both receptive and productive knowledge, the writing of sentences led to greater gains.

Two studies, that of Soleimani and Rahmanian (2015) and Touti and Maleki (2016) tested the involvement load hypothesis by comparing the same two tasks: cloze exercises and reading comprehension exercises with the use of glosses. The participants of both studies were teens and young adults taking English classes at an institute in Iran. The proficiency levels of participants of the studies were advanced and intermediate, respectively. In Soleimani and Rahmanian's (2015) study, both groups initially showed significant gains, but in the case of reading comprehension, all gains were lost during delayed posting (defined as two weeks). In contrast, cloze exercises resulted in no significant difference between initial and posttest results. Touti and Maleki (2016) report similar results, with cloze exercises yielding greater gains. Touti and Maleki, however, do not note the time lapse between the pre-test and delayed posttests except to say that after each session, students were asked to take posttests and were given a comprehensive posttest after the treatment period.

Although Polio and Williams (2011) position writing as form-focused, writing and vocabulary activities testing the involvement load routinely rely on meaning-based activities. In Baba's (2009) incidental vocabulary acquisition study incorporating writing tasks, Baba discovers that writing the definitions of words "made a unique contribution over and above the

other variables” (p. 191), with higher reading comprehension and summary lengths associated with greater summary writing proficiency. It should be mentioned that studies like that of Baba and Kim (2011), discussed below, are unique; research testing the involvement load hypothesis nearly always defines writing as the writing of a single sentence. Through two experiments, Kim (2011) measured the effect of learner proficiency on vocabulary acquired through the involvement load hypothesis, comparing matriculated bilingual learner undergraduate students with young adults in intensive English programs. During experiment 1, students of multiple first languages read modified texts taken from a textbook with target terms glossed in the margins. Unlike involvement load hypothesis studies asking only for sentence writing, Kim’s study is unique in its requirement to write essays. Kim found that composition writing resulted in the greatest gains over other tasks, with a moderate effect size (.40). In experiment 2, differences between essay and sentence writing were measured. Although no significant effects were found, Kim explains that simply using the total participant score from the Vocabulary Knowledge Scale (VKS) may be misleading; a learner with a total score of 20 may have little familiarity with any words, while a score of 18 may represent a student with the ability to productively use some items.

Lin and Kawai (2016) report that through a five-week intervention, they found support of the involvement load hypothesis on receptive and productive vocabulary knowledge. Although few other details of the study are reported, Lin and Kawai explain that participants wrote a 150-word story after being given 148 targets to help facilitate the writing task and with the requirement to include at least 20 target words. The researchers found a 76% gain in productive knowledge of terms participants previously had receptive knowledge about and a 77% gain for terms unfamiliar to them. The researchers interpret their findings as supporting that the “process

of vocabulary knowledge is a development continuum” (p. 274). Students also revised essays after instructional feedback.

Holding task involvement loads constant by choosing tasks with equal loads, Hazrat (2015) compared the vocabulary acquisition of 10 low-frequency target words from three task types: reading, writing, and speaking. He found that the reading and writing groups scored significantly higher than did the speaking group for both receptive and productive vocabulary knowledge. Although differences were not statistically significant, the writing group scored higher than the reading group for receptive vocabulary knowledge. Hazrat’s study therefore provides evidence that reading and writing tasks, when noticing is required, may contribute to vocabulary acquisition; the texts used with the reading group promoted noticing by using bold print. Hazrat explains that the participants “needed to circle the best option” of the target words in bold (p. 83), which presumably required them to search for word meaning. Pourakbari and Biria (2015) examined the involvement load hypothesis through six tasks of varied involvement load, three of which were receptive tasks and the other three, having equal involvement loads, were productive. The researchers generally found support for the involvement load hypothesis, although productive tasks lead to greater gains than receptive tasks of the same involvement load index.

Finding that involvement load and frequency of occurrence had only been examined independently, Eckerth and Tavakoli’s (2012) inventive study compared interaction effects of receptive and productive knowledge (active recognition, active recall, passive recall, and passive recognition) of word knowledge gained through extensive reading of expository texts. Although all groups undertook the intervention over a three-week period and were exposed to the same tasks—with immediate posttesting after each task—the interventions did not take place

concurrently, with three different texts having different target items embedded within. This design allowed the researchers greater participant numbers for each task, and because all participants took the same delayed posttest concurrently, Eckerth and Tavakoli were able to incorporate delay time as an independent variable; the delay time ranged from three to five weeks, depending upon the task and group, whereas the frequency bands were either one occurrence or five. In effect, all participants' scores were grouped together, with group becoming delayed testing time. The researchers found that both frequency and involvement load coincided, in nearly equal measure, with vocabulary gains and had moderate to large effect sizes. Over time, however, only task three resulted in significant gains, with frequency of occurrence not significant. Task three, which Eckerth and Tavakoli refer to as the "input-output task" (p. 240), involved sentence writing using the targeted words. This task better facilitated active vocabulary knowledge. Arguably, a multiple regression analysis may have better served the researchers, allowing them to control for specific variables.

I located two studies, Gohar et al. (2018) and Khoshsima and Eskandari (2017), that compared the predictive power of the involvement load hypothesis to Nation and Webb's (2011) technique feature analysis, with both suggesting that the technique feature analysis may provide a better measure of comparison. Khoshsima and Eskandari (2017) undertook a single study of four groups, assigning each task two indices: one based on the involvement load hypothesis and another based on the technique feature analysis. The researchers found that cloze exercises had higher gains than the rewriting of sentences, which the researchers interpreted as partial support for the involvement load hypothesis. In contrast, indices from the technique feature analysis better predicted posttest outcomes. Gohar (2018) undertook a hierarchical analysis, finding that

regardless of variable order, technique feature analysis accounted for a greater proportion of difference.

In summary, few studies discussed so far examined frequency in terms of dispersion, few investigated the acquisition of lexical sequences through reading and writing, and few studies included secondary school students among their participants. In fact, I could locate only one study that incorporated all three, that of Snoder (2017). Snoder's investigation of Swedish bilingual learners compared intentional with incidental learning and included the effect of dispersion and the involvement load as independent variables. The lexical sequences chosen for Snoder's study were taken from a work that does not include idioms, *The Oxford Collocations Dictionary for Students of English* (2009). The target sequences were chosen because of their absence of literal translations in Swedish. The expanding dispersion treatments are described by Snoder as two separate, graduated intervals occurring over either 16 days or four days. In counter to most other involvement load hypothesis studies, Snoder discovered that higher involvement loads led to fewer gains. Although Snoder found greater gains for dispersion over massed exposure, Snoder uncovered no significant difference. However, the longest treatment period in Snoder's study was two weeks, which may not be a suitably long period for examining the effect of dispersion. Further, most research does not dispute the superiority of intentional learning over incidental learning; the argument advanced by most vocabulary scholars promoting extensive reading for vocabulary acquisition instead finds its value in encouraging long term entrenchment and efficiency.

Summary of Vocabulary Studies

An examination of the literature investigating the promise of reading and writing to facilitate vocabulary acquisition supports this potential, especially when word exposure through

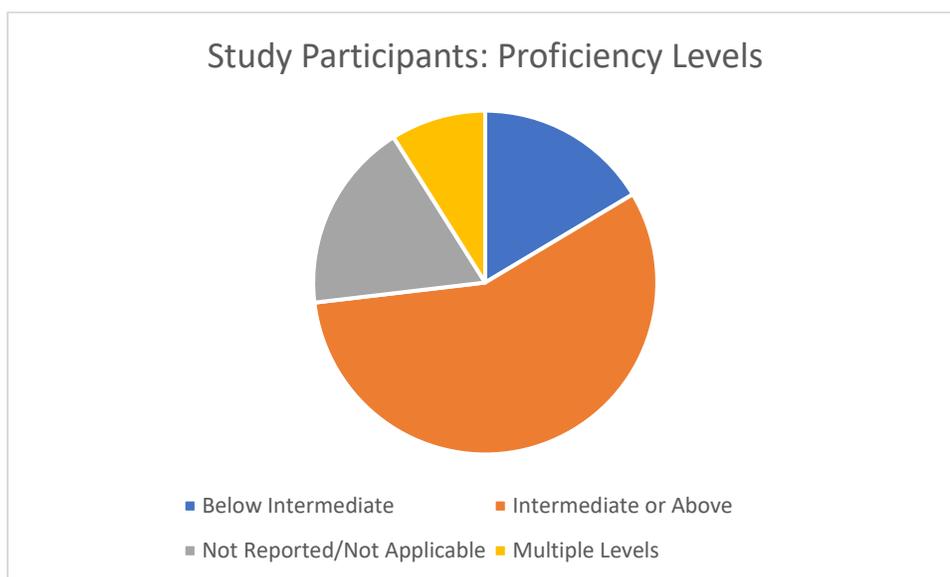
reading occurs with frequency and dispersion. Research also suggests that the glossing of vocabulary during reading may ease this process. However, only a few studies examined bilingual learners in English speaking contexts, and few included secondary students among their participants. Additionally, studies of lexical sequences were less common than those of individual word acquisition, and those examining exposure through reading followed by writing to support entrenchment were also difficult to locate. Although research has established that genres are characterized in large part by the vocabulary incorporated within (Biber, 2009; Coxhead, 2000; Hyland, 2011), there was little attention paid to genre within these studies. In other words, participants asked to complete writing activities were sometimes asked to respond in a genre different from the one they were exposed to. Finally, glossing studies that incorporate dictionaries do not describe the nature of definitions used, despite research showing that dictionaries often incorporate vocabulary more difficult than the head word being defined (Kelley et al., 2010; Restrepo Ramos, 2015).

The methods undertaken in the studies I examined may also highlight areas of concern. Of 73 studies where delayed testing would be relevant, in 30 such studies, delayed testing did not take place. Among the remaining studies in which delayed testing was reported, 21 defined delayed testing as testing that occurred one-two weeks after immediate testing, yet all studies measuring attrition found that it occurs even within this short period. Eighty-six percent of empirical studies are quantitative studies rarely reporting effect sizes. Restrepo Ramos (2015) calls for more studies examining the acquisition of lexical sequences, rather than single words, through reading, a suggestion that my literature review supports: Of the 97 studies I examined, 83 defined vocabulary as single words. Furthermore, as earlier noted, a full 72% of the 96 studies I examined took place in university contexts. While tertiary studies such as these

certainly provide valuable insights into vocabulary learning through reading and writing, the dearth of secondary studies may be concerning, given that students attending university have often passed admission examinations such as TOEFL⁹ and have therefore demonstrated a level of proficiency that assumes students have the tools to successfully complete university studies. This may explain why, in more than half of the studies I examined, the proficiency levels of participants are intermediate or above, as shown below in Figure 4.

Figure 4

Literature Review: Participant Proficiency Level



Conclusion

My review of literature investigating vocabulary acquisition through reading and writing highlights many unanswered questions about how learner variables may affect the process. Among them are the role of gender. To begin, gender has been shown to correlate to bilingual learner perceptions of language learning (Fatemi & Asghari, 2012) and to impact educational outcomes as a factor of self-regulation of behavior and subject matter (Weis et al., 2013).

⁹ Test of English as a Foreign Language

Nevertheless, despite half of the studies in my review having balanced gender participation, only two studies measured the impact of gender. Both are development studies describing changes in reading and writing development (Ahmed et al. 2014) or vocabulary development (Schmitt et al. 2004). Eighteen percent of studies I examined report having participants with multiple first languages, yet none of these studies investigated participants' first languages despite research that has reported large differences for the effect of first languages on aspects of language learning, as reported in Chapter 1.

This literature review supported the need for an intervention study that measured the impact of L1 and gender on interventions of lexical sequence acquisition through reading and writing. The fully integrated mixed methods intervention study intended to provide important insight into how secondary bilingual learners developed lexical sequences from traces to entrenchment—that is, the traces of lexical sequence recognition gained through initial exposure during reading with purposeful writing toward more entrenched word knowledge—while examining factors that impact both intervention results and language learning experiences generally. The literature reviewed supported a study with interventions taking place over an eight-week period and delayed testing taking place after a time period longer than two weeks. The addition of in-depth interviews conducted after the eight-week intervention provided a well-rounded understanding of factors that may help or hinder English language acquisition among bilingual learners.

3: METHODOLOGY

Teddlie and Tashakkori (2012) emphasize the importance of conducting research through appropriate methodology guided by an appropriate epistemological stance. The following chapter discusses how the chosen methodology assisted me in uncovering answers to my research questions as they were found to relate to my participants. After describing researcher subjectivities, I begin with a discussion of mixed methods methodology, connecting it to my questions and study purpose. I then discuss pragmatism as the epistemological stance that guided my research and provide a rationale for it. Next, I provide a detailed description of the study design and method, including a portrait of the study participants of both strands. Following a discussion of the unique challenges posed by research setting choices I made in 2021 (due to a global pandemic) and how I addressed them during the planning and implementation of the study, I provide specificity in the intervention centering my research.

Researcher Subjectivities

My positionality is that of a straight, cisgender, White, middle-aged woman who grew up in a working-class family on the Ohio River. I grew up with no awareness that I possessed privilege, initially learning about it while living in the Middle East at the age of 23. From the moment I arrived in Kuwait, I was rushed to the front of long lines for privileged treatment at government offices because I was a White American woman. However, I was also prohibited from practices reserved only for men: I could not attend sporting events, I was asked to refrain from pumping my own gas, and I was required to stand in segregated lines to buy freshly baked bread. My gendered treatment was framed as being protective and respectful in nature, however, and I envisioned these experiences as stories from my travels that I would later tell with fondness and humor.

Soon, I began to meet women of color whose lives and experiences could never be described as “respectful” or “protective.” While doing volunteer work with a friend at an embassy, I met numerous women who had been raped or beaten by their employers, women who had been locked inside their homes, and women who—despite living roughly 5,000 miles from home—were not allowed to write letters to their families or to make phone calls to their children. I met women who worked to exhaustion for years without ever having had a day off. These women had run away to the Filipino embassy, sometimes jumping from 2nd or 3rd floor windows to escape.

As a woman, I have been subjected to the same sexism and limits that all women of my generation have experienced (e.g., workplace discrimination and sexual harassment; Lease et al., 2020; Stockdale et al., 2020), both in the United States and abroad. At an airport in Tehran, I was once yelled at for having a few threads of hair exposed from under my scarf (while listening to the angered rant of the airport security guard, I noticed several Iranian women walking past me with their scarves hanging down to reveal tufts of thick hair). My life and physical safety were never at stake, however; it was simply an inconvenience that hurt my feelings. My ever-evolving understanding of my privilege and responsibilities as a White woman has led me to a heartfelt belief that activism and antiracism must be a way of life, not simply an academic philosophy or an event. In a very real sense, the physical, emotional, and economic wellbeing of many others is perpetually harmed through systemic racism (Feagin, 2004).

But educational research requires consideration of one’s subjectivities beyond simply stating them (Angrosion & Mays de Perez, 2000; Duff, 2008; Dunbar et al., 2001; Fine et al., 2000; Holstein & Gubrium, 2000; Roulston, 2013; Schwandt, 2000). As human beings, researchers cannot always be cognizant of their own subjectivities. We may not have explicit

knowledge of how our subjectivities may warp data analysis or how the research process itself may violate Kant's categorical imperative (Limes-Taylor Henderson & Esposito, 2017). I try to remain keenly aware that my personal biases are implicitly invisible to me, and I acknowledge the hard work required to surface and challenge them in all aspects of my life.

Limes-Taylor Henderson and Esposito (2017) describe a fundamental ethical problem that institutional research is necessarily fraught with. Regardless of our intentions to destroy Lorde's (1984) proverbial master's house using his tools, in reality research legitimizes "the institution that builds and reinscribes the body of knowledge necessary to justify the decimation of some peoples, and the subjugation of others ... White supremacist colonialist concepts and ways of understanding our world and each other" (Limes-Taylor Henderson & Esposito, 2017, p. 19). With this caution in mind, I tried to adopt an ethic of humility during my research (Grenberg, 2005). I attempted to mitigate the effects of my own subjectivities and to avoid exploitation of my participants through active intent. During interviews, I tried hard to avoid misunderstandings by asking for clarification and repeating what I heard back to them. I paid each participant \$50.00 upon completion of the study, and I paid interview participants an additional \$100.00 to demonstrate that I respected and valued their time. I also offered my participants the opportunity to collaborate with me through the process of member checking, although they declined to do so.

As cross-cultural in nature, I felt that the interview process required special attention to building rapport with my participants through self-disclosure (Briggs, 2000; Ryan, 2000) and by expressing genuine care for their well-being. For instance, I made sure my interviewee who revealed that she is queer has a safe space and someone to protect her, both physically and emotionally, and when two interviewees shared stories of how difficulty with their families lead

them to drug abuse, I paused my audio recorders to tell them personal stories from my own life. When participants told me they were bullied because of their accents, I shared how I felt when I was ridiculed for making errors while practicing conversational Persian in Iran. When one interviewee shared an emotional narrative about the struggles in his past, we cried together. I believe these actions assisted me in my commitment to lowering power dynamics between us *such that I could*, and I hold hope that the sharing of my research findings will contribute to improving the educational experiences of students like them in some way.

Watching my children living their lives with an energized understanding of the urgent need for social justice gives me hope that someday, things may be different for those who are relentlessly marginalized. *Roe v. Wade* was recently overturned at the time of this writing, yet my two transgender children attend pride rallies and wear pendants symbolizing the pride they take in their identities. They challenge my expectations and assumptions and regularly teach me new ways to think about gender. My other children fervently defend what is right—*even when it is hard*—time and time again, without regard for how it will affect their careers or their relationships with those around them. They seem to seek fulfillment in their lives by imagining new ways to help others in need. They continue to teach me how to improve myself as a human being, and they are the inspiration for everything I do, including this research project.

Research Design

The purpose of my study was to examine factors affecting language learning and vocabulary acquisition through reading and writing among secondary bilingual learner students. I sought answers to the following research questions:

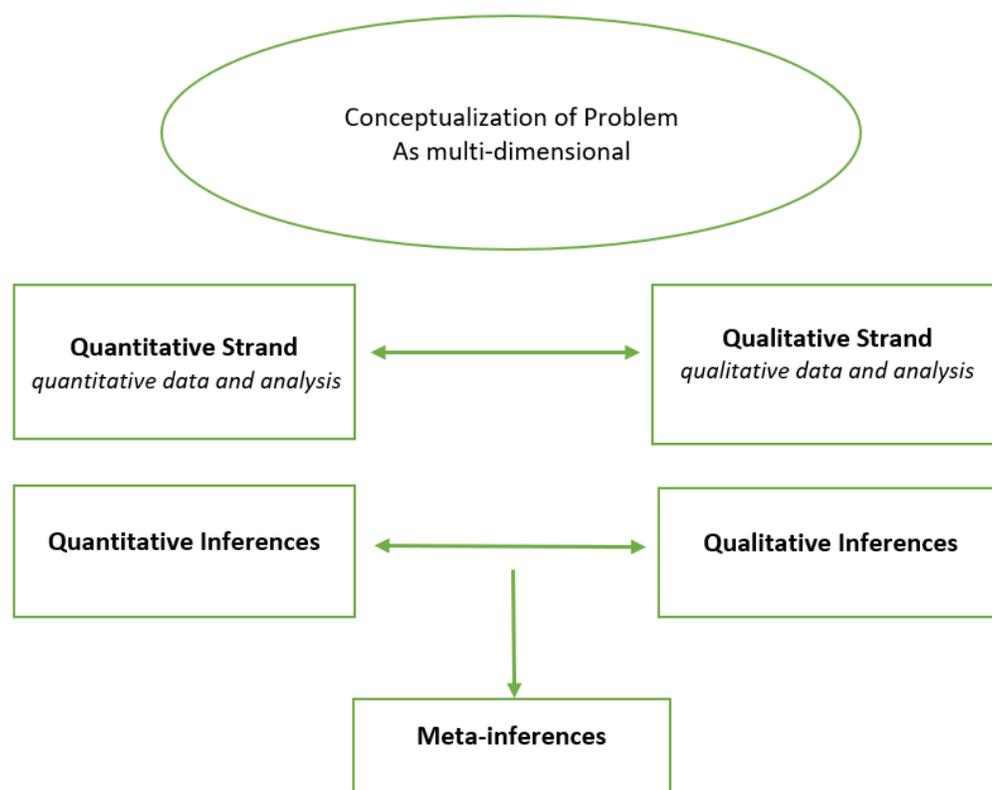
1. (QUAN) Over a 4-week period, can exposure to targeted lexical sequences through reading and writing of academic texts lead to acquisition among secondary school bilingual learners?
 - a. Can exposure through reading lead to targeted lexical sequence acquisition?
 - b. Can exposure through reading followed by intentional use of the sequences lead to acquisition?
 - c. Is one intervention more effective than the other?
 - d. Are any initial gains retained in delayed vocabulary testing?
 - e. Are there any interaction effects due to participants' first language (L1) or gender?
2. (QUAL) What other factors affect intervention results?
3. (QUAL) What factors impact the participants' educational experiences?
4. (MIXED METHODS) What meta-inferences can be drawn from the quantitative and qualitative strands of the study about factors affecting bilingual learners' learning?

To explore the first research question, I conducted a factorial ANOVA with repeated measures using the test scores from a modified version of the Vocabulary Knowledge Scale (VKS) assessment of the targeted vocabulary items as the measure. The repeated measures calculated the impact of intervention (reading alone or reading coupled with writing) and time tested (pretest, posttest, or delayed posttest) as within-group factors and the impact of L1 and gender as between-group factors among 49 participants who underwent both interventions in turn. The second research question required an examination of qualitative data. Through in-depth interviews, I gained deeper understandings of factors that affect my participants' language learning experiences, something that quantitative data alone could not have provided. An

examination of written responses to readings provided further avenues of investigation of quantitative data as described in detail below. Together, the qualitative and quantitative data allowed me to draw meta-inferences about the factors affecting English learning, or the answer to Research Question 4. (See Fig. 5.)

Figure 5

Mixed Methods Fully Integrated Design Model (Riazi, 2017)



Defining Mixed Methods Methodology

Not to be confused with multi-methods—the mixing of research methods *within* quantitative or qualitative research (APA, 2020; Campbell & Fiske, 1959; Morse, 2003; Schoonenboom & Johnson, 2017)—mixed methods is a methodological approach that incorporates both qualitative and quantitative methodologies. Mixed methods has been called the “third tradition” (DeCuir-Gunby & Schutz, 2017, p. 2), the “third wave” (Johnson &

Onwuegbuzie, 2004, p. 14), the “third methodological movement” (Teddlie & Johnson, 2009a, p. 4), the “third research paradigm” (Creswell & Plano Clark, 2018, Chapter 1), and the “third research approach” (Dörnyei, 2018, p. 20). Regardless of what proceeds it, the inclusion of *third* refers to its having arisen in the late 1980s as a result of a paradigm war characterized by dogmatic, dichotomous views of qualitative and quantitative research (Dörnyei, 2018; Riazi, 2017; Teddlie & Johnson, 2009b). In fact, according to Dörnyei (2018), “the terms *qualitative* and *quantitative* were originally introduced as part of, or rather for the purpose of, an ideological confrontation” (p. 26). Thus, the narrative of a paradigm war seems to position mixed methods as having arisen as a peace negotiator.

A Rationale for Mixed Methods Methodology

Despite the paradigm war narrative, some methodologists believe that mixed method research has been conducted in the social sciences for nearly a century or longer (e.g., Burke et al., 2007; Dörnyei, 2018; McKim, 2017; Teddlie & Tashakkori, 2009), or that strict boundaries of paradigms have never truly existed (Mallet et al., 2013; Teddlie & Johnson, 2009a). Bergman (2008) questions whether any scholar truly espoused the incompatibility thesis, arguing instead that given the heterogeneity of social science research, a list of diametrically opposed elements as belonging to one or the other paradigm would be impractical. For instance, Onwuegbuzie and Hitchcock (2015) claim, “we do not know what mono-method research looks like” (p. 276). Biddle and Shaftt (2015) characterize the increase in mixed methods publications in the 21st century as “near-exponential growth” (p. 321) and present a growth chart demonstrating incremental increases in its popularity over the past two decades, while Dörnyei (2018) calls it the new “zeitgeist” (p. 42). In short, there seems to be increasing acceptance of mixed methods studies overall (Denzin, 2010; Riazi, 2017; Suter, 2012), and perhaps for good reason. As Ladner

(2019) contends, sometimes choosing between the benefits of quantitative studies and qualitative studies is “a terrible trade-off” (Chapter 1).

Although Hesse-Biber (2010) believe that a single researcher may lack the expertise to undertake mixed methods, others imply a responsibility for learning all three. For example, Teddlie and Tashakkori (2012) have proposed that the student researcher become “a connoisseur of methods” (p. 777), whereas Denzin and Lincoln (2000) stress that “a researcher-as-methodological-*bricoleur*” must know and gain experience in multiple techniques of data collection and data analysis (p. 6). Some scholars have suggested that because quantitative and qualitative methods may well complement one another, mixed methods methodology may be essential to good research. For example, Dörnyei (2018) expresses a sentiment similar to Denzin and Lincoln when he states, “I strongly believe that the ‘good enough researcher’ needs to master some knowledge of both quantitative and qualitative research, as well as ways of combining them” (p. 10). Dörnyei also notes that adherents of mixed methods sometimes propose that mono-methodology threatens the advancement of the social sciences. Within literacy studies, the view of mixed methods studies as important research can be seen in Goldberg et al.’s (2005) assertion that “mixed methods are essential to advancing our field” (p. 42). Mixed methods research also reaches a large audience (Dörnyei, 2018) and allows for the expansion of a study’s results (Schutz et al., 2004).

Mixed Methods as Fully Integrated

I conducted what Bogdan and Biklen (2007) referred to as “lone ranger research” (p. 75), which required great attention to detail, and above all, the use of a well-considered mixed methods model to assist me in providing the rigor required for meaningful and generalizable results. Tashakkori and Cresswell (2007) as well as Creswell and Plano Clark (2018)

differentiate between mixed methods as a *method* of data collection and analysis and as a *methodology* that integrates two approaches. I undertook mixed methods as a fully integrated methodology.

Mallett et al. (2013) emphasize the importance of an iterative approach to mixed methods analysis within which each approach informs the other. As Schutz et al. (2004) state, “Simply adding open-ended questions to a larger quantitative study will probably not meet the guidelines for conducting useful inquiry” (p. 275); Dörnyei (2018), Curry et al. (2009), and Onwuegbuzie and Mallette (2011) believe that mixed methods allows the strengths of qualitative and quantitative studies to be combined such that, in the words of Dörnyei (2018), “the resulting mixture or combination is likely to result in complementary strengths and non-overlapping weaknesses” (p. 167). Thus, scholars are often steadfast in demanding high standards for the “mixed” part of mixed methods, and indeed, mixed methodologies very often define and describe sophisticated approaches to mixed methods (e.g., Hitchcock & Onwuegbuzie, 2019; Onwuegbuzie, & Hitchcock, 2015; Uprichard & Dawney, 2019). Teddlie and Tashakkori (2009) describe fully integrated mixed methods as one that involves integration of both qualitative and quantitative methods at nearly every stage of research. As Riazi (2017) explains, principled, innovative research requires this. Using Teddlie and Tashakkori’s (2007) typology, my study can be characterized as multi-stranded. As such, the qualitative collection and analysis of my study took place alongside the collection and analysis of quantitative data, but not necessarily as a result of its findings.

Pragmatism as a Guiding Paradigm

My research was guided by a pragmatic paradigm. Often associated with Charles Sanders Peirce (1933), William James (1907/1995), and John Dewey (1916/1967), pragmatism is best

thought of as a paradigm package, complete with an epistemology, ontology, axiology, and methodology (Dillon et al., 2000; Kivunja & Kuyini, 2017; Kaushik & Walsh, 2019). Whether overtly stated or not, paradigms inform methodological choices all researchers make.

Consequently, some (e.g., Kivunja & Kuyini, 2017; Koro-Ljungberg et. al., 2009) have argued that researchers must identify and justify paradigmatic choices.

Pragmatic inquiry frames research decisions on the meaningfulness of goals and the appropriateness of method while standing in opposition to actions that would “limit freedom of inquiry and possibilities for social justice” (Weaver, 2018). Because, as Morgan (2014) argues, “without preconceptions, inquiry is impossible” (p. 66), there is an inherence that pragmatic inquiry is *abductive* in nature (Feilzer, 2009; Morgan, 2020; Sullivan, 2015). Abductive reasoning simply means that our expectations are driven by previous experiences. It has been described as “the cornerstone of scientific methodology,” and it is also what makes language comprehension possible (Douven, 2017, “The ubiquity of abduction”). People hold assumptions they have abductively acquired until “the production of new insights through the creation of a hypothesis that offers an explanation for one’s observation” come into being (Morgan, 2020, p. 67). My own study relies on inductive and deductive reasoning through quantitative and qualitative analyses respectively, but it began and ended with abductive reasoning: Abductive reasoning informed my research choices—the decision to conduct a mixed methods study that is fully integrative—and it informed the questions I sought answers to, which are grounded in my teaching experiences. It also assisted me in constructing an understanding of the results.

Several scholars find promise in pragmatism as a guiding principle for educational research (Cresswell & Plano Clark, 2018; Dillon et al., 2000; Koopman & Garside, 2019; Onwuegbuzi & Mallette, 2011; Teddlie & Tashakkori, 2009) as well as literacy research in

particular (e.g., Dillon et al., 2000; Dillon & O'Brien, 2019; Dressman & McCarthy, 2011). However, as Dillon et al., (2000) have discussed, literacy researchers rarely take a pragmatist stance, perhaps because it is so often misrepresented and misunderstood (Dillon & O'Brien, 2019). Perhaps the most common misassumption about pragmatism is that it simply means *what works*, (Dillon & O'Brien, 2019; Fox & Alldred, 2018; Hall, 2013; Morgan, 2014). Such an oversimplification, however, overlooks the criticality that is inherently part of pragmatist thought (Deans, 2009; Dillon & O'Feilzer, 2011; Dressman & McCarthy, 2011; Jones, 2002; Kaushik & Walsh, 2019; Morgan, 2014; Sullivan, 2015).

Writing in 1997, Bereiter et al. described concerns about the placement of scientific knowledge alongside magic and pseudo-science. They view science not as the seeking of truth but as “improvement on existing knowledge” and theory (p. 331), and while they invite skepticism, the researchers emphasize that skepticism need not be anti-scientific. In educational contexts, they argue, science requires commitment to advance knowledge and to propose ideas that may be tested empirically—which requires making one’s position vulnerable.

Bereiter et al.’s (1997) cautions have arguably never been more relevant than today. In the current political and cultural environment, it seems particularly important for any researcher to confront the meaning and limits of what counts as *truth*. A purely objective reality, free from individual interpretation of meaning may not be possible (Bruffee, 1986). Derrida (1967/2016), for instance, proposed that writing is twice removed from consciousness, and as such, it cannot possibly represent truth, and as Schwandt (2000) explains, “we do not construct our interpretations in isolation but against a backdrop of shared understandings, practices, language, and so forth” (p. 197). Yet an underlying truth—as Luke (2018) names it—may still be, and must be, sought.

Hall (2013) contends that within pragmatic paradigm, truth is “never completely settled and continues ... as new understandings come into focus” (p. 18). Dewey’s conception of pragmatism replaces the notion of truth with what he terms “warranted assertions” (Hall, 2013; Morgan, 2014), which arise from inquiry. Inquiry is therefore a central part of Dewey’s pragmatism as well; pragmatism embraces a disposition towards action and a dependence on experience in guiding quests for answers (Koopman & Garside, 2019; Morgan, 2014; Morgan, 2020). Within pragmatism, experience is inescapably emotional, social, cultural, embodied, historical (Morgan, 2014) and is bound to language (Jones, 2002). As a mixed-methods study incorporating both qualitative and quantitative data to examine how participants’ lived experiences, first languages, and gender affect their language learning experiences, my study is both pragmatic and critical in nature.

Research Challenges During COVID-19

The COVID-19 pandemic affected all aspects of human life globally, and our world is forever and fundamentally changed as a result (Ogodo, 2022; Sharma & Borah, 2020). One global theme that has emerged from the pandemic has been the capricious nature of human behavior that the story of COVID-19 seemed to depend upon. To design a study in the depths of this pandemic felt akin to embarking on a journey without a clear path and without an understanding of where the path would lead. On the other hand, perhaps all burgeoning researchers feel this to a lesser degree. It seemed unclear, when my research was planned, when the pandemic would end or what “end” would even mean. At the time that my study was planned, two vaccines had been approved by the FDA (Center for Disease Control and Prevention, 2021). However, there was uncertainty about how long the vaccine would provide immunity from COVID-19 (Center for Disease Control and Prevention [CDC], 2021; Hopkins &

Loftus, 2020). Furthermore, conspiracy theorists had proven remarkably adept at convincing people to treat the vaccine with suspicion (Romer & Jamieson, 2020), which added an additional layer of future uncertainty. The pandemic's invasive nature meant that researchers were compelled to recognize not only how the COVID-19 pandemic would affect research design and data analysis, but additionally, how it would affect the research experience and how it unfolded. These complications and uncertainties therefore had to be considered in any research study planned after March 2020—including, of course, my own.

For guidance in understanding how the pandemic would affect my research and how I could best prepare for the challenges it might bring, I sought answers from critically reflective scholarship (e.g., Duff, 2008; Palmer, 2014; Ravitch, 2021). I learned that in the absence of preparing for specific challenges, as is true for all research, I must instead prepare for the unforeseen. That is, I gained awareness that my study would be characterized by unexpected turns. One way I prepared for such uncertainty was through the acceptance of research design as a flexible and responsive plan. Although I originally intended to conduct a case study, the qualitative strand became dependent upon interview analysis. While I intended to have many face-to-face interactions with my participants, I only met them briefly to provide them with remuneration. I imagined visiting my interviewees' homes and meeting their families, but instead, we sat in a conference room at their high school speaking through masks and following a strict schedule that others had created for me.

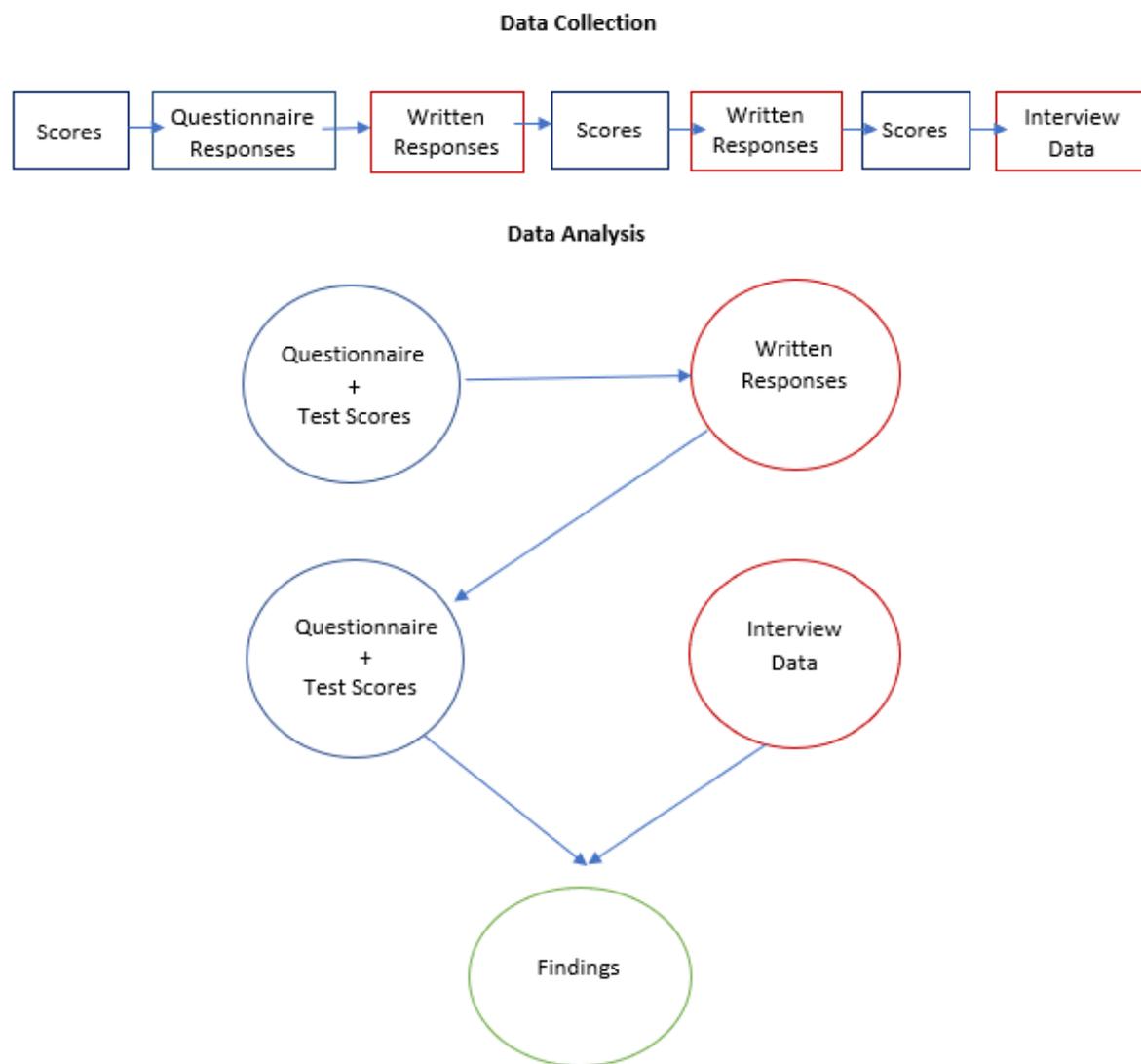
Study Procedures

Because of the nonlinearity that characterizes integrative mixed methods analysis, an explanation of how the two study strands unfolded is warranted as a starting point. Figure 6 provides a visual presentation of how the mixed methods analysis evolved. It is important to note

that although the initial quantitative analysis occurred first, the integrative nature of the study meant that examinations occurred iteratively.

Figure 6

Evolution of the Mixed Methods Study



The top of the figure demonstrates the order that data were collected, with blue indicating quantitative data and red indicating qualitative data. First, pretest scores and questionnaire data were collected for all participants. Written responses to readings were then collected for Group

B, who underwent the writing intervention. Next, the first posttest scores were collected, followed by written responses to readings collected during the writing intervention for Group A. Delayed test scores were then collected for both groups, followed at last by the collection of interview data. During the analysis stage, quantitative data were informed by the inclusion of qualitative data, which was quantified and examined as part of a second quantitative analysis and would inform my interpretation of results. Once the interview data were examined, both strands came together to inform my findings and discussion.

Quantitative Strand

Because the quantitative strand of the data analysis occurred first, I discuss the participants, instruments and materials, study procedures, and study validity for the quantitative strand first. This is followed by a discussion of the participants, instruments and materials, study procedures, and study validity for the qualitative phase. Finally, I briefly discuss how the two strands informed one another during the mixed methods analysis phase.

Participants. The participants for my study were high school bilingual learner students designated as English language learners by the State of Georgia. As noted in Chapter 1, Limited English proficient was defined by the English Language Acquisition, Language Enhancement, and Academic Achievement Act, a grant program within Title III of the Elementary and Secondary Education Act (2001). In 2015, the term was updated to English language learner (ELL) when the law was amended by the Every Student Succeeds Act (Darling-Hammond, et al., 2016). Because these terms advance a harmful deficit view of these learners, I refer to these learners as bilingual learners, a term that acknowledges their proficiency in their L1s.

Sampling Size and Sampling Scheme. To ensure that the quantitative sample size took into account elements that Cohen (1988) considered important for quantitative research—that is,

power, alpha level, and effect size—I used G*Power for calculating the sample size by determining these elements *a priori*. The calculated estimate—which arose as a function of the number of independent variables (4: time, gender, L1, and group), of the alpha level I set (.05), the power I set at .8 and a moderate effect size of $R^2 = .25$ —was 36. Because attrition was a concern about bilingual learner student participants (Dörnyei, 2018; Duff, 2008), to help account for the effects of participant attrition, I aimed for a participant sample size of 50: 37% higher than the number G*Power calculated to be sufficient. The only inclusion criteria for the study were that participants were bilingual learner high school students in metro Atlanta who scored an 80% or below on a modified version of the Vocabulary Knowledge Scale, which measured depth of knowledge of the targeted lexical sequences.

Recruitment Process. Once the IRB granted permission for my study in September of 2021, I attempted to recruit participants. The scope of my initial recruitment efforts included communities and schools in a Metropolitan Statistical Area identified by the United States Census Bureau (2012), often used for statistical purposes (Nussle, 2008). First, I reached out through email using addresses found online for ESOL instructors within the largest school districts located throughout this region. To find these emails, I searched public directories for each school district. I also conducted an online search for churches offering ESOL classes in Metro Atlanta. I called the first five churches that appeared in a Google search using the terms ESOL classes, church, and Metro Atlanta, asking them to share my contact information with interested individuals. Nevertheless, despite my efforts, from September 2021 to December 2021, I found little interest. When one ESOL teacher expressed concern that I was cold-calling teachers without going through her county’s district office, I contacted IRB offices for each school district having high schools on my contact list. All IRB offices requested that I seek

formal permission from them before contact ESOL teachers employed by their school districts. However, the IRB processes would have delayed my study for an additional six months, and in the case of one school district, it would require me to reduce the length of my study by one month.

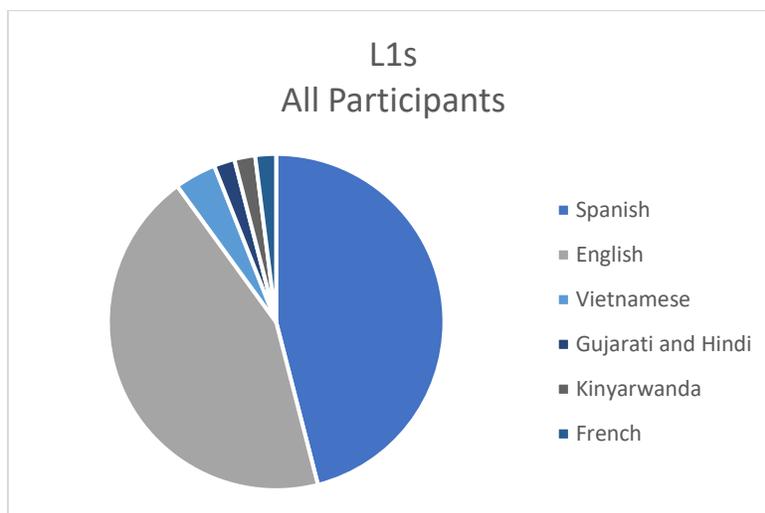
To improve my chances of recruiting participants, in December 2021, I modified my IRB application and proposed study by changing the participant target number from 100 to 50 and increasing the remuneration from \$25.00 or an hour of tutoring to \$50.00. My original plan for the study was a two-group comparison study. I changed the study to a repeated measures design with one group of 50 participants who would undergo both interventions in turn. These changes required several amendments to my IRB application as well as new consent and assent forms, new translations, and new approvals for each.

In January 2022, I reached out to the principal at a large suburban high school known for having a large bilingual learner population (U.S. Department of Education, 2017), and the absence of a district IRB office meant that I would not have to wait several months to secure all needed permissions or to change the length of my study. The principal of the high school seemed eager to help and arranged for me to meet with three ESOL teachers to allow me to explain my study to the teachers personally. The ESOL teachers happily agreed to ask their students and even offered me advice on ways to improve my study. Interested students were sent a link to complete the pretest through Google Forms, which measured the degree of knowledge they had of eight lexical sequences chosen for the study. Interested students were asked to include both a pseudonym and their real names on the form. When I received new pretest results, they were immediately printed out, and their results were deleted electronically from Google Forms to help retain confidentiality. Between early January and February of 2022 and with the encouragement

of their ESOL teachers, 58 participants took the pretest. To curb the potential impact of a ceiling effect found in some of the literature I reviewed (e.g., Keyvanshekouh, 2012), three participants were excluded because they had scores of higher than 80% on the vocabulary pretest. The first 50 who met the pretest criteria were invited to take part in the quantitative strand of the study.

Demographics. During Week 3, it became clear that one participant was unable to complete the weekly tasks within the given time frame, and consequently, the participant was dropped from the study, leaving the sample size reduced to 49. These 49 participants included 26 females and 23 males. A questionnaire elicited demographic information that has been shown to impact schooling: L1 (Ellis, 2006; Liu & Chen, 2019; Masrai & Milton, 2015; Paquot, 2013), gender (Agustín Llach & Terrazas, 2012; Roohani & Akbarpou, 2016), age of arrival (Roessingh et al., 2005), number of years and months living in the United States (Lane et al, 2019), number of years of formal schooling (Custodio & O’Loughlin, 2020; Hos, 2016), race (Kubota, 2015; Liggett, 2014; Monzó & Rueda, 2009; Roxas & Roy, 2012), ethnicity (Fish, et al., 2019), country of origin (Castro-Olivo et al., 2015; Lee, 2006), and refugee status (Pittaway & Bartolomei, 2001; Roy, 2015; Roy, 2017). This questionnaire appears as Appendix B.

L1 was defined in the questionnaire as the language the participant feels most comfortable using and the one that they know best. Nearly half of the participants reported that they were born in the United States, so perhaps unsurprisingly, 45% reported English as their L1. Twenty-three participants indicated that they speak Spanish as their L1, two reported Vietnamese, and one each reported Gujarati, Kinyarwanda, and French as their L1. (See Fig. 7.)

Figure 7*Participant Self-Reported L1s*

Because the small group numbers of languages other than English and Spanish would lead to a violation of the assumption of equal variances for L1s (Stevens, 2007), I decided to measure L1 in terms of the impact of whether a participant identified English or non-English as their L1. Fourteen students chose not to report their AOA, and another 15 stated that they were born in the United States. The average AOA for the remaining 13 students was 9.0. Of participants who reported AOA, all who reported English as their L1 also reported being born in the United States.

As culturally constructed, ways of defining race and ethnicity are dependent on geographics and culture; lists of “races” and “ethnic groups” from place to place differ greatly from country to country (Flanagin et al., 2021), yet Gaias et al. (2020) argue that reporting race and ethnicity in intervention studies is critical to avoid underrepresentation of some groups. Unfortunately, however, the constraints that I faced in participant recruitment limited my ability to actively seek balanced representation of race and ethnicity. I collected data on race and ethnicity of participants following recommendations by some researchers (e.g., Ross et al., 2020)

to allow participants to define their own race and ethnicity. Therefore, participants were asked to report their race and ethnicity as an open-ended response. In doing so, they often used terms that may reflect a different understanding of race and ethnicity than is commonly reflected in preprepared lists. Participants reported a total of nine ethnicities, with Hispanic comprising the greatest majority by far. However, of the 46 participants who reported race, there was no clear majority choice. (See Fig. 8 and Fig. 9.) One participant reported having refugee status; no participants had experienced interrupted schooling.

Figure 8

Study Participants: Ethnicity as Described by Participants

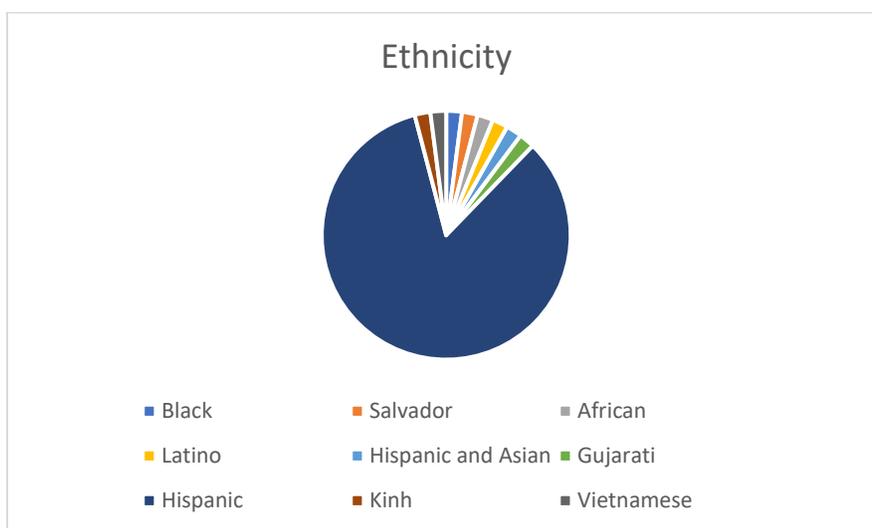
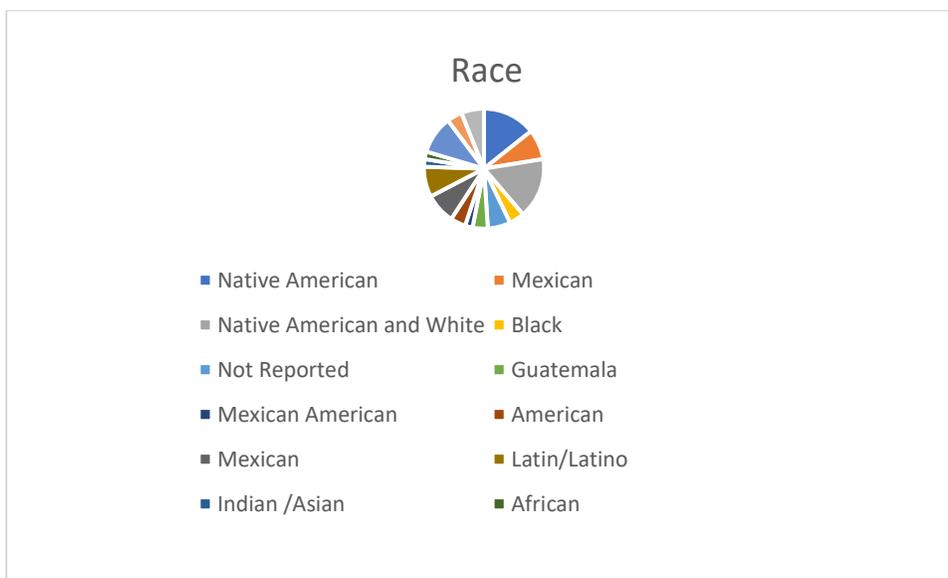


Figure 9

Study Participants: Race as Described by Participants



Instruments and Materials

The quantitative strand of the study required eight lexical sequence targets, a vocabulary assessment, a questionnaire, and 16 reading passages. The passages, lexical sequence targets, and instrument choices are described below. The questionnaire appears in Appendix B, and the vocabulary assessment appears in Appendix C.

Lexical Sequence Targets

To locate lexical sequences for my study, I examined several lists identified in authentic texts by previous researchers as common to academic writing. I then created a list of lexical sequences that appeared in at least two published lists. I looked for sequences characterized by some degree of opacity in meaning because I wanted to be confident that a learner who acquired the sequence had acquired it as a prefabricated unit. I also looked for sequences with one- or two-word synonyms so that students encountering the phrases during reading could access meaning through interlinear glosses without excessive interruption of reading (Lofgren, 2022).

Because a Google Scholar exact phrase search counts the number of independent articles using the phrase at least once, Google Scholar’s exact phrase search was used to measure how often the lexical phrases were found within peer-reviewed, published journal articles. I removed several targets that were found to have low frequency of occurrence relative to the other potential lexical sequence choices. For example, I removed the lexical sequence “has been identified as” because the phrase was only found in 716,000 academic works. I added one lexical sequence not found in any published lists: “referred to as.” At 4,260,000, the phrase is less common than others on my list. However, the phrase itself appeared in many articles publishing the lists (e.g., Biber et al., 2004; Hyland, 2008; Yland & Jiang, 2018), and because of its opacity in meaning despite its easy definition, I found it to be a suitable choice. The final list of eight target items appears with set numbers in Table 1 along with Google Scholar phrase search result count, simple English gloss, and corpus study sources.

Table 1

Lexical Sequence Targets

Lexical Sequence	Set	Publications	English Gloss	Corpus Studies
a number of	B	7,980,000	many	Hyland, 2008; Shirazizadeh & Amirfazlian, 2021
referred to as	B	4,260,000	is	none
a wide range of	B	6,130,000	many different	Hyland, 2008; Ozturk & Durmusoglu, 2016; Pan et al., 2016; Shirazizadeh & Amirfazlian, 2021
as well as	B	6,430,000	also	Biber et al., 2004; Cortes, 2006; Biber et al., 2004; Hyland, 2008; Jiang & Nekrasova, 2007; Ozturk & Durmusoglu, 2016
in addition to	A	7,500,000	also	Ozturk & Durmusoglu, 2016; Shirazizadeh & Amirfazlian, 2021

Lexical Sequence	Set	Publications	English Gloss	Corpus Studies
the extent to which	A	4,890,000	how much	Cortes, 2006; Martinez & Schmitt, 2012; Ozturk & Durmusoglu, 2016; Shirazizadeh & Amirfazlian, 2021
as a result	A	5,520,000	so	Appel & Murray, 2020; Byrd & Coxhead, 2010; Cortes, 2006; Biber et al., 2004; Hyland, 2008; Jiang & Nekrasova, 2007; Ozturk & Durmusoglu, 2016
in terms of	A	6,510,000	about	Gilmore & Millar, 2018; Hsu, 2014; Hyland, 2008; Pan e al., Ozturk & Durmusoglu, 2016; Shirazizadeh & Amirfazlian, 2021; Simpson-Vlach & Elis, 2010

Reading Passages

The texts that I created for my participants are abridged and adapted Wikipedia articles that include topics I felt that participants would find engaging. The texts also include articles about social justice-oriented topics, articles about social movements and philanthropic organizations, and a few topics I believed teenagers would find relevant and relatable. I define relatable texts as those that connect to students' cultural backgrounds and interests (Guthrie et al., 2012; McDevitt, 2021). I also incorporated critical texts, defined in accordance with Poulus and Exley (2018) as those used to promote transformation of power in sociocultural and political spheres of influence. I agree with Mertens (2007) and Mertens et al.'s (2010) way of defining critical text, also expressed by Teddlie and Tashakkori (2012) as "tools that are used in the service of value systems that are foremost to (researcher) perspectives" (p. 781). As a former ESOL teacher, I have discovered that students find critical texts relevant and therefore interesting, an important consideration when fostering good reading habits (Tinker Sachs, 2001).

To create the texts, I brainstormed a list of 18 topics I felt would have relevance for the population of students I sought to recruit. Next, I located Wikipedia articles and read them. One article at a time, I copied sections of the Wikipedia article defining the concept, relaying important concepts associated with the topic and discussing critical aspects of the topic. Next, I simplified the text using the most common English words whenever possible and shortening sentences to independent clauses. As I completed this part of the task, I continually checked the readability level, aiming for a readability below 5.5 on the Fleish Kincaid readability scale and a text length of between 100-150 words. At times, readings ranging from 5.5-6.0 were the lowest reading level possible to remain true to the text's meaning, and at other times, longer lengths were required to relay all important concepts or to create a coherent paragraph. I then placed the texts in the order they now appear. With targeted lexical sequences having already been arranged in sets of two, I incorporated lexical sequences where they best fit semantically. Because the lexical sequences I chose have simple meanings, this part did not require many changes to the text. These readings appear in Appendix D, and the links to original Wikipedia articles appear in Appendix E. The dispersion schedule appears as Table 2 along with the word counts and reading levels of the passages.

Table 2

Passage Readability and Schedule of Lexical Sequence Dispersion

Text Title	Week	Session	Set	Word Count	Fleish Kincaid
1. Gen Z	1	1A	1	111	3.4
2. Implicit Ideas	1	1B	2	82	4.6
3. Social Movement	2	2A	1	104	4.5
4. Youth Activism	2	2B	2	85	5.3
5. Refugees	3	3A	1	86	5.8
6. Immigration	3	3B	2	114	5.5
7. Climate Change	4	4A	1	108	5.5
8. Universal Declaration of Human Rights	4	4B	2	104	3.5

Text Title	Week	Session	Set	Word Count	Fleish Kincaid
9. Dragon Con	5	5A	3	117	4.5
10. What is an ELL?	5	5B	4	129	5.0
11. Social Justice	6	6A	3	84	5.2
12. Fundamental Attribution Error	6	6B	4	97	5.9
13. Black Lives Matter	7	7A	3	114	5.9
14. Amnesty International	7	7B	4	131	3.3
15. Greta Thunberg	8	8A	3	113	3.9
16. Doctors Without Borders	8	8B	4	99	4.9

Vocabulary Assessment

Participants took a vocabulary assessment four times during the study, with each test including the lexical sequences relevant for that specific test. The test is based on the Vocabulary Knowledge Scale (VKS) test developed by Wesche and Paribakht (1996). Often adopted for vocabulary acquisition intervention studies (Kremmel, 2020), the assessment allows participants to report their knowledge of specific vocabulary items on a scale of 1-5 and is intended to measure small gains in vocabulary knowledge of bilingual learners along the receptive and productive continuum. The modified test, which appears in Appendix C, measures the breadth and depth of knowledge of all eight targeted lexical sequences.

Test Scoring. The modified tests presented participants with targeted lexical sequences, asking them to rate their knowledge for each lexical sequence and asking them to use the targeted lexical sequences in a sentence. Sequences having responses of “I have never seen this lexical sequence before” were given a score of 0. Choosing “I have seen this lexical sequence before, but I do not know what it means” earned participants a score of 1 for that sequence, whereas “I know what this lexical sequence means, but I can’t use it correctly in a sentence” was assigned a score of 2. All participants who used the lexical sequence correctly in a sentence were granted a score of 3.

Because of the subjective nature of determining whether a given sequence was used correctly by a participant, for the sake of validity, two individuals independently judged whether the sequences was correctly used: myself and an assistant who holds a BA in English. The paid assistant was trained in the process of determining whether usage of the lexical sequence had used the phrase correctly in a sentence. The assistant was instructed to only grant credit for correct usage if the answer was written in English. He was asked to focus on the semantics of the lexical sequence in context, specifically whether the meaning as intended was apparent in the writing with no consideration for usage errors, punctuation usage, or spelling. Of 348 instances of student attempts to demonstrate correct usage of a lexical sequence across all tests, only 16 responses resulted in different ratings between raters; Cohen's kappa was calculated and found to be .953, or near perfect agreement (McHugh, 2012).

Data Collection

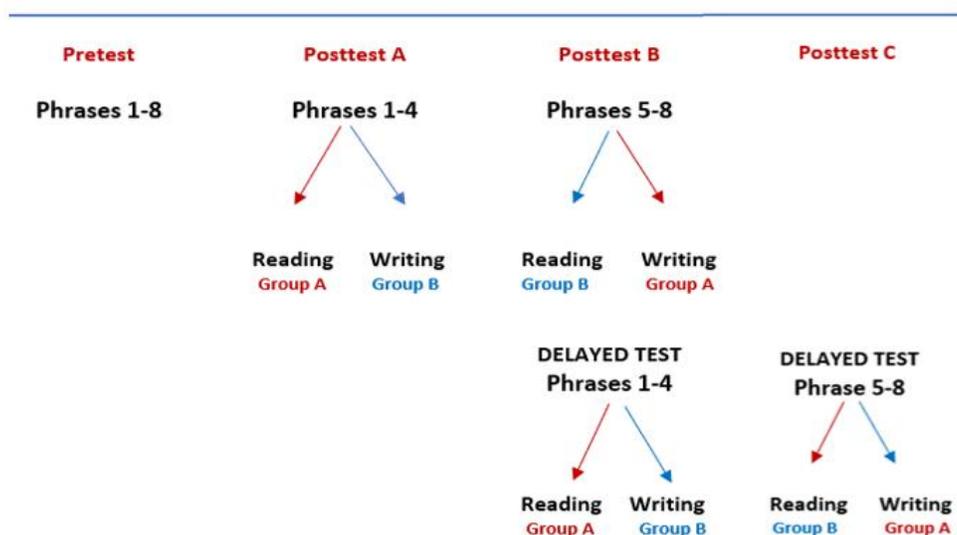
The quantitative strand of my study was a pretest, posttest design with randomization used to mitigate the effects of individual differences as extraneous factors (Torgerson & Torgerson, 2012). Along with answers to the questionnaire, vocabulary assessment scores comprised the quantitative data collected. Group A underwent the reading intervention followed by the writing intervention, whereas Group B performed the writing intervention first, followed by the reading intervention. The quantitative data collected included the questionnaire and the test scores for all four vocabulary tests. Participants were randomly assigned into one of two experimental groups (Group A and Group B) using the online app QuickCalc (Motulsky, 2021).

As shown in Figure 10, students were given the test four times over an 11-week period: a pretest before the study began, a posttest of sequences 1-4 at week 4, a posttest of sequences 5-8 coupled with a delayed posttest of sequences 1-4, and a delayed posttest of sequences 5-8 at

week 11. For lexical sequences 1-4, Group A underwent a reading intervention whereas group B underwent a writing intervention. For lexical sequences 5-8, Group A underwent the writing intervention, whereas Group B undertook the reading intervention. Each reading passage was seeded with two target vocabulary sequences. Hence, each week, participants were exposed to four targeted vocabulary sequences. In this way, all participants underwent both conditions. Both groups had delayed posttests after both interventions, as seen in Figure 10.

Figure 10

Testing of Lexical Sequences by Group



Time Frame. The pretest was emailed to interested participants throughout the months of January and February of 2022. Once 50 qualifying participants were found, they were asked to sign an assent form in their L1 or in English. Participants were also given paper copies of parental consent forms translated and approved by the university's IRB and were asked to return them. Once all paper copies of the parental consent forms and assent forms were collected from the participants at the school, the study began on March 3, 2022. Information linking the pseudonyms to participants' real names was stored on a single sheet of paper kept in a fireproof

lockbox with a combination code in the researcher's home. The testing data were collected in accordance with the schedule appearing in Table 3.

Table 3

Testing Schedule

Date Range	Test
February 1-Feb 28, 2022	Test 1
March 29-April 1, 2022	Test 2
April 26-April 29, 2022	Test 3
May 16-May 19, 2022	Test 4

Beginning March 3, 2022, all participants were sent two links each week through email, one for each of two weekly readings. I found two recommendations from the ESOL teachers at Ridgemont High School (a pseudonym) helpful, and in fact, I believe that adopting these recommendations improved the study's validity. The first recommendation was a suggestion that I use Google Forms to collect data instead of the paper booklets that I had originally planned to disseminate. My paper booklet plan relied heavily on trusting the teenage participants to faithfully complete the tasks in accordance with an 8-week schedule I would provide. The plan required participants to reliably open the booklets twice per week, for eight weeks—given only weekly reminders to do so through email. In contrast, Google Forms allowed me to create the booklet pages electronically and send them to the participants following the dispersion schedule.

Once a participant completed a task or test through a Google Form, verification was automatically sent to a Google Sheet created for the purpose of collecting it. This process allowed me to see each week which participants had completed the tasks, thereby alerting me to reach out to participants who needed an occasional nudge. The second recommendation offered to me was to create images of the readings so that participants would not simply translate the

entire page and read it in a language other than English. One English teacher in particular went to great pains to ensure that students completed their readings within the allotted time frame. She contacted me weekly to ask which students had not yet completed their tasks and reminded these students to do so with fidelity. She collected parental permission forms as well as consent forms from those interested in participating in the study. Without her assistance, I would have found it difficult to ensure that all participants completed the readings following the dispersion schedule. According to participant interviews, their teachers also often allowed time in class for them to complete the research study assignments. Occasionally, a participant completed a task outside of the week in which it was required, but in all cases, participants' exposure to targeted lexical sequences were well dispersed throughout a four-week period.

Each reading was seeded with two target vocabulary items such that within a given week, participants had a single exposure to a total of four different target vocabulary sequences through reading. Once they received a link, the participants accessed a Google Form, which included a reading as an embedded image. At the top of the electronic form, in both English and in the participant's L1, participants were reminded not to use their real names and were asked to type in their pseudonym. At times, participants would choose new pseudonyms without telling me, which sometimes caused difficulty in determining whose data had been collected. Fortunately, this only occurred occasionally, so a missing score easily matched up with an extra pseudonym. When it did happen, I changed all records to include both pseudonyms separated with a forward slash (e.g., Black love/Purple flower; Beargamer/George Brown). For writing tasks, the Google Form included a writing prompt in English as well as a translation into the participant's L1. An example of a Google Form appears as Appendix D. Once they had completed the reading task, participants clicked on a box next to the statement, "I have read the paragraph." When tasks

required participants to write a response to readings, the participants were unable to submit the form until they responded to the prompt by writing something in the area allotted for this purpose.

The school's ESOL Department requested that the study end before May 19, 2022, one week before the end of the original study plan. Consequently, the delayed posttest of lexical sequences 5-8 took place three weeks, instead of four weeks, after the posttest. Vocabulary acquisition studies examined as part of the literature review did not always include a delayed posttest, and when they did include one, it was most often defined as a period of two weeks and often showed attrition after that period of time. (See Appendix A for a reporting of all delayed testing for literature reviewed.) Therefore, I determined that a delayed posttest of three weeks would still be useful in determining whether attrition occurred.

Quantitative Data Analysis

Quantitative data collected through Google Sheets were copied into an Excel spreadsheet. All statistical analyses were undertaken using SPSS 28.0 with confidence levels set at .05 with group, L1, and gender analyzed as between-subject factors and test and intervention as within-subject factors. A repeated measure test was conducted to examine the impact of group number, L1, and gender on intervention outcomes and to examine whether and to what extent attrition occurred after a one-month delay in the case of sequences 1-4 or a three-week delay in the case of sequences 5-8.

Quantitative Validity. Internal threats may include history, maturation, testing, instrumentation, selection of subjects, mortality, and selection-maturation (Campbell & Stanley, 1963; Dörnyei, 2018). While some of these threats can be ameliorated through statistical testing, some cannot. For instance, I could not control attrition, nor could I control events that occurred

between tests. However, I found the use of qualitative investigation helpful in awareness and reporting of these potential threats. Fortunately, only one participant was removed because of failure to complete the tasks in the given amount of time, and interviewees did not report any difficulties or events that would have impacted study results.

Conducting statistical tests requires great care, given the pitfalls and apparent propensity for educational and social science researchers to fall into them (Coe, 2002; Cohen, 1988; Keselman et al., 2001; Osborne & Waters, 2002; Oshima & Neel, 2003). For instance, assumptions of statistical tests are often violated (Osborne & Waters, 2003), and studies often omit whether assumptions were tested (Osborne & Waters, 2002). According to Cohen (1988), published research reveals that “statistical power is frequently not understood” (p. 1). Keselman et al. (2001) express disapproval of using repeated measures ANOVA as a catch-all phrase, failing to reveal the type of analysis or the calculated validity of the p value. Moreover, APA requests that reliability estimates be reported, yet this advice is rarely followed (Dörnyei, 2018). Finally, (and perhaps most notoriously), researchers of all stripes often fail to report effect sizes (Coe, 2002; Maher et al., 2017; Schuele & Justice, 2006; Sullivan & Feinn, 2012). I hope that careful attention paid to such advice ensured me steady footing.

For trustworthiness in results, I ensured that all assumptions of repeated measures ANOVA were met: the presence of at least two categorical variables, no significant outliers, approximate normal distribution of the dependent variable, independence of observations, and sphericity (Verma, 2015). When assumptions were violated, I followed procedures recommended by research to obtain reliable results. For instance, during the repeated measures test, sphericity was found to be violated, leading me to use results from Mauchley’s test to determine significance (Vogt, 2011).

I also heeded advice in the use of tools for avoidance of Type I and Type II errors and cautions against making common errors found in published research. Onwuegbuzie and Collins (2007) claim that most quantitative studies use samples sizes that are not large enough to detect difference, which leads to inadequate power, and in their estimation, nearly half of all published mixed methods studies have such low power that they report significant findings erroneously. I took precautions to avoid these errors. For instance, I used G*Power to calculate the needed number of participants for the chosen statistical test, but I recruited 36% more participants than G*Power recommended. I attended to robustness by ensuring that the actual alpha level was near the nominal alpha by having equal group numbers and through the use of *F* statistics (Stevens, 2007). I strove for content validity by using the VKS simplified by Brown (2008) for all tests because it has a long history of providing valid results, having earlier undergone validity tests for my participant population (e.g., Stewart et al. 2012¹⁰; Tan, 2016).

Notwithstanding the pretest results, participants' scores were automatically sent to Google Sheets as soon as they were submitted, which were then copied and pasted into an Excel spreadsheet. This process helped me avoid reporting errors. I copied the data into SPSS twice, on two separate occasions, and undertook the repeated measures testing twice for the sake of ruling out errors. Pretest scores were collected through Google Forms before I became aware that a setting allowed the scores to be transferred automatically to a database for easy transfer into SPSS. Consequently, only pretest scores were manually entered into the spreadsheet.

Qualitative Strand

As Dörnyei (2018) puts it, quantitative data alone “(cannot) do justice to the subjective variety of an individual life” (p. 35). For this reason, I also conducted qualitative research as a

¹⁰ The adapted version of the VKS recommended by Brown (2008) aligns with Stewart et al.'s (2012) proposal that of the original 5 stage scale, Stages 3 and 4 was collapsed into a single descriptor.

separate strand to determine how L1 and gender in addition to other factors could affect the intervention results of my study (e.g., lived experiences as defined in Chapter 1, including educational experiences, past experiences, responsibilities, and family) and language learning in general.

Participants

The participants for the qualitative strand were 12 students chosen from the quantitative strand who were invited to take part in hour-long interviews¹¹. The demographic information for the interviewees appears after a description of the recruitment process.

Sampling Size and Sampling Scheme

The purposive sampling I employed is what Teddlie and Yu (2007) call “maximum variation sampling,” designed to “achieve representativeness or comparability” (pp. 80-81), and what Onwuegbuzie and Collins (2007) call *criterion-based*. Such sampling allowed me to choose participants that were representative of a variety of backgrounds, and with it, the potential of rich and diversified data that helped inform the quantitative phase of my research. A minimum sample size of 10 has been recommended for interviews in mixed-methods studies by Onwuegbuzie and Collins. However, for greater representation, at the end of the second intervention, I invited 12 participants who represented the gender, L1, ethnicity, AOA, and intervention outcome range of the larger participant group for interviews. All invited participants agreed to be interviewed. As can be seen in Table 4, The demographics of the 12 interview participants are comparable to those of the total sample except in the case of race. Because of the strong impact of being Black on educational experiences (Ladson-Billings, 2021), I interviewed

¹¹ Although participants were invited to hour-long interviews, the interviews typically lasted between 30-45 minutes.

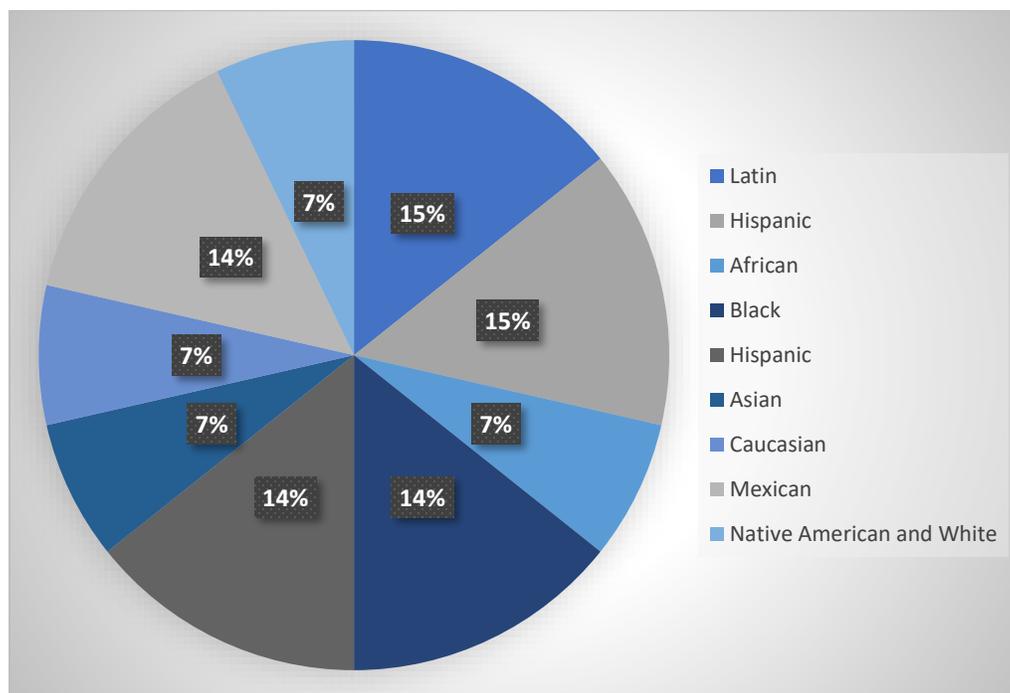
both Black participants. I also interviewed the only participant who identified as a refugee. No participants revealed having had interrupted schooling.

Table 4

Comparison of Demographics, Participants of Both Strands

	Total Sample	Interviewees
Mean AOA	9	9.8
Male	45%	50%
Female	55%	50%
Mean Current Age	14.76	14.75
Born in the US	31%	42%
Race		
Black or African	4%	17%
Latin, Latina, Latino, Hispanic, Mexican, Mexican American, Native American	71%	75%
Asian	6%	8%

The interview participants reported race and ethnicity with the same variation found in the larger sample (See Figs. 11 and 12). Because I invited participants to describe their race and ethnicity in their own way, several unexpected descriptions appear.

Figure 11*Race as Described by Participants***Figure 12***Ethnicity as Described by Interviewees***Instruments and Materials**

To plan for the interviews, I followed advice offered by Castillo-Montoya (2016) by designing a matrix that connected interview questions, and I adapted some of their questions to

elicit data about participants' lived experience (Fig. 13; for the interview guide, which includes probing questions, see Appendix F). These questions included how learning may be impacted by school experiences (Lin Lee, 2020; Sandh, et al., 2020), family (Das & M, 2017; Maroun, 2018; Spagnola & Fiese, 2007) responsibilities (Mortimer, 2010; Staff et al., 2020) or past experiences (Dube, 2018; Garner, et al., 2012).

Figure 13*Interview Matrix*

	Procedures	Gender	L1	Education	Family	Responsibilities	Past Experiences
Tell me what you think about the lessons I asked you to complete.	X	X	X	X	X	X	X
Did you have any problems understanding the directions?	X	X	X	X	X	X	X
What exactly did you do for the reading part of the lesson?	X	X					
What exactly did you do when you wrote the paragraphs?	X	X					
What exactly did you do when you saw words in bold?	X	X					
Did you have any interruptions during reading or writing?	X	X			X	X	
What is hard about reading in English?	X	X	X	X	X	X	X
What is hard about writing in English?	X	X	X	X	X	X	X
What is easy about reading in English?	X	X	X	X	X	X	X
What is easy about writing in English?	X	X	X	X	X	X	X
What was the hardest thing about the lessons?	X	X	X	X	X	X	X
What was the easiest thing about the lessons?	X	X	X	X	X	X	X
In lesson --, you wrote about ---. Tell me more about that.	X	X	X	X	X	X	X
Is there anything else you would like to say about things that make learning English easy or hard?	X	X	X	X	X	X	X
Tell me about your school.		X		X			
Are there things at school that make learning English hard? Follow up: What things? Tell me more.		X		X			

Are there things at school that can make learning English easy? Follow up: What things? Tell me more.		X		X			
Sometimes there are differences in how people are treated at school. Do you think you are treated or viewed in any special way at school, by students, teachers, or anyone else? <i>Follow up:</i> In what ways? <i>Follow up:</i> Who does this?		X		X			
Let's go back and look at some of what you wrote. I noticed that in lesson (s)---- you wrote less (or more) than other lessons. Why do you think that is?	X	X	X	X	X	X	X
Do you have a job? Follow up: What is your job? What do you do at work? How many hours do you work? Are your hours usually the same?		X	X		X	X	
Tell me what your average day is like. What happens when you get home from school?		X	X		X	X	
Tell me about your weekends. What do you do on the weekends?		X	X		X	X	
Tell me about your family.					X		
Are there things at home that make learning English hard? <i>Follow up:</i> What things?		X			X	X	
Are there things at home that sometimes make learning English easy? Follow up: What things?		X			X	X	
What types of things do your (mom/dad/guardian/etc.) believe about school?		X			X		
How did you come to the United States?		X					X
Tell me about the ways that life is different in the United States.		X					
<i>(For Refugee Students)</i> In your questionnaire, you said that you are a refugee student. What was school like in ...?		X					X
<i>(For Students with Interrupted Schooling)</i> In your questionnaire,		X					X

you said that you have gone to school for ... years. Tell me more about the times you were not in school.							
---	--	--	--	--	--	--	--

Data Collection

The qualitative strand data consist of semi-formal interviews scheduled for one hour, but lasting 30-45 minutes, with 12 participants. All interviews took place over a two-day period during the week of May 9, 2022. At the request of the principal, interviews were conducted in person during school hours in a conference room at Ridgemont High School. According to U.S. News and World Reports (2023), the school has a population of over 2,000, with over 80% Black, Asian, or Latinx; over half of students are economically disadvantaged. All interviews were audio recorded and transcribed with the initial help of Trint software; after allowing the software to make initial transcription, I listened carefully to the recordings several times. Because of the importance of transcription quality to rigorous research (Poland, 2000), I adjusted the software generated transcriptions using conventions recommended by Roulston (2013), which are intended to help the researcher capture nuances in communication and dynamics between the researcher and the interviewee. These conventions fostered reflection on how well I listened to the interviewees and how well I offered relevant follow-up questions. Roulston's transcription conventions include notations for changes in speech tempo, inflections, and volume used for emphasis as well as a mechanism for recording interruptions. I listened numerous times to each recording to represent the interviewees spoken words as faithful as I could. Even now, when I read the paper transcriptions, I can hear the voices of my participants and the nuances of our discourse. Once the transcriptions were complete, audio recordings were immediately

deleted from my computer and from my telephone. To retain confidentiality, all data were deleted from Trint Software as well.

Qualitative Data Analysis

Once interview data were collected, the second stage of the qualitative data analysis occurred. The analysis of interview data were arguably the most time-consuming part of my dissertation project. I conducted the analysis multiple times with guidance from my advisor at multiple steps. I followed advice from Peel (2020) and Xu and Zammit (2020), who believe that a thematic analysis approach is a suitable choice for fledgling researchers, in part because of its flexible nature. I choose Braun and Clarke's (2006) thematic analysis guide, using inductive coding as an iterative process.

The six-phase process began as Braun and Clarke (2006) suggest, with gaining familiarity with the data and with attention to patterns. First, I read all of the transcripts. Braun and Clarke explain, "the keyness of a theme is not necessarily dependent on quantifiable measures – but in terms of whether it captures something important in relation to the overall research question" (p. 10). Consequently, during a second pass over the data, I made a list of 18 initial codes, or ideas that occurred frequently, as well as those that seemed most important in answering the two qualitative research questions. This list appears below in Figure 14; the qualitative questions are the following:

(QUAL) What other factors impact intervention results?

(QUAL) What factors affect the educational experiences of high school bilingual learners?

During Phase 2, the initial coding process began. Braun and Clarke (2006) describe this process as involving a return to the data again to closely examine how prevalent each of these 18

initial codes were, and as an iterative process, it also entailed making changes to the codes. I kept my preliminary list by my side as I looked at the interview data line-by-line to see how well these initial codes could guide my process and to capture any new codes as well. During this review, I began to notice that the experiences of some individuals seemed to dramatically impact their schooling, and I noted that these experiences were closely tied to their identity constructs. For instance, students who reported having been bullied for being bilingual learners also reported a reluctance to engage with other students at school as a result. With this observation in mind, I added a new code called *social identities*. At the end of the process, I had formed a list of 19 codes with the understanding that many of these initial codes would be collapsed later.

I found it useful to describe these initial codes in the form of a table to aid me in the process of inferring the meaning of the codes. The initial codes and their accompanying descriptions appear as Figure 14. These descriptions would be later extended and refined.

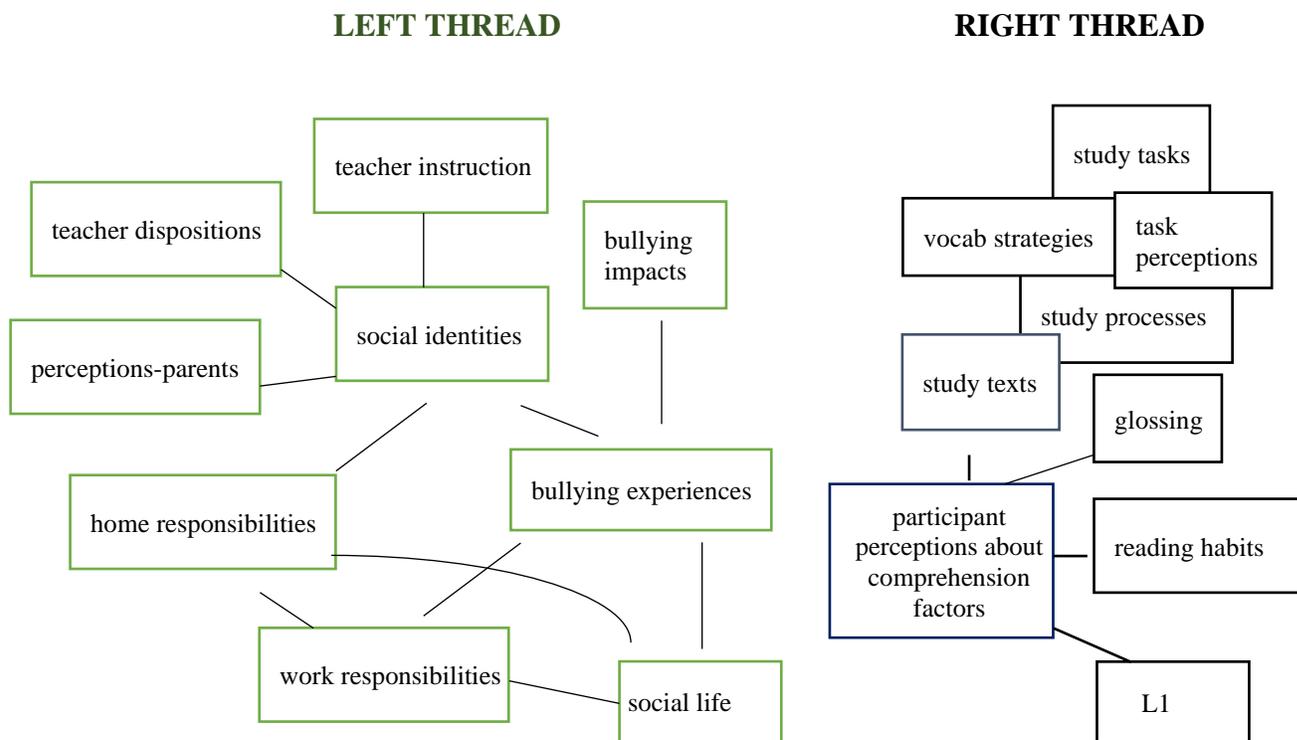
Figure 14

Initial Codes and Descriptions

Initial Codes	Initial Descriptions
Teacher dispositions	Adjectives used to describe teachers
Teacher instruction	Descriptions of instruction and teaching practices
Work responsibilities	Jobs that participants hold (paid)
Home responsibilities	Responsibilities at home, including caring for children and housekeeping
Perceptions about parents	Descriptions of parents
Family life	Descriptions of family activities
Study intervention processes	Approaches taken to complete tasks
Quarantine	All discussions about the quarantine
Perceptions of study texts	What participants thought about the readings
Perceptions about the study tasks	What participants thought about the tasks
Vocabulary strategies	Strategies participants use for unfamiliar vocabulary they encounter
Bullying experiences	Any mentioning of bullying that has occurred at school
L1	Any mentioning of the use of L1

Bullying impact	Any mentioning of the impact of bullying, including a reluctance to speak because of bullying
Participant perceptions about comprehension factors	Factors that participants believe makes reading difficult or easy
Social identities	How participants describe themselves in terms of their social identities
Reading habits	Any mention of reading habits, including how they have changed over time and why
Perceptions of glossing aspect of study	Any mentioning of glosses
Social life	Discussions about social life of participants, including drug use and time spent with friends outside of school

To help me clarify my data, I created a concept map (See Fig. 15). The creation of this concept map led me to a key understanding about my data. I realized that my data followed two discrete threads—one confined to literacy (right thread) and another concerning lived experiences (left thread). There seemed to be no overlap between perceptions about reading and writing processes and life experiences as they were relayed to me. Connections I made between initial codes within the left thread assisted me in understanding how social identities served as a driving factor within these topics, and in reviewing the data, I noticed there was often a disconnect between things implied and stated by the participants. I grew to understand that this was a general pattern found throughout the data.

Figure 15*Concept Map Linking Initial Codes*

The next step in the process required me to begin coding the data. Educational research and literacy theories could assist me in understanding the right side of the concept map (items from Figure 16 with a black outline—those related to literacy), but I would have struggled to gain an understand of the left side of my concept map (lived experiences) without the use of a framework. In accordance with Braun and Clarke (2006), I therefore sought and found an applicable model in Núñez’s (2014) multilevel model of intersectionality, which I had encountered during the planning stages of my project. I kept the model in mind because it had described the educational experiences of immigrants and because it framed these experiences in terms of social identity intersections. Because my participants were immigrants and my study

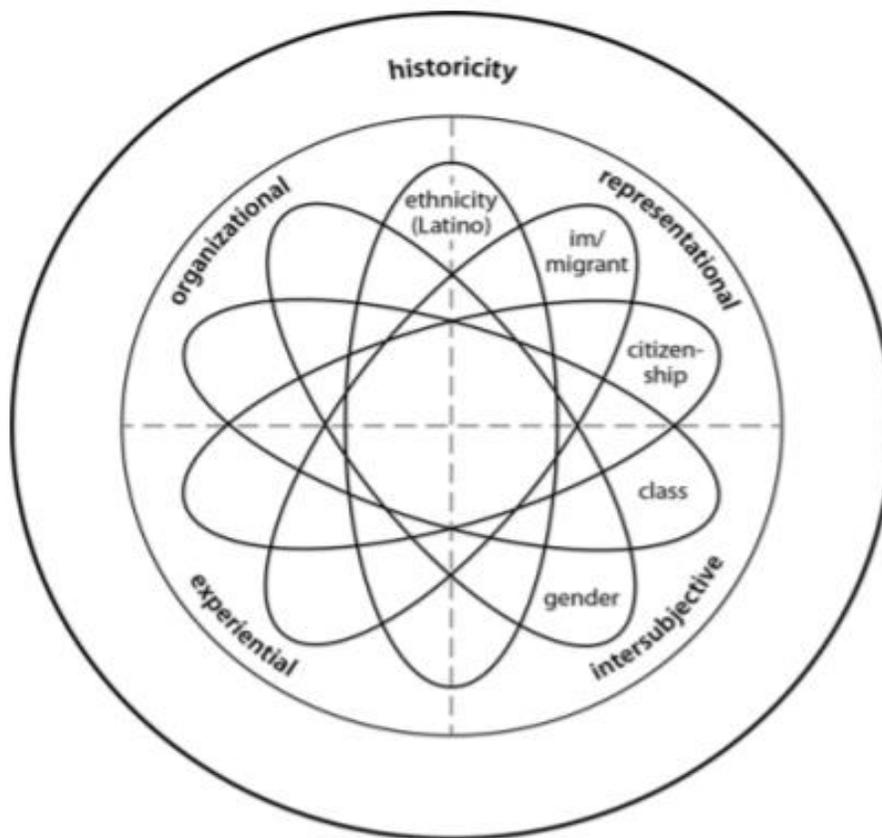
sought to examine their educational experiences in terms of their social identities, the model held promise to assist me in coding my data.

Model of Multilevel Intersectionality

In 2012, Floya Anthias introduced a framework intended to assist intersectionality researchers by differentiating among social categories investigated, social arenas where intersectionality is enacted, and historicity. Núñez's (2014) multilevel model of intersectionality provides a visual representation of Anthias's framework. Núñez also provides a concrete example of how to apply the model in an educational context among immigrant students, which served as an example for my own analysis. The model appears as Figure 16.

Figure 16

Multilevel Model of Intersectionality



Núñez's (2014) model examines the impact of multiple arenas on the educational experiences of immigrant youth. The innermost level of Núñez's model includes social identities, such as immigrant status, bilingual learner status, L1, and citizenship status, race, gender, and ethnicity. The middle level of the model describes four arenas within which social identities may be embodied, including organizational arenas (spaces where interactions occur), representational arenas (how individuals are represented within the spaces), intersubjective (interactions occurring within these spaces) and experiential arenas (how individuals make meaning of experiences occurring within the spaces). Anthias (2012) views these delineations as a heuristic device, explaining that "setting them out helps to organize the types of issues that we focus on and allows comparisons across these" (p. 11). The outer layer of the model, historicity, includes synchronic outcomes and diachronic processes, thereby calling attention to temporal context. In other words, this layer emphasizes that the influence of a given social identity and its intersection with other social identities vary in impact within context and time. I believed that the second layer of this model—the four arenas—would be suitable for helping me make sense of much of my qualitative data and would provide relevant and useful typology that would help me with the subsequent steps of the coding process.

Anthias (2012) describes the organizational arena as a structural one, with a focus on how social identities are categorized within institutions. She includes family structures, networks, and educational systems. Núñez (2014) similarly defines this area as "positions in structures of society such as work, family, and education." Language acquisition may be impacted by family structures and routines (Das & M, 2017; Spagnola & Fiese, 2007), adolescent employment (Mortimer, 2010; Staff et al., 2020), and school policies and practices (Lin Lee, 2020), which are recurring concepts in my data. While organizational influences

include the institutional spaces that individuals inhabit, representational influences include evidence of how individuals are represented in those spaces. Núñez (2014) explains that representations of individuals within society lead to discriminatory actions. For example, teacher perceptions of students may impact their expectations for the students. Núñez defines the intersubjective arena as one that “concerns how people and groups relate to each other and influence educational opportunities” (p. 89). It concerns how relationships and interactions affect learning. Intersubjective practices therefore include practices in relation to others, including intercommunication. In this way, the model accounts for the impact of interactions with others, which is prevalent within my study’s qualitative data. A consideration of the representational arena helped me make sense of how social identity representation impacted my participants’ schooling experiences. For instance, the representational arena accounted for how the perceptions of other students about minority social identities—such as bilingual learner status and disability—led to bullying. Finally, according to Anthias, the experiential arena includes “the affective, the emotional, and the body. This includes narrations of identification, distinction, and othering” (p. 11). Considering the experiential arena led me to the realization that one participant’s narrative of self-identity served to empower her.

Anthias (2012) and Núñez’s (2014) descriptions lent me confidence that examining my data with the use of the multilevel model of intersectionality in mind would allow me to construct meaningful understandings of how my participants’ experiences and identity factors—including status as an ELL, L1, race, and gender—affect their lives and language learning journeys.

In coding the data, I arranged individual passages in accordance with the codes, adding the corresponding participant pseudonyms next to each. This process required carefully

considering where individual passages from the data would belong, and in turn, led to several instances of collapsing codes. These cases included *teacher dispositions* and *perceptions of teaching quality* collapsed into the larger units of analysis *teachers* and *family life* and *perceptions about family* being collapsed into the single code *family*. In both cases, combining these excerpts under a single umbrella would highlight conflicting accounts offered by single participants, such as participants' descriptions of teachers as both effective and ineffective or the absence and presence of bullying. The codes relating to the study intervention were also collapsed into *study intervention*. Some codes were reconsidered as a result of close examination of the data being coded. For example, in the case of *quarantine*, two different discussions were apparent in the data: one regarding the impact of the quarantine on reading habits and a second one discussing the impact of the quarantine on social life. Consequently, the *quarantine* code was recategorized into two other existing codes: *social life* and *reading habits*. I found that items originally coded *reading comprehension factors* either fell under other codes or were confined to discussions about vocabulary, so this code was replaced with the new code *vocabulary*.

In reviewing my work, I had difficulty in deciding whether some specific passages belonged where I had placed them. For assistance, I returned to read Anthias's descriptions for clarification whenever needed. For example, the importance and support of family were prominent concepts throughout interview transcripts, but I was initially unsure which arena they best aligned with; as family, they were inherently organizational, yet they were also intersubjective because they involved interactions with others. Revisiting Anthias's description of organizational arenas led me to determine that discussions of family responsibilities aligned with the organizational arena. However, as interactional, discussions about the support that parents provided belonged to the intersubjective arena. Likewise, a problem arose in determining

which arena passages coded as *teacher* should be aligned because both could be considered representational as well as intersubjective. I solved this dilemma by making a clear distinction between intersubjective and representational influences. Intersubjective was clarified as influences that were interactional in nature, whereas representation was reframed as being more passive in nature. I completed this entire process a second time to ensure that other coding conflicts were not apparent.

In keeping with Phase 3 of Braun and Clarke (2006), I began to describe my codes with more depth in my codebook. I continually reviewed my codes to see how well the arenas of influence were supported by the interview data. The final list of codes and their modified descriptions appear in Figure 17.

Figure 17*Modified Codes and Descriptions*

Code	Description
Study intervention	Any discussions about the study tasks, including perceptions about the text, approaches to the tasks, use of glosses, novelty of the tasks, content of texts, genre of text
Literacy habits	How participants discuss literacy habits, including the impact of the quarantine on changing habits, the impact of the study intervention on changing habits, and outside influences on literacy habits
Jobs	How participants discuss their jobs, whether paid or otherwise. Does not include responsibilities at home.
Home responsibilities	How participants discuss responsibilities at home, including childcare and housework
Social identities	How participants discuss themselves in terms of social identities, including gender, L1, race, ethnicity, status as ELL, or LGBTQ+
Intercommunication	How participants discuss their intercommunication with their peers, including reluctance to speak
Family	How participants describe their families, including interactions with family: siblings, parents, and others in the household; how participants describe their parents, including parent perceptions about education and things parents teach them, values parents instill, and difficulties faced with family
Teachers	How participants describe their teachers' dispositions or the quality of teaching they receive. Includes teaching strategies and their effectiveness or lack of effectiveness
L1	All description of the use of L1, including literacy practices in L1
Vocabulary	Any discussion about the use of strategies taken to understand unfamiliar vocabulary encountered while reading and discussions of vocabulary as causing reading difficulties

The next phase required me to look at each individual code and to construct categories that took into account the passages coded within them. I returned to the coded passages and considered whether the arenas of influence from the multilevel model of intersectionality could assist in the construction of categories. An experienced qualitative researcher who examined my coded data assisted in my understanding that only three arenas were supported by the data; another experienced qualitative researcher, my advisor, gave me confidence that the model could

still assist me, even if my codes only aligned with three of the four arenas. I found that my codes supported all arenas except the experiential arena, and I found that three of my codes—interventions, literacy habits, and vocabulary—did not find a home in any arena. Consequently, I created a new arena and called it *literacy*. I aligned the remaining codes to the four arenas of influence.

Looking closely at my original codes with conjunction with the aligned arenas, I began to identify categories for each by examining what coded passages had in common with one another. For example, job and family responsibilities both fell under the organizational arena, and both included responsibilities that participants often described in contradictory ways. As revealed by the data, I found that excerpts coded as *intercommunication* and *social identities* were often two sides of bullying: the action and its impact. They inherently involved intersubjective experiences with others, yet they are forms of discrimination often having arisen from representational influences (Boske, 2015).

The descriptions and the connections to the arenas that I generated during Phase 3 helped me understand how elements of my data related to one another in a manner that also helped in the generation of themes. For example, the application of the intersubjective arena code helped me make a connection between language use at home with family members and the interactions my participants had with their teachers. With codes aligned to arenas side-by-side, I began to name categories that would describe the codes with detail in light of their respective arenas. The naming process helped create a space for me to understand how I would discuss these codes as themes. For example, the process allowed me to combine bullying into a single category that would include both its representational as well as its intersubjective aspects.

The movement from categories to themes was an organic process during which I simply collapsed the codes in accordance with their categories. To construct these themes, I returned to coded passages for more specificity. For example, two codes that I had categorized as *responsibilities* became the *impact of responsibilities* because passages coded under this category most often described the impact of responsibilities on the students' lives. In the case of literacy, the naming of the categories led me to understand that the passages coded as literacy seemed to reveal influences on literacy practices. Following this process, I created three additional temporary themes from the remaining categories: *literacy practices*, *the impact of responsibilities*, *the impact of bullying on language use*, and *the impact of relationships*. During Phase 4, in accordance with Braun and Clarke's (2006) recommendations, I reviewed the themes to seek understanding of how they were supported by the interview data. To do this, I looked carefully at each theme and their coding to ensure that the themes I constructed were distinctive from one another, with no overlap. I checked the themes against the coded extracts and found that they were suitable, so I did not make any changes.

Following Phase 5 of Braun and Clarke's process, I began to define and describe my themes more distinctly. The naming process led me to clarify the relationships among my themes and how they helped to answer my qualitative research questions. I also wished to make my themes more descriptive, concise, and powerful. This led to several iterations of theme titles. Themes became renamed as *novel practices and motivations*, *claimed and unclaimed responsibilities*, *bullying as a life-altering event*, and *real and imagined intersubjectivities*. Figure 18 delineates the codes I constructed as well as the relationships among codes, arenas, categories, and the four themes, along with their final names.

Figure 18*Alignment of Codes, Categories, and Themes*

Code	Arena	Category	Theme
Intervention	Literacy	Influence on reading and writing	Influences on Literacy Practices
Literacy habits	Literacy	Influence on reading	
Vocabulary	Literacy	Influence on reading	
Jobs	Organizational	Responsibilities	Claimed and Unclaimed Responsibilities
Home responsibilities	Organizational	Responsibilities	
Social identities	Representational	Bullying	Bullying as a Life Altering Event
Intercommunication	Intersubjective	Bullying	
Family	Intersubjective	Interactions	Contrasting Intersubjective Experiences
Teachers	Intersubjective	Interactions	
L1	Intersubjective	Interactions	

Validity. Within qualitative research, research validity spans the entire research process (Dörnyei, 2018). Internal validity is “soundness of research,” while construct validity lies in “truthfulness in interpretation” (Dörnyei, 2018, p. 51). Validity has been described as credibility and trustworthiness (Duff, 2008; Stake, 2005), “meaningfulness of results” (Riazi, 2017, p. 16), “trustworthiness, authenticity, credibility, rigor or veracity” (Dörnyei, 2018, p. 49), and “the boundary for what is acceptable and not acceptable in research” (Freeman et al., 2007, p. 27). I tried to carefully examine my processes to ensure that others would find veracity in my study and therefore find applications for it in teaching and learning. I attempted to engage in member checking using recommendations outlined by Amankwaa (2016). Three weeks after the interviews, I provided interviewees with their individual interview transcripts that were password protected with their interview pseudonyms as their passwords. I asked them to contact me by phone if they had forgotten their pseudonyms. I asked them to read the transcript to tell me if

they believe it is accurate. I intended to change any transcripts to reflect any misunderstandings that I discovered in participants response. At the end of the data analysis, I provided them with a brief description of themes and invited them to respond. Unfortunately, however, my interviewee participants declined to take part. To strengthen the validity and reliability of my coding process, I asked two experienced qualitative researchers to examine sections of coded interview transcripts and my codebook and elicited their feedback.

Braun and Clark (2019) outline misunderstandings that researchers ostensibly using thematic analysis for data analysis sometimes have, which leads to missteps in research reporting. First, researchers sometimes describe themes that they claim have “emerged” and fail to “transparently describe the processes engaged in to produce the themes reported” (p. 591). Themes do not emerge from data (Braun & Clarke, 2020; Braun & Clarke, 2007; Brown, 2018; Rintala & Nokelainen, 2019; Smith & Sparkes, 2020) but are generated by the researcher. To state that themes emerge from data is to ignore the subjective nature of all research. When describing thematic patterns, I acknowledged my role in the generation of themes from the qualitative data, with care given to the language I used to describe the process. I also revised the qualitative strand section of my methodology chapter to describe the data analysis process with enough detail that a replication of the process would lead to similar results.

Braun and Clarke (2006) explain that another common error in thematic analysis research is the confusing of data domains and topics with themes. I made the error that Braun and Clarke describe when I confused the four arenas with four themes and therefore can attest to the ease in making it. The error forced me to undertake the qualitative analysis a second time. Although the findings remained the same, the construction of themes *informed* by my codes and their corresponding arenas—as opposed to my *conflation* of them—lead to more sound organization

of my findings section, thereby aiding both myself and the reader in understanding how my findings helped to answer my research questions.

Mixed Methods as Intergrative

Data collection, data analysis, and interpretation of results integrated both quantitative and qualitative strands. I collected both quantitative and qualitative data concurrently throughout the study. Quantitative data were collected as testing data at the beginning of the study, at the end of the study, and at two points in between. Between testing events, qualitative data in the form of written responses to readings were collected while each group completed the writing intervention twice per week for four weeks. As I analyzed the effect of group, intervention, L1, and gender on intervention results, I sought explanations from the qualitative data to help me understand the quantitative results. For example, the examination of written responses led me to quantify the written results and measure the impact of usage of individual lexical sequences on participants results. In the end, I was able to examine the effect of group, intervention, L1, gender, and lexical sequence usage on results.

At the end of both interventions, qualitative data were also collected in the form of interviews with 12 participants. From the qualitative strand, I constructed four themes based on the findings: *influences on literacy practices*, *claimed and unclaimed responsibilities*, *bullying as a life altering event*, and *contrasting intersubjective experiences*. Once the analysis of both strands was completed, a final mixed methods analysis was undertaken. This involved an iterative examination of the intervention results as well as themes I constructed from the qualitative data analysis.

To conduct the mixed methods analysis, I placed the findings from the quantitative data—intervention results and the impact of L1, gender, and group membership—next to my

discussions of the four themes. Examining the quantitative data through these themes assisted in my discovery of how not only L1 and gender, but other social identity factors—such as race, bilingual learner status, and disability—have impacted the educational experiences of my participants. I found that collectively, the quantitative and qualitative data fell into two categories: factors influencing intervention outcomes and factors influencing educational experiences. This led me to determine that lexical sequences, text relatability, and usage during the writing intervention were the most relevant factors influencing intervention outcomes. The factors influencing language learning and the lives of participants were found within the qualitative data to include work, family, school, bullying, and the pandemic. These factors were driven by social identities that included bilingual learner status, L1, race, gender, and disability status. The findings of both strands, as well as my interpretations of how together, the strands created a greater understanding of language learning among my participants, are reported in Chapter 4.

4 FINDINGS

The purpose of this study was to uncover factors affecting bilingual learners' acquisition of lexical sequences through reading and through reading followed by the intentional use of lexical sequences through writing. The quantitative study strand intended to measure the effect of a learners' first language (L1), gender, and intervention on lexical sequence acquisition. The qualitative study strand intended to uncover any additional factors influencing lexical sequence acquisition and language learning generally. Together, analyses of findings from both strands determined which combination of factors influenced the intervention results and the educational experiences of the high school bilingual learner participants of the study. In this chapter, I report all findings.

Quantitative Findings

As earlier mentioned, all statistical analyses were undertaken using SPSS 28.0 with confidence levels set at .05. During the initial quantitative strand analysis, a factorial ANOVA with repeated measures was performed. This test was conducted to examine the impact of group number, L1¹², and gender on intervention outcomes and to examine whether and to what extent attrition occurred after a one-month delay in the case of sequences 1-4 or a three-week delay in the case of sequences 5-8. The within-subject factors were therefore intervention (two levels: reading intervention and writing intervention, and test (three levels: pretest, posttest, and delayed posttest), whereas the between-subject factors were group, L1, and gender. Because the data were normally distributed, outliers were identified using the criterion that any z -score greater than 3 or less than -3 on the dependent variable was a univariate outlier (Geraghty, 2022). The z

¹² L1 is defined as the language individuals identified as the one they felt most comfortable using. As noted in Chapter 3, because the sample sizes of all languages other than Spanish or English were no greater than $n = 2$, the language comparison made was L1 reported as English or non-English.

score method was chosen because it is the most common method of identifying outliers across all disciplines (Erkuş & Purutçuoglu, 2021). No outliers were found. The z scores for all tests appear in Appendix G. Mauchly's test of sphericity indicated a violation of sphericity ($p < .001$).

Consequently, Huynh-Feldt values were used to determine significance for the within-subjects analysis and are reported in this chapter.

A cursory glance of Tables 5 and 6, which provides descriptive statistics for the initial repeated measures test broken down by intervention, shows that Group A had higher scores for the reading intervention, Group B had higher scores for the writing intervention, and both groups made gains during both interventions. The tests of between-subject and within-subject effects showed whether these differences were significant and whether any interaction effects occurred.

Table 5

Descriptive Statistics—Reading Intervention

Test	Independent Variable	Mean	SD	n
Pretest	Group A	6.8750	2.17321	24
	Group B	6.6400	2.44745	25
	Female	6.9259	2.41670	27
	Male	6.5455	2.17622	22
	Non-English	6.7143	2.22539	28
	English	6.8095	2.44170	21
Posttest	Group A	8.9583	1.89918	24
	Group B	9.6800	1.90875	25
	Female	9.2222	2.02548	27
	Male	9.4545	1.81861	22
	Non-English	9.0357	2.02726	28
	English	9.7143	1.73617	21
Delayed Posttest	Group A	8.6250	2.31840	24
	Group B	9.9600	1.94679	25
	Female	9.5185	2.34308	27
	Male	9.0455	2.08115	22
	Non-English	9.0000	2.43432	28
	English	9.7143	1.87464	21

Table 6*Descriptive Statistics—Writing Intervention*

Test	Independent Variable	Mean	SD	<i>n</i>
Pretest	Group A	7.1667	2.80786	24
	Group B	6.6800	2.24944	25
	Female	6.9259	2.51038	27
	Male	6.9091	2.59870	22
	Non-English	7.0000	2.40370	28
	English	6.8095	2.73165	21
Posttest	Group A	10.3750	2.08123	24
	Group B	9.0000	1.82574	25
	Female	9.4444	2.02548	27
	Male	9.9545	2.10390	22
	Non-English	9.2857	2.33900	28
	English	10.0952	1.70014	21
Delayed Posttest	Group A	10.7083	1.65448	24
	Group B	8.6000	2.00000	25
	Female	9.2222	2.25889	27
	Male	10.1364	1.83343	22
	Non-English	9.2857	2.33900	28
	English	10.0952	1.70014	21

Within-Subjects Effects

Main Effects. Research Question 1 and 1a asked if exposure to targeted lexical sequences through two interventions—hereafter referred to as *reading intervention* and *writing intervention*—can lead to acquisition of these sequences. To answer this question, I conducted a factorial repeated measures ANOVA. The within-subjects effects showed significance among at least two tests ($F(2, 41) = 36.17, p < .001$) with a large effect size ($\eta^2 = .64$). Research Question 1b asked whether there is an advantage of one intervention over another. The repeated measure within-subjects effects showed no advantage for any intervention between any two tests ($F(1,41) = 2.6, p = .114$).

Simple Effects. Pairwise comparisons were examined to see where the statistical significance among tests lay. Differences were found between the pretest and posttest ($p < .001$)

and between pretest and delayed posttest ($p < .001$). However, there was no statistical significance found between the posttest and delayed posttest ($p = .99$). This was interpreted to mean that both interventions led to gains in lexical sequence acquisition generally, with no significant losses three to four weeks later.

Between-Subject Effects

Main Effects. To answer Research Question 1d—whether any interaction effects exist for L1 or gender—between-subject effects were analyzed to determine the impact of group number, L1, and gender on test results. As can be seen in Table 7, no significance was found for group, L1, or gender, and no interaction effects were seen among these variables. This was initially interpreted to mean that order of presentation did not impact results and that there was no impact of these variables on test differences. However, a repeated measures statistical test by group later showed this to be inaccurate interpretation. (See “Interaction Effects” below.)

Table 7

Between-Subject Effects: Group, L1, and Gender

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^a
Group	4.28	1	4.28	.26	.612	.006	.261	.079
Gender	.005	1	.005	.00	.986	.000	.000	.050
Language	4.69	1	4.69	.29	.596	.007	.285	.082
Group * Gender	7.67	1	7.67	.47	.498	.011	.467	.102
Group * Language	3.14	1	3.14	.19	.664	.005	.191	.071
Gender * Language	1.74	1	1.74	.11	.746	.003	.106	.062
Group * Gender * Language	.86	1	.86	.05	.820	.001	.052	.056
Error	673.47	41	16.43					

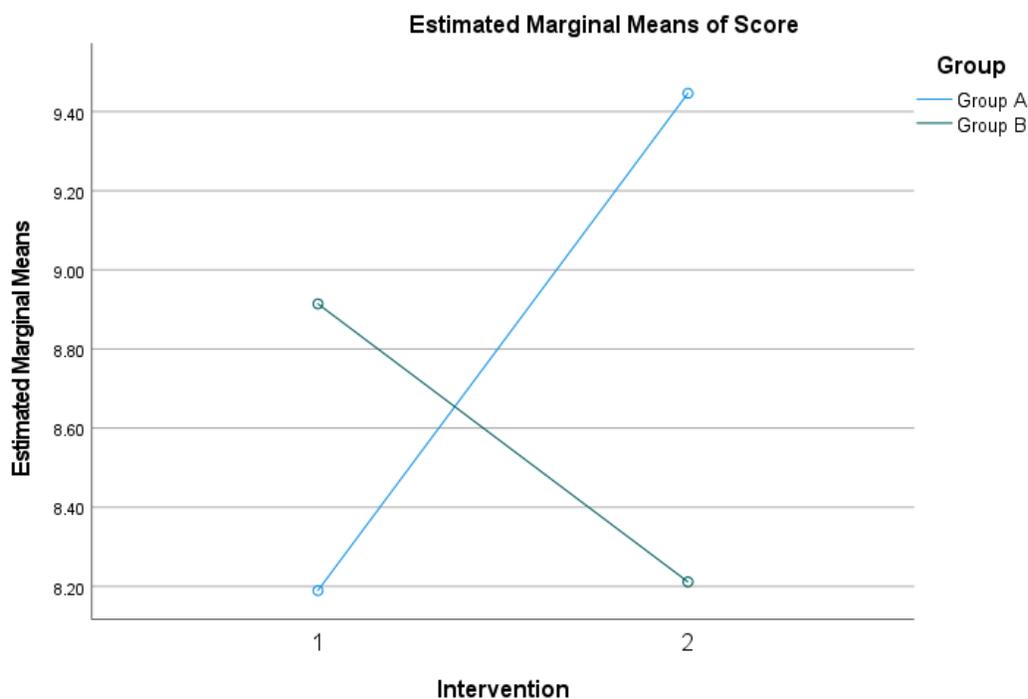
a. Computed using alpha = .05

Interaction Effects

Interaction effects for both between-subject and within-subject factors were examined and found between group and intervention ($F(1,41) = 32.53, p < .001$) with a large effect size ($\eta^2 = .44$). This was interpreted to mean that there were overall mean differences dependent upon the order of intervention, irrespective of test. However, interaction effects were also found among intervention, group, and test ($F(2,40) = 8.28, p < .001$) with a large effect size ($\eta^2 = .29$). This was interpreted to mean that differences between at least two tests and intervention were dependent upon group membership. As seen in Figure 19, Group A had higher overall means for the writing intervention than for the reading intervention, whereas Group B had higher overall means for the reading intervention.

Figure 19

Group and Intervention Interactions

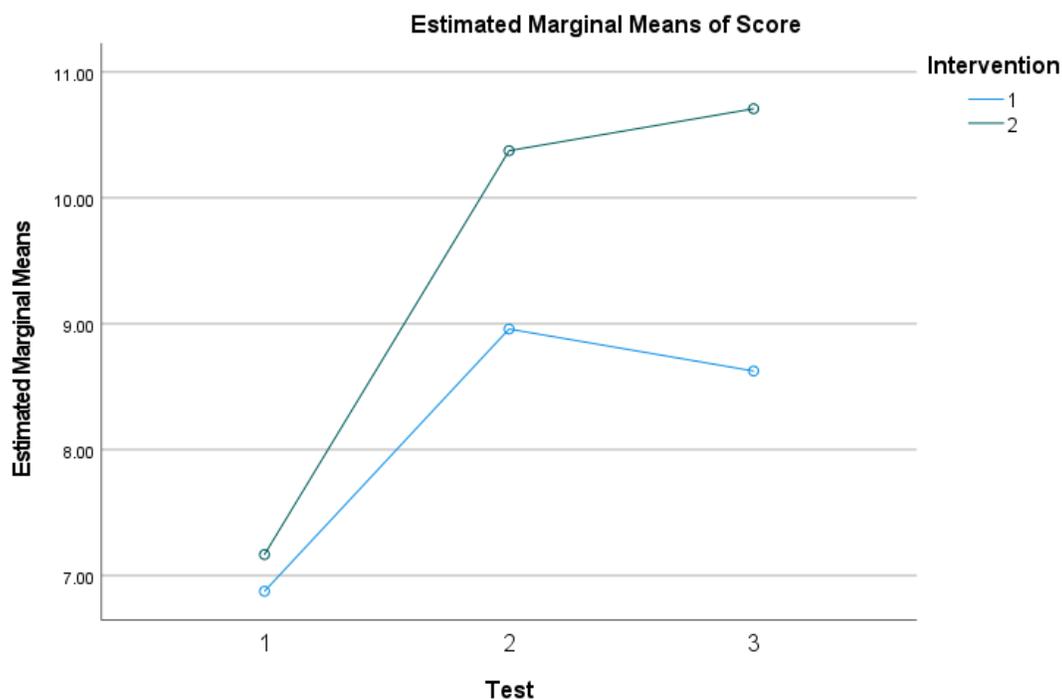


Group A

To untangle these effects and to examine where these differences lay, I conducted a repeated measures analysis for both groups independently. For Group A, there was significant difference for interventions between the pretest and the posttest ($p = .005$). Figure 20 demonstrates this difference. In short, members of Group A made higher gains with writing than they did with reading, and they were more likely to retain what they learned through writing in the delayed posttest. In contrast, Group A suffered significant loss during the delayed posttest of lexical sequences they acquired through reading alone ($p < .001$).

Figure 20

Group A Scores by Intervention and Test



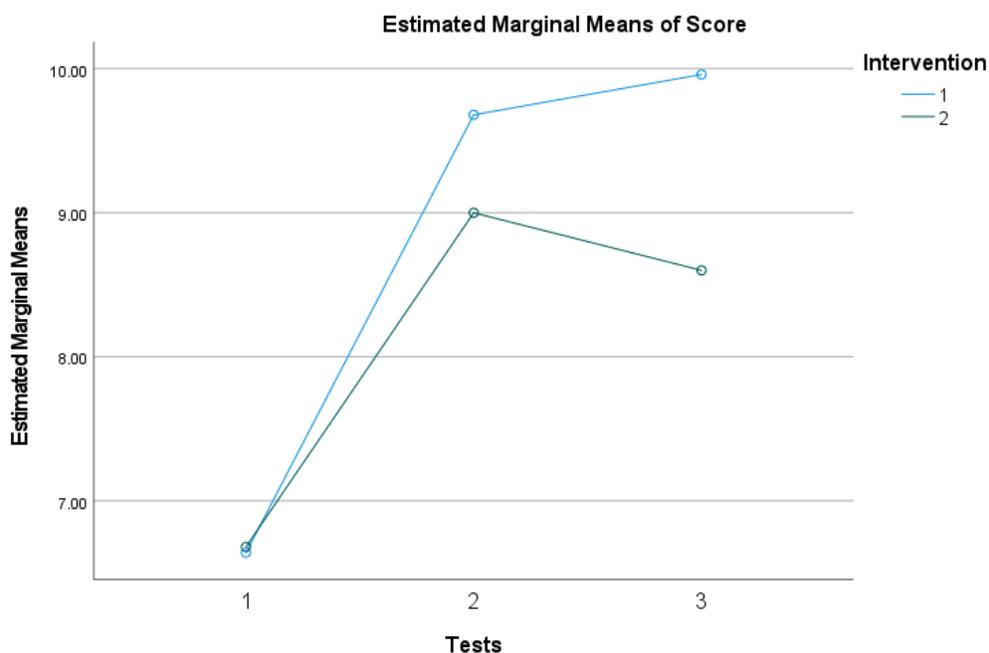
Group B

Group B also had significant gains for both interventions, although in contrast to Group A, the reading intervention was significantly more effective than the writing intervention ($F(1) =$

6.42, $p = .018$) with a large effect size of $\eta^2 = .211$. A pairwise comparison found significance between the pretest and delayed posttest ($p < .001$) and between the pretest and the posttest ($p < .001$), but not between the posttest and delayed posttest. This was interpreted to mean that for Group B, the reading intervention led to greater gains, although both interventions led to significant learning, and that there was no significant loss of learning for either intervention after three to four weeks.

Figure 21.

Group B Scores by Intervention and Test



The Role of Lexical Sequences

The finding that one group made greater gains during the reading intervention and the other group made greater gains with the writing intervention perplexed me. To gain a clearer understanding of this curious outcome, I considered the fact that lexical sequences used for reading or writing differed depending group number. For instance, Group A practiced *reading* of sequences 1-4; Group B practiced *writing* of these sequences. Group A practiced *writing* of

sequences 5-8, whereas Group B were exposed to these sequences through *reading*. I therefore wondered if perhaps the lexical sequences themselves lay behind this finding. In effect, both groups learned the *second* set of lexical sequences more readily, regardless of whether the intervention was reading or writing. If so, such a finding would imply that the sequences themselves are a more important factor than the intervention used to lead students to acquire them. Other questions arose as I considered the impact of individual sequences:

1. Did participants attempt to use the lexical sequences in their written responses, and if so, did they use them as expected?
2. Were participants more likely to practice some lexical sequences than others during the writing intervention?
3. Did any differences in usage among lexical sequences impact results?

I began to answer these questions by examining participant writing samples. Specifically, I examined patterns of how individual lexical sequences were used during writing tasks. A perusal of written responses revealed that participants were more likely to correctly use some lexical sequences than others during the writing intervention. Furthermore, in roughly 36% of cases, participants did not attempt to use the targeted sequences when responding to readings, even when they wrote a measured response in English. The written responses also revealed a difference in the accuracy of usage of the targeted sequences. In other words, even when participants did attempt to incorporate the lexical sequences in their writing during the writing intervention, they did not always use the targeted sequence as expected or with its intended meaning. For example, Table 12 shows that Group B, who wrote using sequences 1-4, were much more likely to use the lexical sequences in unexpected ways—whether incorrectly or

whether using an alternative definition for a phrase than the one modeled and defined through glosses in the readings.

To examine the usage of lexical sequences during the writing intervention, an index was created for the writing intervention results for each participant and the lexical sequence they used during the writing intervention. As is true for the subjective nature of determining usage in testing, determining correct usage during the written responses required the assistance of another individual to maintain the integrity of the study's validity. An assistant followed the same guidelines for determining whether usage was correct as he did for the assessments. When participants did not attempt to use a lexical sequence, a zero was assigned. When the participant attempted to use the lexical sequence but did not use it correctly or with the intended meaning, a score of 1 was assigned. Finally, a score of 3 was assigned when the participant used the lexical sequence as intended in the written response.

Table 8

Usage of Sequences During Writing Intervention by Group

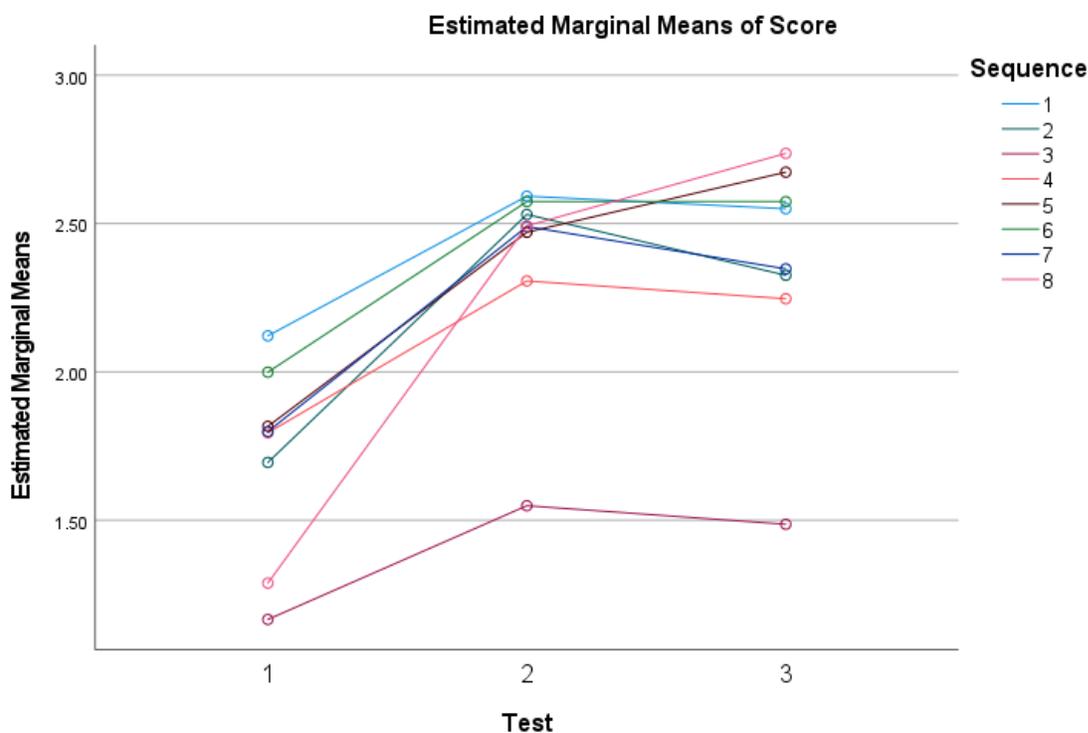
Group	Sequences	Total Sample	No Attempt	Unexpected Usage	Percent "Correct" (Expected Usage)
A	5-8	384	153	8	60%
B	1-4	400	138	51	66%

Lexical Sequence Usage Scores. As shown in Table 9, participants were more likely use some lexical sequences correctly than others during writing tasks.

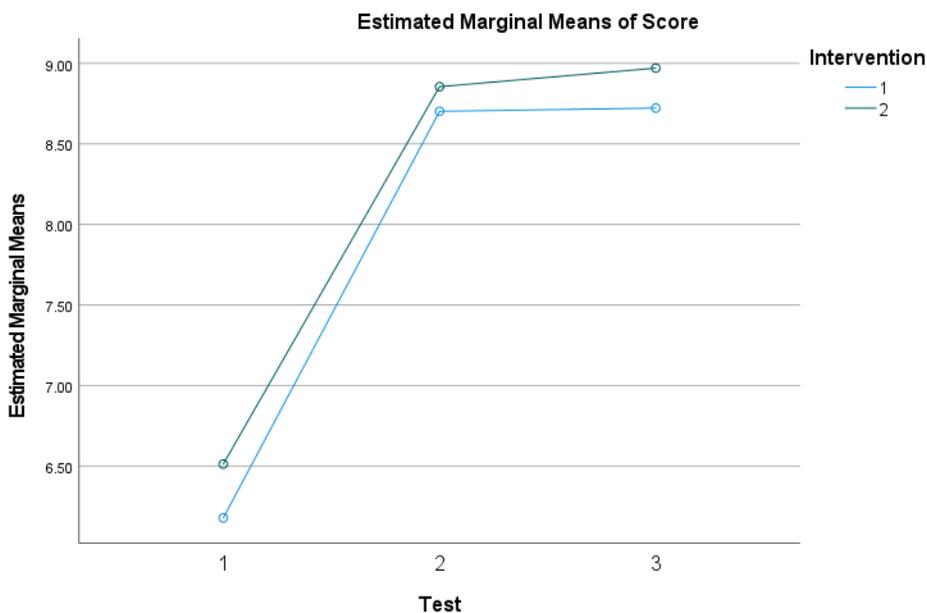
Table 9*Usage Scores by Lexical Sequence*

Lexical Sequence	Usage Score	Set Number
Lexical sequence 1 (<i>as a result</i>)	124	set A
Lexical sequence 2 (<i>in terms of</i>)	108	set A
Lexical sequence 3 (<i>the extent to which</i>)	89	set A
Lexical sequence 4 (<i>in addition to</i>)	120	set A
Lexical sequence 5 (<i>a number of</i>)	88	set B
Lexical sequence 6 (<i>as well as</i>)	127	set B
Lexical sequence 7 (<i>referred to as</i>)	74	set B
Lexical sequence 8 (<i>a wide range of</i>)	118	set B

To determine whether individual lexical sequences had an impact on test results, I conducted a repeated measures analysis with test and each of the eight lexical sequences as within-subject factors and with group number as the between-subject factor. The test showed that sequence was a significant factor on test results ($F(14, 34) = 4.97, p < .001$) with a large effect size ($\eta^2 = .67$). To determine which lexical sequence(s) had statistical significance from others, I made pairwise comparisons among sequences. The pairwise comparison demonstrated significance between Sequence 3 and all other sequences for both groups and for all tests. (See Fig. 17 for visual.)

Figure 22.*Test Scores by Lexical Sequence***Subsequent Repeated Measures Analysis**

As a consequence of this finding, Sequence 3—*the extent to which*—was considered an outlier and removed from the data set. A second repeated measures was then conducted in the same fashion, with intervention and test as within-subject factors and group, L1, and gender as between-subject factors. Scores were still found to be significant for both groups from pretest to posttest ($p < .001$) and from pretest to delayed posttest ($p < .001$). No losses were uncovered for the delayed posttest ($p < 1.000$). Overall results indicated a slight advantage of writing over reading (See Fig. 20), although this difference was not statistically significant between any two tests ($p = .171$).

Figure 23.*Intervention and Test Results, Outlier Sequence Removed*

No significance was found for group ($F(1) = .197, p = .660$), L1 ($F(1) = .409, p = .526$), or gender ($F(1) = .010, p = .920$). However, interaction effects were seen again for group and intervention ($F(1,14) = 187.00, p < .001$) and for intervention, group, and test ($F(1,14) = 14.96, p < .001$). The large F value and effect size for group and intervention interaction compared to when tests were added (e.g., $\eta^2 = .82$) led me to an interpretation that this difference may largely be a result of differences among lexical sequence knowledge as well as gains. In other words, lexical sequences with large posttest results also had greater pretest averages. Group A had much larger gains when asked to write (Sequence 4-8) and Group B had much larger gains when asked to read (Sequence 4-8). Again, this suggests that sequences 4-8 were more readily acquired, whether through reading or writing. This remained the case even when the easiest lexical sequence was removed from sequences 1-4. Of all sequences, Sequence 8, *a wide range of*, had the greatest gains, whether acquisition occurred through reading or writing, and participants

made even further gains for this lexical sequence during the delayed posttest. It should be noted that Sequence 8 is part of the second set, which both groups more easily acquired, regardless of intervention. Table 10 ranks the order of ease in learning each lexical sequence overall and by intervention.

Table 10

Lexical Sequence Gains, Highest to Lowest, by Intervention

Exposure	Lexical Sequence	Set	Gains Pretest-Posttest
Both Interventions			
	a wide range of	B	59
	in terms of	A	41
	referred to as	B	34
	a number of	B	32
	as well as	B	28
	in addition to	A	25
	as a result	A	23
	the extent to which	A	19
Reading Intervention			
	a wide range of	B	30
	referred to as	B	21
	in terms of	A	19
	a number of	B	15
	in addition to	A	14
	as a result	A	13
	as well as	B	10
	extent to which	A	4
Writing Intervention			
	a wide range of	B	29
	in terms of	A	22
	as well as	B	18
	a number of	B	17
	extent to which	A	15
	referred to as	B	13
	in addition to	A	11
	as a result	A	10

Summary of Quantitative Findings

Quantitative findings demonstrate that both interventions were effective in helping learners acquire vocabulary sequences, yet group membership determined which of the two interventions most effectively fostered this. An examination of the written responses of participants as part of the writing intervention led to questions about whether unexpected usage or failure to use lexical sequences may have impacted the individual's propensity to learn a given sequence. As a consequence, participant writing task responses were quantified. It was found that in many cases, students did not attempt to use the targeted lexical sequence during the writing intervention, and they often used them in unexpected ways when they did. I also discovered that some lexical sequences were more likely to be acquired than others, with *the extent to which* the least likely to be acquired and the least likely to be used as expected during the writing tasks. A *wide range of* was by far the most easily acquired and retained lexical sequence while also having a comparatively high usage score.

Qualitative Findings

It seems clear that as individuals, the interviewees have had unique and complex life experiences, each with their own set of learning influences, both positive and negative. It is not my intention to essentialize participants' social identities, to generalize about specific participants, or to propose that absolutes can be established based on interviews with 12 high school teenagers. However, I believe that my participants' stories reveal basic truths about how interactions with others at home and at school may have far reaching consequences, not only on learning outcomes, but also on their social lives and emotional health.

As described in Chapter 3, data in the qualitative strand of the study were primarily comprised of interviews from 12 participants and written responses to readings examined as part

of the quantitative strand. The interview data were found to align with four constructed themes. The first theme regards literacy practices, such as participants' perceptions of the intervention and how it has impacted literacy habits. The remaining three themes—claimed and unclaimed responsibilities, bullying as a life-altering event, and contradictory intersubjective experiences—describe the impact of work and relationships on the educational experiences and lives of the participants, with interpretation informed through the lens of Núñez's (2013) multilevel model of intersectionality and Anthias's (2012) levels of analysis. The levels of analysis and Núñez's model are described in detail in the methodology chapter.

Theme 1: Influences on Literacy Practices

Influences on the participants' literacy practices arose in part through discussions of the difficulty of vocabulary in learning English and the longstanding impact of the 2020 COVID-19 quarantine. Discussions of the interventions I asked participants to undergo revealed other influences on literacy practices as well, including text interest, the presence of glossing, the repetition of targeted vocabulary, and the requirement to practice using the lexical sequences in writing.

A common refrain from participants was the identification of vocabulary as the primary reason for difficulties in reading English. When asked directly what makes reading difficult, all but one of the 12 participants mentioned vocabulary. As Paper Lamp explained, "words can be challenging? to understand and read." Sophie and Harold also deemed vocabulary a primary cause of reading difficulty, but they attributed the difficulty to differences between English and Spanish. Sophia attributed the difficulty in part to the confusion that arises from false cognates, whereas Harold emphasized that unlike Spanish, difficulty with vocabulary sometimes relates to the nonphonemic spelling of some English words:

Harold cause you see a word and its very different to spelling from when it is spoken ... also you hear something some (.) English speaker saying something i:t it would really be different to write it

All participants responded positively to questions about the study tasks, often revealing their novel aspects as helpful in aiding reading, although the individual aspects of the study that participants found helpful varied from student to student. For example, Apollo mentioned the use of phrases instead of single vocabulary words as helpful to his learning, whereas three interviewees—John Doe, Wolf, and Jenny—mentioned the glossing aspect as helpful to completing the reading tasks:

Jenny I was just reading it (.) I didn't know what that means (.) it had like (.) the meaning in it? yeah and that REALLY help me ... it was gre:at

Other reasons included how the structured nature of the reading tasks led to reading as habit forming:

Harold I'm getting used to reading in ENGLISH ... twice a week (.) and then (2.0) and the way that I <had to do it> twice a week it make me to think that I have to read more (2.0) li:ke every week even if it not (.) for this

and the repetitive nature of exposure to the phrases:

Jenny I thought it was helpful for me to learn more ... it was repeate:d? and its like different words than I use ... like bigger cause like (.) I didn't know like (.) the words? so it really: help me to make more sentence

One participant, Stephanie, also described the cognitive difficulty in completing the writing task as helpful in her learning:

Stephanie like (.) like (.) I heard of the words but I never type them (.) it took me a lot of time? because I had to think about it (.) then (.) like (.) write it? like over again? they were not hard:?

Four interviewees (33%) discussed the relatability of the texts as an important element in finding interest in the tasks and in having an incentive for attempting to understand the text. For example, Harold explained that he enjoyed reading “Black Lives Matter” because the topic was familiar to him, explaining that “everyone knows about it.” Jenny, Sophie, and Stephanie all explained that they enjoyed the texts that were relevant to them; Sophie explained that she reread the text to ensure her understanding of it because she found the topic relevant to her. Jenny found the immigration reading engaging “cau:se immigration kind of talks about MY family too,” while for Stephanie, it was a reading about refugees: “the refugees (.) I liked that one cause I’m a refugee too you know?” Indeed, Sophia and Wolf found some topics so relevant that they used them as springboards to discuss them with me in depth. Both participants seemed compelled to explain how the text related to their lives personally and to share their opinions on the topics. For Sophie, this reading was Gen Z, a topic she discussed at length when she relayed her views about the readings and a topic she later revisited when she discussed her family. Sophie explained that older generations do not understand her own, a truth she believed to be reflected by the hurtful statements her own mother makes: “She says if I lose weight I will be prettier,” as well as those of her friend’s mother, “her mom calls her fat.” For Wolf, the topic was social justice, which she discussed at length:

Wolf you know, I just think there's not social fairness? because if people choose how you live your life and you have to work hard sometimes people (.) a lot of people work hard for their dream and the other people just don't work that hard for it? ... in ↑Africa (.) I saw a lot of kid that didn't go to school (.) that didn't have the opportunity to stud::y? so they just have to sell thing? they have to work hard? to sell thin::g? to have money for their family? (.) I just think that we have the right to go to school and them not having that right and other people have better school than that (.) they may the think that's not ↑fair.

Here, Wolf connects the reading about social justice to her own life experiences in Africa, explaining that children in Africa do not get the same educational opportunities as American children and noting that it does not seem fair to her.

Participants often discussed strategies they take when encountering unfamiliar words, whether in describing the approaches they took when undertaking the intervention tasks or in reflective descriptions of what makes reading difficult for them. Nearly half of interviewees stated that they often look up words they do not know. Other strategies mentioned include the use of context clues and trying to pronounce the word aloud.

The data also revealed that the quarantine had a positive impact on learning for two interviewees. For instance, Luis turned to schoolwork when he could not visit friends, and Jenny explained that while under quarantine, she discovered a newfound interest in reading. She explained that she now reads in English every day—both novels and nonfiction—and wants to become a writer when she grows up.

Jenny: you know how we had quarantine? before? so like (.) I was like
 (.) oh wait (.) this is very interesting (.) I started reading books
 during the quarantine ... I have to read like (.) every day at
 least a little bit

The most recurring influence on literacy practices that participants shared was the identification of vocabulary as the most salient influence on reading difficulty. However, strategies for overcoming reading difficulties and the underlying reasons attributed to them differed from participant to participant, varying from L1 semantic distance from English to the distance between phonetics and spelling that is characteristic of English. In a similar vein, participants stated that they found the intervention helpful, often describing aspects of the intervention tasks as something they had not previously been asked to do. The participants also mentioned specific texts that they found relevant to their experiences and identities as powerful motivators for comprehension. Other novel aspects of the study that participants discussed include the glossing of easy definitions within the text and the regularity of required reading that, for two participants, established reading as a daily habit. One surprising finding was the report from two participants that the quarantine inadvertently fostered positive changes in reading and learning for them that persist today.

Theme 2: Claimed and Unclaimed Responsibilities

The impact of responsibilities on the lives of participants concerned work that the participants have outside of school, whether paid jobs or responsibilities at home. Nearly half of the interviewees hold jobs, with Apollo and Paper Lamp working over 20 hours per week. Apollo explained that he holds four part-time positions (neighborhood babysitting, food delivery, food preparation, and housekeeping), working eight hours per day on average in addition to

attending school nearly every day. However, participants did not always *claim* these responsibilities. For example, when asked directly if they had jobs, two interviewees—Harold and Paper Lamp—denied having jobs, yet when asked to describe a typical day after school, they discussed work they conduct from home: Paper Lamp works 4–6 hours per day as a nail tech, whereas Harold provides support services for his father’s at-home business by interacting with customers on Ebay.

Unclaimed responsibilities were also apparent in how participants discussed childcare. Three participants take care of their siblings when they return home from school—one provides care to siblings until early morning—but none of the interviewees described childcare as a job. Both male and female interviewees often described housekeeping duties as a family event. For instance, they often described chores performed by using “we” or positioning them as family activities in other ways (i.e., “I help my mom with the dishes,” or “we make dinner”). In contrast to chores such as cleaning one’s room, caring for siblings was largely a gendered activity. Sophie described both the impact on schooling that caring for her younger sister has on her as well as its gendered aspect:

Sophia: and studying with her? is really hard (.) my: (.) brother: because
 he's the only boy? of the house? he doesn't really have
 responsibilities (.) he like (.) plays his Roblox? and (.) is on the
 phone with his girlfriend? and I will be like (.) can you HELP me?

In summary, the theme of claimed and unclaimed responsibilities was interpreted from responsibilities that participants discussed, including both paid jobs and household chores. Participants confined the term *job* to include only activities that required them to leave their homes to undertake, even when they received payment for services. One

participant described her responsibility for childcare as one assigned to her because she is a female.

Theme 3: Contrasting Intersubjective Experiences

The impact of relationships on the lives of the participants were seen in how they discussed their families and teachers, sometimes revealing contradictions. One of the most uniformly recurring ideas, however, was the overwhelming importance of family to the participants. Many participants stated that they spend nearly all of their free time with their families, whether having cookouts on the weekends, attending church together, playing sports together, or visiting other families to go to parties together. For nearly all of these teenagers, family is everything. Paper Lamp explained that parents are there for support and to help their children “to be better people.” Nearly all interviewees viewed family members as their life coaches, most important teachers, and primary support systems. GenZ and Harold provide two such examples. GenZ’s account demonstrates that he has meaningful relationships with all members of his family,

GenZ: my family:? they help me a lot with everything? they (.) they teach me EVERYTHING I know (2.0) my dad teach me a lot about working? working on cars (.) my mom she helps me a lot with my life basically? (.) and my brothers (3.0) my siblings they help me like play sports: play video ga:mes (2.0) that’s all.

whereas Harold explained that he has special relationships with his parents, who play different roles in his life but always support his decisions:

Harold: I am very close to my family ... I’m very close with them so: I don’t have a job? but sometimes I: just like my dad has jobs out of

town? so sometimes I go with him ...also my mom (.) li:ke (3.0)
I'm really_close with her

All interviewees who shared their parents' perceptions about education described parents as having strong, positive views about school. For instance Harold, Stephanie, John Doe, and GenZ talked about how their parents support their goal of attending college. Harold's discussion demonstrates that many participants' parents place value on a good education:

Harold: OH they think that school is the best thing that you can do in your
LIFE cause um: that's how you're going to be when you grow up
(.) school teach you like (.) the things (2.0) how to be in LIFE

The importance of family was also apparent in narratives that describe the consequences of interruptions to families, seen in narratives from Sophia and Alex. Alex explained how worrying about his little brother impacted him after his father left the family:

Alex: we:ll he was the one who would cause a struggle for me and my
mom cause mostly <cause he's my little brother you know?> a:nd
I'd get worried

Alex continued, explaining that the absence of his father also led him to drug use and in turn, a difficult relationship with his mother that continues today, despite the fact that he no longer uses drugs. Sophie described the unfortunate consequences to her life when her mother's lack of support for her sexuality led her to gravitate towards friends that she described as bad influences:

Sophie: I would sneak out a little and then my friends (.) they were super
influencers (.) so like they would tell me to smoke and I would do an to
drink and I would do it too like (.) let's go

Relationships and perceptions about teaching were also abundant within the qualitative data. Most participants described their teachers as having an impact on their learning, whether positive or negative. Nearly all participants described their teachers as “friendly” or “helpful.” On the other hand, representational influences are reflected in teaching practices borne from implicit beliefs about bilingual learners (Denessen et al., 2022). Consequently, implicit beliefs were also apparent in the more nuanced ways that participants critiqued their teachers, which were often veiled within contradictory statements.

Students were prompted to discuss their schools and teachers, with no specific queries about the quality of teaching the students receive. Interviewees overwhelmingly expressed positive opinions about their teachers’ dispositions. However, among those who discussed the quality of teaching they receive, the results were less decisive, arguably following a gendered pattern. For instance, female interviewees Stephanie, Wolf, and Paper Lamp discussed positive learning experiences in detail; Stephanie and Paper Lamp described the scaffolding techniques their teachers use. For instance, when asked to clarify how teachers are helpful, Paper Lamp replied:

Paper Lamp: try to go on the easier side? and try to make us understand the easier sides? and when we are used to the easier side of like (.) the learning um: try to go to the harder

Wolf and Paper Lamp offered some critiques of their classroom experiences as well, but they couched their criticism in positive commentary about their teachers. For points of comparison, see Fig. 23. Criticism appears in red, while positive comments appear in blue.

Figure 24

Gendered Aspect of Teacher Critique

Wolf (Female)	Paper Lamp (Female)
<p>IE: and my teacher to help me?</p> <p>IE: the ESOL teachers? umm. I think they're great? cause they tryna help you learn the language at the same time? which is I think is something really good? to help people learn? at the same time trying to get better in their English? think is something very helpfu:l</p> <p>IE: um::m (2.0) there is some ↑classes: that ... you just have to aware yourself? ... it kind of make it (.) kind of sometime hard to: to understand them? um::m sometime just reading <I just don't understand it sometime></p> <p>IE: um. an example? um my teachers? they always help me: they give me dictionary? when I don't understand a word (.) an:d they help me try improve</p>	<p>IE: mm (2.0) I feel lik:e. (3.0) they do help me (.) but I just don't understand what they mean (.) when they: say some of the stuff</p> <p>IE: when they are teaching I don't really understand</p> <p>IE: I just think that (.) they don't explain and they just throw in the words</p> <p>IE: mm: I think that they (3.0) teach me good</p> <p>IE: learning (.) just to try to help us understand what (.) what is easy and what's hard.</p>

<p>my English? when I need help (.) they</p> <p>help me to learn English more better</p> <p>IE: they just help me, you know?</p>	
---	--

Early in their interviews, Wolf and Paper Lamp both call their teachers *helpful*. Later, Wolf reiterates this sentiment by giving a concrete example in her teachers' use of content-based instruction, saying that it helps her learn English. Later in the interview, Wolf somewhat contradicts this statement by implying that because she cannot understand her teachers, she sometimes has to try to teach herself. Likewise, Paper Lamp immediately follows her initial statement that her teachers help her with statements that she cannot understand, and both participants end by reiterating that their teachers are helpful: Wolf says, "they just help me, you know?" and Paper Lamp says, "I think that they teach me good."

In contrast, male participants were more direct in critiquing their teachers, as the following excerpt from Apollo demonstrates:

Apollo I want to say:: um: want to say about a teacher but it's not (.) but is not personal? ... [redacted teacher name] does not teach good

In contrast to female participants who were reluctant to admit their learning experiences were not always perfect, Apollo shared his teacher's name, which was redacted. Apollo is one of two participants who seemed particularly impacted by teaching practices that minimized expectations for them. For example, he opined that "there are teachers that are not very:: they don't put any effort to make us study." He also explained that his teachers sometimes treat him as though he is stupid. John Doe expressed a similar sentiment when he described some of the assignments he is asked to complete as busywork:

John Doe: it's just an extra grade (.) it's like an extra grade (.) um: (2.0) I
don't want to say it doesn't help me (.) but:

IR: it doesn't help you

John Doe: no

Intersubjective experiences—interactions with individuals that may impact educational experiences of learners—were most salient in how participants discussed their parents. Nearly all participants described warm, caring home environments where caregivers offer social experiences, emotional support, and educational support; when family bonds were severed, participants engaged in risky behaviors that caused their learning to suffer. Experiences were often described by female participants in contradictory ways: On the surface, teachers were kind and instructive individuals who always helped them, but when asked to expand on their experiences, female participants revealed that their learning experiences were not always effective. Female participants tamped their criticisms heavily with positive descriptions of their teachers. In contrast, male participants were more direct in exposing their teachers' shortcomings.

Theme 4: The Life-Altering Impact of Bullying

Stories of bullying were unfortunately not uncommon, and because instances of bullying often co-occur with discriminatory practice (Raveche et al., 2014), participants' accounts of bullying revealed that they were most often correlated to the identify factors of gender, race, and bilingual learner status. In the case of Apollo, identity factors included a perceived nonnative accent coupled with a communication disorder. The only two interviewees whose bullying continues—Apollo and Stephanie—offered strikingly similar accounts of their experiences. For example, both denied being bullied at least once during the interview but later admitted they had

been or continue to be bullied. Both participants also minimized the experience in some way. Apollo did so by emphasizing that no physical abuse occurred, whereas Stephanie did so by denying any racial motivation behind being called a “bad word” (she explained she could not say the word aloud) because of her race. I asked if the word was really, really bad and she replied that it was. Other than Apollo, no males reported having been bullied. Alex explained that although personally, he is treated fairly by the students in the school, he has witnessed the bullying of others because of their socioeconomic status:

Alex: they just make fun of them (.) like (.) for no reason (.) I don't think that the
 (.) the students that (.) that are being bullied deserves that ... they make
 fun of like their clothes and their shoes ... and they can't afford them (.)
 those types of things

One of only two Black female participants, Stephanie's egregious treatment by other students, which may have included racial slurs, provides as snapshot of the experiences of a Black female ELL. In the excerpt below, Stephanie discusses differential treatment she has received because of how these social identities intersect:

Stephanie: sometimes (.) um: it's like (.) bad (.) bad words sometimes (2.0) and
 they just put it in like (.) sentences (.) and I just be there ...
 like (.) most of the time they: I can't say it (.) like most of the time
 they call you a African (.) like ((sigh)) like (4.0) ... most people get
 caught at that (.) they judge you because of where you came from

As a White male participant who spoke with a near natively accent, Harold described a starkly different schooling experience and even expressed disbelief at the notion that anyone

would be treated differently at his school. When the question was posed, he interrupted me to declare, “We are treated the same way cause ALL the students matter.”

Clearly, there is great variation in how the interviewees experience social life at school. Males were less likely to experience bullying as were White students and those with natively like accents. The most apparent impact of how school experiences impact language learning was evidenced in descriptions of how bullying experiences continue to adversely affect the interviewees’ language learning, even when several years have passed since the bullying took place. Three female participants—Paper Lamp, Millie, and Stephanie—revealed that one long-lasting impact of being bullied as a bilingual learner is an unwillingness to engage with other students at school. Paper Lamp explained that she avoids talking to others, stating, “I want to be by myself.” Stephanie explained that because she doesn’t talk much, other students think that she doesn’t understand English, which leads them to talk openly about her.

As noted in Chapter 3, half of the interviewees identified English as their first language, and one interviewee, Alex, rejected the English language learner label until I sought clarification, which led him to reluctantly agree that he qualified.

This is not to say all interviewees passively accept how others characterize them. For instance, Sophia confidently rejected the notion that the homophobia of others at school could negatively affect her, instead positioning homophobic students as unfortunate victims of their own prejudice.

Sophia: First of all (.) it's okay (.) <I really don't care> (.) <it doesn't affect me> (.) it does affect them (.) the fact that they don't have a gay friend (.) I really don't care because there are straight people in my life (.) and it's like (.) I have

people that support me and I don't need you to support me because I already have people that do.

Likewise, when discussing her mother's judgmental behavior towards her and the Church's position on her sexuality, Sophia positioned herself as being in control of choosing whether to spend time with her mother and implied that it is the church, and not her sexuality, that is wrong.

Like many other of my findings, the bullying experiences that participants shared followed a gendered pattern, with females more likely to be subjected to bullying and its aftermath than males. Bullying arose as a product of social identities, including bilingual learner status, presence of an accent, race, and sexual orientation. Bullying that occurred years earlier continues its impact today, often leading the students to avoid interactions with others or to avoid speaking English altogether. One participant shared a powerful narrative that has allowed her to challenge perpetrators of bullying in her life.

Summary of Qualitative Findings

Qualitative results of interview data taken from 12 participants fell into four major themes, which include influences on literacy practices, claimed and unclaimed responsibilities, the life-altering impacts of bullying, and the often contradictory accounts of intersubjective experiences. Data revealed that connections participants made to the reading topics often determined how invested the participants were in the readings, and in turn, whether or not the participants made an effort to reread the text for meaning. Participants often commented on the novel approach to vocabulary learning that the interventions offered, such as the inclusion of interlinear glosses within the text, the use of academic vocabulary, the repetitive and dispersed exposure to the phrases over multiple readings, and the encouragement to practice using

vocabulary after contextual exposure to it. Nevertheless, the most commonly occurring refrain within this theme was the finding that vocabulary is the primary cause of difficulty participants face in learning English as an additional language. Two interviewees stated that the quarantine led them to establish reading and study habits that continue today; one explained that the study intervention led him to improved reading habits.

Family was found to be of central importance to those interviewed, who stated that they spend nearly all of their free time with their families and depend on their families heavily for emotional support, learning, and guidance for future educational decisions. Data suggested that the parents of interviewees offer support to their children in graduating from high school and share their aspirations for postsecondary school. Data also revealed gendered patterns in the types of work interviewees undertook at home, with childcare being primarily a female occupation but with both genders often undertaking housekeeping chores as a family activity. Differentiation was found in how “job” was defined by the interviewees; several interviewees who stated they did not have a job later described working, yet payment for services did not seem to be the defining characteristic. Finally, critiques of teaching practices seemed to follow a curious gendered pattern, with females more likely to tiptoe around critiques of their teachers and males more likely to offer direct criticism, even calling a teacher out by name in one case.

While data offered some evidence that implicit beliefs of teachers may impact the learning experiences of a few interviewees, a much more prevalent influence was seen in interviewees’ descriptions of their own and other’s bullying. Students had been bullied because of their socioeconomic status, race, status as a bilingual learner, for having a nonnative accent, for having a speech impediment, and for combination of these social identities. The longstanding impact of bullying, even among interviewees who had not been bullied for several years, became

apparent as numerous interviewees described their reluctance to speak to others even now for fear of being ridiculed by other students. The narratives of one interviewee, Sophia, provided important data about the experiences of a gay Latinx teenager, who offered descriptions of the control she wields over homophobic students and her mother by repositioning the locus of control onto herself. Her narratives also revealed an understanding of her identity as representative of a generation whose values differ from that of their parents.

Mixed Methods Integration

By combining the results of both the quantitative and qualitative analysis, I was able to identify factors that influence language learning generally within the multilevel model of intersectionality. These factors appear in Figure 25.

Figure 25.*Mixed Methods Findings*

Mixed Methods Research Question: What meta-inferences can be drawn from the quantitative and qualitative strands of the study about factors affecting bilingual learners' learning?			
Data	Finding		Mixed Methods Inference
QUAN	Students acquired lexical sequence depth of knowledge through both interventions with large effect sizes.	→	<p>Bilingual learners can readily acquire depth of vocabulary knowledge through reading and writing. Text relevance, innerlinear glosses, and dispersion may enhance bilingual learner vocabulary acquisition through reading. However, practicing newly acquired vocabulary should be accompanied by corrective feedback.</p> <p>Asking students to routinely read interesting texts will foster habits that they may continue at home.</p> <p>Some lexical sequences are more readily acquired than others.</p> <p>Employment and gendered responsibilities at home are gendered and may impact learning.</p> <p>Bullying of bilingual learners (based on intersections of identity factors including bilingual learner status, perceived non-nativelike accent, race, gender, L1, and disability) leads bilingual learners to avoid communicating in English with others.</p> <p>Implicit beliefs of ESOL teachers may detrimentally impact learning for bilingual learners.</p> <p>LGBTQ bilingual learner students may find empowerment by shifting the locus of control to themselves.</p> <p>Families provide an important source of social life, support for educational aspirations, and emotional support for bilingual learner students.</p>
QUAN QUAL	Participants more readily learned lexical sequences that they used correctly during the writing intervention	→	
QUAL	Theme 1: Influences on Literacy Practices	→	
QUAN	The targeted lexical sequence outweighed the impact of the intervention	→	
QUAL	Theme 2: Claimed and Unclaimed Responsibilities	→	
QUAL	Theme 3: Bullying as a Life Altering Event	→	
QUAL	Theme 4: Contrasting Intersubjective Experiences	→	

In combining the findings from both the quantitative and qualitative strand, I inferred that bilingual learner students can acquire lexical sequences through reading and through writing with

repeated, dispersed exposure in multiple reading contexts. This is especially true when students find the readings relevant to them personally. The reading and writing interventions were found to be a successful way for bilingual learners to acquire depth in knowledge of lexical sequences common to academic writing, and these gains were largely retained three to four weeks after the intervention. The factors found most salient in influencing intervention results among participants were the targeted lexical sequence being learned, the degree to which participants found the text relatable, and for some sequences, the degree to which participants used the lexical sequence as expected when responding to texts during the writing intervention. Interview data suggested that the quarantine had a positive impact on learning for a few participants, who used the time to establish reading and studying habits that continue today. Furthermore, one participant explained that the routine reading that characterized the intervention tasks led him to adopt reading as a daily habit.

Although L1, gender, and type of intervention were not found to have a statistically significant impact on intervention results, L1 and gender impacted language learning through their function within other factors found to be important. These factors include bullying and responsibilities of family and work. For instance, gender was found to influence how participants critiqued their learning experiences as well as home responsibilities. Participants described most chores as a family activity, but the gendered activity of childcare had a direct impact on schoolwork as reported by one interviewee.

Bullying was found to be a function of numerous social identities and their intersections. These included bilingual learner status, race, L1, gender, and disability. Earlier experiences of being bullied continued to impact language learning experiences for some participants through a continued reluctance to engage with other students at school, and this impact had a gendered

aspect among the interviewees, with females more likely to report bullying and to isolate themselves from others. The narratives of one interviewee provides insight into how one lesbian bilingual learner facing homophobia at school and at home found empowerment by positioning herself in the locus of control and others as victims of ignorance and intolerance. In most cases, participants occupied their free time with family activities. Parents provided knowledge and emotional support to their children and supported their educational aspirations.

5 DISCUSSION

Through integrative mixed methods, this study sought to determine whether high school bilingual learners could acquire and retain knowledge of lexical sequences through two interventions: one requiring participants to read short informational texts seeded with targeted lexical sequences, and a second requiring participants to respond to these readings using the targeted sequences in writing. The study compared the two interventions to learn whether one would prove more effective than the other while also measuring the impact of L1 and gender on results. The study also aimed to identify factors that have influenced the learning experiences of bilingual learner participants. In this chapter, I discuss each major finding from the mixed methods analysis and its implications vis-à-vis relevant research. Next, I describe the study limitations and make recommendations for policy, teaching practices, and further research.

Factors Influencing Literacy Practices

Several findings concern literacy practices and factors that influence study intervention results. Vocabulary has been shown to be an important predictor of both reading and writing proficiencies (Karakoç & Köse, 2017; Sarbazi et al., 2021; Pilar & Llach, 2007; Wright & Cervetti, 2016; Wu, et al., 2021), and findings of the qualitative strand of my study align with research holding that the reverse is also true: In keeping with a substantial body of vocabulary studies in both bilingual learner and native-speaking contexts (e.g., Kwon, 2009; Laufer, 1997; Salam, 2021), both male and female participants of my study overwhelmingly reported that vocabulary was the greatest obstacle in reading comprehension, regardless of their first language. Vocabulary knowledge is associated with educational success (Fernald et al., 2012; Hoff, 2003), and the advantages of a hearty vocabulary extend into realms of authority that grant social privileges to some (Hyland, 2011). Therefore, it is difficult to overstate the implication that

effective vocabulary instruction techniques—as described below—are important in both bilingual learner and in native-speaker contexts.

In accordance with entrenchment and usage-based theories of language learning, my findings also confirm that repeated, dispersed exposure to target lexical sequences through reading short informational texts leads to increased depth of knowledge of those lexical sequences among bilingual learners, with little attrition after three to four weeks. The fundamental finding that reading enhances vocabulary is secure within entrenchment theory (Blumenthal-Dramé, 2012) and usage-based theory (Ellis, 2003) as well as a plethora of studies, from Nagy et al.'s (1987) foundations to Yeatman et al.'s (2022) groundbreaking attempts to understand how learning vocabulary through reading impacts our neural circuits.

Numerous theories—namely the output hypothesis (Swain & Lapkin, 1995), Robinson's cognition hypothesis (2001), Laufer and Hulstijn's (2001) involvement load hypothesis, Long's (1981) interaction hypothesis, and Nation and Webb's (2011) technique feature analysis—have proposed that meaning-focused activities will lead to cognitive processing that better facilitates vocabulary acquisition than exposure alone. In other words, these theories suggest that practiced usage of targeted sequences should lead to greater gains than exposure alone through reading. However, in my own study, this was found to be the case *only for Group A*. Finding an advantage of reading over writing for Group B contrasts sharply with nearly all other studies that aver an advantage of reading and writing over reading alone for the sake of vocabulary acquisition (e.g., Bai, 2016; Gohar et al., 2018; Hazrat, 2015; Hulstijn & Laufer, 2001; Pichette et al., 2012). On the other hand, most of these studies defined vocabulary as single words, whereas in my study, vocabulary was defined as lexical sequences.

I propose three potential explanations for the quantitative finding that participants in Group B were more likely to acquire vocabulary through reading alone. As discussed in Chapter 4, a secondary quantitative analysis measured the impact of the individual lexical sequences themselves, finding that one set of lexical sequences (set B, which included the sequences *a number of*, *as well as*, *referred to as*, and *a wide range of*) was more readily acquired than set A, regardless of whether the sequences were incorporated into the reading or writing intervention. Therefore, one alternative explanation for the finding may be that the impact of the ease in learning the first set of lexical sequences outweighed the potential impact of the intervention difference; *it may simply be that some lexical sequences are more easily learned than others.* (See Table 10 for a breakdown of ease in learning by lexical sequence and intervention.)

Entrenchment theorists have proposed several potential factors that may influence the degree of strength for the entrenchment of a lexical sequence, including aspects of the lexical sequence itself. For instance, lexical sequences representing more concrete concepts may lead to stronger entrenchment with fewer exposures (Hilpert & Diessel, 2017) as might lexical sequences that are less complex and shorter (Blumenthal-Dramé, 2012). The two most difficult sequences for participants, *the extent to which* and *in terms of* may qualify here; although none of the sequences in my study can be described as semantically concrete, these two sequences from set A rely on sentence frames with multiple slots. As discussed in chapter one, phrases held in the mind as fixed units—those that can be replaced with an alternative word, for example—are more quickly processed than sequences that are part of a greater frame (Schmid, 2017). *In terms of* requires not only knowledge that a noun clause must follow it; it also requires knowledge that the entire prepositional phrase should be followed by an independent clause. *The extent to which* not only requires similar complex structural knowledge as *in terms of*, but also among all eight

lexical sequences in my study, it happens to be the sequence with the second lowest frequency in a Google Scholar search. (See Table 1.) Finally, pretest results revealed that participants had fewer encounters with the lexical sequence *the extent to which* before exposure to the study tasks, even though it may have required more exposure than sequences that are more fixed in nature.

A second explanation may be that the individuals in my study became accustomed to the intervention process during the first four weeks, in turn leading them to more ease in taking part in the intervention during the second four weeks, when the lexical sequences were more readily acquired, regardless of intervention.

A third explanation lies within research examining the testing effect. Having arisen from cognitive research on memory retrieval, the testing effect refers to a finding offered by a large and diverse body of research: When students are asked to learn content through practice testing as a method of retrieval practice, they are more successful at learning the content than when rereading or reviewing content (Agarwal et al. 2012). The testing effect is the mechanism that makes flash cards an effective method of studying (Greving et al., 2022), for instance.

The testing effect generally requires learners to engage in free or cued recall of text for comprehension or asks learners to take multiple-choice or true/false tests. Following practice testing, learners are provided with the correct answers (Eisenkraemer et al., 2013). Although my study tasks differed from those typically required in research examining the testing effect, the writing intervention in my study asked participants to complete a task requiring cognitive elaboration similar to practice testing. After exposure to the targeted lexical sequences through reading, the writing intervention required participants to practice retrieval of content to which

they had just been exposed (i.e., the targeted lexical sequences). In short, both the writing intervention and practice testing require memory retrieval to spur learning.

Much recent research on the testing effect has led to speculation that for efficacy, students must be provided with feedback after practice testing (e.g., Agarwal et al., 2012; McLaughlin & Coderre, 2015; Ruiz-Martin & Bybee, 2022; Van den Broek et al., 2016). In examining cognitive theories alongside fMRI studies measuring the impact of the testing effect, Van den Broek et al. (2016) found that when participants were provided with immediate feedback after an incorrect practice test answer, extra brain activity occurred—but *only if that participant remembered the correct answer later*. Consequently, Van den Broek proposes that making comparisons between prior answers and correct answers as feedback creates a new learning context that “enriches memory representation” (p. 63).

Van den Broek’s (2016) findings imply that individuals not only need corrective feedback when practicing content retrieval to learn; they also need to engage with their wrong answers. For the greatest and least acquired lexical sequences among my participants, my study found an alignment between usage score (whether the participant attempted to use the lexical sequence and whether they used it as expected) and depth of knowledge gained for the sequence. Unfortunately, however, I did not respond to usage errors with corrective feedback during the writing intervention. As discussed in Chapter 4, participant usage errors varied greatly depending on the lexical sequence in question. In short, the potential for a learning advantage through the writing task may have been lost without the feedback necessary for students to make comparisons, thereby encouraging the cognitive elaboration needed to facilitate entrenchment.

Within entrenchment and usage-based theory, the process of language acquisition is highly dependent upon an individual language learner’s experiences (Blumenthal-Dramé, 2012;

Ellis & Wulff, 2014). No two language learners have identical semantic maps because no two individuals have had the same life and language experiences. Consequently, individual learner differences may also have impacted the degree and ease with which my participants acquired depth in knowledge of the sequences. Entrenchment theory implies that when acquiring lexical sequences, L1 to L2 language distance may also be an important factor, specifically when one language has a larger degree of fixed word order than the other. In other words, because lexical sequences are linear, languages less dependent on linearity for meaning may be less amenable to acquisition of sequences than of single words (Schmid, 2017). Unfortunately, L1 differences among my participants were not large enough to be measured through quantitative analysis. Another important factor in depth of entrenchment may lie within a bilingual learner's English proficiency level; proficiency levels are positively correlated with lexical sequence acquisition generally (Ehsanzadeh, 2012; Schoonen et al., 2011; Schmitt et al., 2004; Tekman and Daloğlu, 2006; Zhao, 2016). Because vocabulary sophistication improves with increased proficiency level (Levitzky-Avid & Laufer, 2013; Muncie, 2002; White, 2015), learner proficiency level and the lexical sequence may have interacted to impact acquisition in my own study.

The lexical approach (Lewis, 1994) recommends that students at earlier stages of learning be encouraged in such a way as to avoid "the self-monitoring that would inhibit its use," which implies that "acceptable pidgin must be perceived as successful elementary and intermediate behavior, rather than defective" (p. 116). In keeping with this perception of language teaching and learning, incorrect usage may be understood as a necessary stage towards competency and therefore be recognized and supported as such. Van den Broek (2016) explains that corrective feedback must be offered to enhance vocabulary acquisition, yet students' experimental use of language should be commended as well, with the expectation that depth of knowledge will be

acquired through repeated exposure to the sequences in multiple contexts over time and with repeated usage supplemented by corrective feedback. I therefore agree with Choi (2017), who suggests that encouraging learners to practice the use of vocabulary newly encountered in reading be considered an essential part of vocabulary and reading instruction.

The life experiences of my participants impacted study results through how well bilingual learners could relate to the passages on a personal level, which sometimes determined their desire to reread a text for understanding. Tan and Mante-Estaci's (2021) study of high school students found that they are better engaged with texts that they find relevant, and this finding is supported by the work of many others as well: Aronson and Laughter (2016) reviewed literature reporting positive outcomes across content areas among students when tenets of culturally relevant, responsive, and sustaining pedagogy were adopted in classrooms. Likewise, a meta-analysis of intervention studies by McBreen and Savage (2021) found interventions focusing on improving student motivation highly effective in leading to reading gains. My results suggest that motivation may be fostered when students are presented with relevant texts. The reading passages presented to the participants provided contextual examples required for comprehension. When students in my study found the passages relatable, they reported that they returned to reread for better understanding or were motivated to comprehend the text, leading them to employ strategies for aiding in understanding. My findings emphasize that choosing texts that learners will find personally relevant will motivate them to read for learning.

When asked for their opinions about the tasks the study required them to undertake, students often commented on the novel aspect of the assignments. As discussed in Chapter 4, text relatability, the presence of interlinear glosses, the inclusion of lexical sequences instead of single words, and repeated exposure to the target phrases were some of the aspects of this study

that participants seemed to find unique and useful. One participant reported that the systematic nature of the study led reading to become a habit for him. The embedded nature of the sequences provided contextual examples that encouraged participants to practice using the sequences on their own, a process that many participants reported that they preferred over traditional assignments they are asked to complete in their classrooms. Given the interest the participants expressed in completing the tasks and the positive results of the study, these vocabulary teaching techniques hold promise for classroom teaching.

Factors Influencing Language Learning

Qualitative data led me to several inferences about factors influencing the educational experiences of the participants as discussed below. According to Schmid (2017), entrenchment is influenced by social interactions, which may include the degree of solidarity shared with the interlocuter and social dynamics between them. Although gender and L1 did not have a statistically significant impact on quantitative intervention results in my study, these two social identity factors—as well as race and disability—were found to be salient to the lives and language learning practices through their experiences with being bullied at school.

Bullying

Bullying was not a factor I considered while planning my study. Although participants were not asked about bullying, the issue of bullying was raised by most interviewees and was found to have a profound impact on the lives and learning experiences of participants. Bullying arose from social identity factors that included gender, L1, bilingual learner status, disability, and race.

Half of my participants identified English as their first language. Nearly all other participants identified Spanish as a first language, whereas the representation of other

languages—French, Vietnamese, Hindi, Kinyarwanda, and Gujarati/Hindi—was too small in the sample to be accurately measured as a between-subject factor. Consequently, I examined L1 in terms of English or non-English as reported by participants, finding that those who identified English as their L1 were less likely to report bullying, as were White students and male students. As noted earlier, the impact of bullying on language learning manifested as a reluctance among many learners to speak in front of others because of fear of being ridiculed. This was especially true for females and for those who identified a language other than English as their L1, which among my interviewees was an indication that the student did not have a nativelike accent. The most egregious cases of L1 as a predictor of bullying lay in participants who spoke a language other than Spanish, including Paper Lamp and Stephanie, and in Apollo, who had a speech impediment. Stephanie’s narrative provides a key example of how the intersection of race and bilingual learner status negatively impacted her experiences at school. Although the bullying that Stephanie experienced included being called racial epithets, she did not attribute this abuse to her race, but instead, her status as an African immigrant.

School bullying was identified by Pratt-Johnson (2015) as one of four major stressors commonly experienced by immigrant families, and one meta-analysis found that more than half of adults and children who experienced bullying exhibited symptoms of post-traumatic stress disorder (PTSD; Nielson, et al., 2015). Peker (2020) notes that bilingual learners’ perceptions of identity change as they engage with individuals in their new community and attempt to reconcile “current identity and the identity-to-be through the interactions with others in the new community” (p. 187). Consequently, bullying may profoundly impact bilingual learners’ identity perception. My study expands these findings, suggesting that bullying of bilingual learners may be widespread even in school environments with large populations of Latinx and bilingual

learner students (Sugarman & Gear, 2018). My study further found that bullying may impact language learning by creating a reluctance among learners to engage with others. This finding aligns with Hoffman (2015) and Mendez et al. (2012), who found that experiences with language learning may impact an individual's motivation for future learning, and with Piker's (2016) finding that it may lead some language learners to refrain from interactions or identification with the language learner community.

My study also adds to this literature by revealing that bullying of students because of bilingual learner status and perceived accent may be gendered and racialized, even when unrecognized by the students themselves. Furthermore, participants in my study not only discussed their experiences with bullying because of their social identities; they also gave specific accounts of how it manifested. Students reported that they were ridiculed and embarrassed by fluent and native English speakers while practicing English in activities such as group work, and participants who were bullied for any reason described a type of alienation this bullying caused. Participants who faced bullying for speaking with nonnative accents revealed to me that they do not speak to others. They explained that they like being left alone and that they do not have many friends at school. This finding is supported by Gluszek et al. (2010), who found that strength of accent predicted a "lack of social belonging" (p. 28).

The output hypothesis (Swain & Lapkin, 1995) established the importance of language use for language acquisition, and further, the lexical approach (Lewis, 1994) emphasizes the importance having freedom to experiment with language without self-monitoring for successful language acquisition. Consequently, when students are unwilling to engage in language use for fear of being ridiculed, as the participants in my study reported, their language acquisition may

suffer. Finally, my study reveals that the impact of bullying for bilingual learner students may have a compounding effect—as it continues to impact an adolescent even years later.

Claimed and Unclaimed Responsibilities

Like discussions of bullying, participants in my study often described their responsibilities inside and outside of the home in contradictory ways. For instance, several participants denied having jobs although they later described work they do. In some cases, this work was paid, albeit it was performed at home. In other cases, it included household chores, babysitting, and contributing to family-owned small businesses conducted from home.

Provis (2009) notes that *work* “is a vague and ambiguous idea that is not suited to the demands made on it” (p. 123), yet given the central place of work to one’s life and identity, how it is defined matters. I did not define the terms *work* and *job* for my participants, although I had in mind that it would include any activities for which students would be paid. Half of my participants reported that they work. When coupled with Passel’s (2011) finding that immigrant youth are more likely to work than other high school students, my finding aligns with research suggesting that 17% of 16 to 19-year-olds worked during the school year in 2021 (Forum on Child and Family Statistics, 2021). On the other hand, my study suggests that this number may be underreported, as two of my participants did not consider themselves having a job if they conducted their work from home, even when they were being paid to perform it.

Graves et al. (2017) found that teenage employment may be beneficial to teenagers, provided they work less than 20 hours per week, are 16 or older, and that they participate in after-school activities in addition to having employment. However, for students younger than 16 and those who work over 20 hours per week, there is an adverse impact on the students’ quality of life index. Graves et al.’s finding may imply that some of the participants in my study are in

danger of dropping out of school. My study also found that the way in which immigrant students undertake responsibilities at home may follow a gendered pattern. For example, although both male and female participants took part in family chores, females were the sole caregivers to children. A potential explanation may be seen in Blau et al. (2020), who propose that,

“first-generation immigrants, both women and men, from source countries with more gender equality (as measured by the World Economic Forum’s Global Gender Gap Index) allocate tasks more equally, while those from less gender equal source countries allocate tasks more traditionally ... Our findings suggest that broader cultural factors do influence the gender division of labor in the household” (p. 907).

The researchers also add that differences among source countries significantly shrunk for immigrants who arrived before the age of 18.

Contrasting Intersubjective Experiences

Bilingual learners’ language needs are informed by numerous factors—including their social identities and life experiences—making effective instruction dependent on highly individualized instruction (Leung, 1997). The responses of my participants when discussing their teachers revealed that although “helpful” and “nice,” their teachers did not always attempt to meet their learning needs. An explanation for indifference towards bilingual learner student success may be found in literature examining teacher beliefs about bilingual learners. In 2000, Sharkey and Layzer found that teachers of bilingual learners often attribute academic issues to the family’s cultural values. Indeed, research regularly finds evidence that mainstream teachers hold negative implicit beliefs about bilingual learners, often as a curious juxtaposition to their reported positive explicit beliefs about the same students (e.g., Harrison & Lakin, 2018;

McLaughlin & Pettit, 2013; Reeves, 2004; Walker et al., 2004). For instance, the teacher participants in Walker et al.'s study admitted that they did not want bilingual learners in their classrooms and did not want professional development to assist them in strategies for teaching bilingual students. However, the same teacher participants believed that their schools respected and embraced bilingual students, including their languages and cultures, and they reported that bilingual learners contributed to needed diversity in their schools. A similar discrepancy is reflected among my own participants' views of their teachers as nice but ineffective. Participants in my study were placed in classrooms where teachers sometimes seemed to lack knowledge of effective strategies for teaching them. While the studies of Sharkey and Layzer and Walker et al. are roughly 20 years old, not much seems to have changed.

Research suggests that teacher implicit bias may be culpable for the lack of language support common among my participants. Using the Implicit Association Test–EL, Harrison and Lakin (2018) found that among 197 mainstream teachers across the content areas, 55% held negative attitudes towards bilingual learners despite an explicit beliefs assessment finding strong positive attitudes about bilingual learners. Although nearly all students in my study emphasized that their teachers were kind to them, some students explained that they felt the work they were given was not challenging enough, which may indicate implicit beliefs that bilingual learners are unable or unwilling to complete challenging work or that they do not have aspirations towards higher education—a “benevolent conspiracy”—as Hatch (1992, p. 67) termed it. Sharkey and Layzer (2000) argue that when bilingual learner teachers focus solely on a positive environment but fail to provide the students with challenging content, they are, in effect, blocking access needed for college preparedness.

An explanation for the benevolent conspiracy might be offered by McLaughlin and Pettit (2014), whose study revealed that teachers often hold strong deficit views of bilingual learner students and their families. Teacher participants in McLaughlin and Pettit's study stated that Latinx families lacked the right priorities for their children to be successful in school, that the values of Latinx families interrupted their children's school success, and that Latinx students had no positive role models to keep them in school. Indeed, one teacher participant in McLaughlin and Pettit's (2014) study said that she "struggled with 'instilling' the right kind of values in her Latino students" (p. 120). Another teacher participant implied that Latinx families do not want their children to complete high school, saying that Latinx families believe that "at that point, the person needs to be doing something else, not being in school" (p. 120).

In contrast to the deficit perceptions of Latinx students often held by teachers, I found that families of nearly all interviewees expect their children to succeed in school and emphasize the importance of education to their children. Although some participants noted a disconnect between their school experiences and their parents' belief that school is a wholesome place of learning—void of such things as bullying, apathetic teachers, and drug use—nearly all students in my study described parents who provide warm, loving homes and who teach them skills for success in all facets of life. Parents also provided an abundance in the funds of knowledge often unavailable in traditional school settings. For instance, participants discussed learning such things as car maintenance and small business management from their parents. The parents of my participants want their children to finish high school, to attend college, and to succeed in life. Most bilingual learners in my study described high educational aspirations for themselves and explained enthusiastically that their parents supported and encouraged these goals.

An important contrast between perceptions of learners and their realities can also be seen in the narrative of Sophia. As a lesbian student, Sophia offered a powerful counter to deficit views of LGBTQ+ youth when she discussed homophobia in her life by pivoting the problem away from her identity and towards those who engage in homophobia. She radiated exuberance throughout the interview, and the anecdotes she shared about her approach to anti-LGBTQ+ sentiments at home and at school demonstrated confidence and positive self-talk. She positioned herself as having an advantage over those at her school who do not support LGBTQ+ lifestyles by explaining that it is their problem to grapple with; because they are homophobic, they miss an opportunity to get to know her, which is unfortunate for them. When she described her mother's homophobia, she implied that her mother's disparaging comments mean that her mother does not always get to enjoy the pleasure of her company. Her defiance also appeared when she discussed how she would react if someone at school accosted her: "I would definitely defend myself (.) because I'm not just going to stand there and (.) take it." Lu et al.'s (2021) study of gay Latinx men in college demonstrated that "positive reframing" can provide a powerful means of emotional regulation to help students cope with social stress in school.

Study Limitations

Despite several important implications of my findings, there are some limitations to my study. For instance, all participants came from a single high school environment, and nearly all were native Spanish speakers. Because participant proficiency data were not collected, the impact of proficiency on results was not measured. During data collection, participants were not provided with feedback on their written responses—a measure that may have helped me untangle the reasons for my finding that writing only strengthened reading for one group. Nevertheless, it should be noted that for both interventions, participants made significant gains in learning and

retained their learning several weeks later. Because the lexical sequences were chosen before participants were given pretests, the pretest results could not be considered when determining which lexical sequences would optimize acquisition; using sequences that participants had initial exposure to may have better facilitated acquisition of further depth of knowledge. AOA has been linked to educational outcomes for bilingual learners (Basu, 2018), and gender has been linked to accent attainment (Polat & Mahalingappa, 2010). Consequently, a further exploration of these two factors may have added insight to the study results.

The sample size of the quantitative strand was 36% higher than the recommended sample size for a repeated measures factorial ANOVA. All participants underwent both interventions, effectually making the participants their own comparison group, yet it necessitated the counterbalancing of interventions and the use of different lexical sequences for reading and writing interventions between groups. A larger two-group study of 100 participants may have eliminated the possibility of intervention order as having an impact on my quantitative finding that for one group, reading alone was more impactful, and for the second group, writing coupled with reading was more effective. Finally and most importantly, in retrospect, a final questionnaire given to all participants—one that reflected the results of the interview findings—could have been used to determine the extent to which such things as bullying from social identity factors, the quarantine, and text relatability impacted all 49 participants.

Recommendations for Policy and Practice

As Gloria Ladson-Billings (2021) has noted, despite a significant amount of educational research devoted to Black, Latinx, and immigrant students, we “rarely provide the kind of remedies that help them to solve their problems” (p. 60). Thus, recommendations for policy and

teaching practices having arisen from research findings are arguably the core reasons for undertaking educational research.

Policy Recommendations

Although all of my participants were enrolled in their school's ESOL program, nearly half identified English as their first language, with many reporting that they were born in the United States. The process of identifying students in need of English language support services may begin at an early age, but many bilingual students remain in ESOL programs well into high school. Such students are referred to as long-term English learners and “provoke particular concern” (Clark-Gareca et al., 2019, p. 1). The bilingual learners in my study who qualify as part of this subgroup often described teaching practices in ESOL classrooms that lacked the rigor needed for their exit from the language support program. This self-perpetuating cycle suggests needed changes in the policies that guide diagnostic and standardized language assessments, which seem to prevent them from entering mainstream classrooms and have arisen from policies guided by deficit views about them. Given that bilingual learners are among the fastest growing groups in American schools (Clark-Gareca et al.), attention to seeking ways to understand the role of segregated instruction and standardized assessment practices in perpetuating this cycle seems warranted. As the process currently stands, placing some bilingual learners into ESOL classrooms at an early age sets a trajectory for them that has consequences for the remainder of their lives.

Vocabulary Teaching

Usage-based theory and entrenchment hold that exposure to lexical sequences provides not only advantages for learning vocabulary, but also for the acquisition of language structures, and teaching bilingual learners lexical sequences common to academic genres may empower

students by allowing them freedom to experiment with the language of academic success. As Cancio and Iturrieta (2022) explain,

Advocates for the LA argue that classes focusing on learning as a product should be left behind, as teachers need to go beyond delivering content and completing textbook activities that highlight accuracy. Instead, this approach requires learners to engage in the observation of language features, to hypothesize on how such features can be used to communicate, and to experiment with language to communicate successfully” (p. 43)

The lexical approach proposes shifting an understanding of language as “grammaticalized lexis” as opposed to “lexicalized grammar” (Lewis, 1994, p. vi). Advice for incorporating the lexical approach into teaching was first offered by Lewis in 1994, but its principles remain sound, especially the promotion of extensive, vocabulary-rich reading that exposes students to language in varied contexts.

While it is sometimes necessary to provide a translation for single words for beginning bilingual learners, the reduction of vocabulary into lists of L1 translations to memorize may not be as effective as contextual learning through reading and writing. Students in my study successfully learned lexical sequences introduced through the reading of short texts, without being asked to memorize or study the vocabulary items, and they retained their learning in delayed posttesting. Several participants reported that they enjoyed the experience, particularly when they found the readings personally relevant to them. However, tasks requiring participants to recall or use newly learned vocabulary should include feedback to maximize the benefit of practiced learning, yet students’ initial usage practice must be encouraged and commended;

students should not fear making mistakes while experimenting with language usage. Indeed, fear of making mistakes led some of my participants to avoid using English altogether.

While general frequency may be a useful heuristic in choosing targeted lexical sequences for teaching, other factors may be at play in determining which sequences are most likely to be learned. Because some sequences are more readily acquired than others (e.g., *a wide range of*), some may require more exposure in varied context than others. When this process is viewed as necessary for language acquisition, concerns about whether a student has learned a specific word or phrase within a standardized, measured time frame are incompatible with accepted theories of language learning that my study supports (i.e., usage-based theory and entrenchment theory). When choosing or writing texts for exposure to sequences, teachers should take care to ensure that students will find the readings relatable and interesting because, as reported by my participants, doing so may increase the propensity for bilingual learners to better comprehend readings. Relatable readings, defined as those materials that connect with the students' interests or cultural backgrounds, as well as texts that address current events or social issues, motivate learners and enhance reading comprehension (Guthrie et al., 2012; McDevitt, 2021), particularly when texts have personal relevance to them. When my participants wanted to better understand a reading they found relatable, they returned to the texts to read it again or looked up unfamiliar vocabulary. Webb's (2020) vocabulary research handbook devotes equal time to lexical sequences and single terms for every topic presented, an indication that the examination of lexical sequences as vocabulary is becoming increasingly common in vocabulary research. However, students in my study found phrases as units of vocabulary learning a novelty, which highlights a disconnect between research and teaching practices. The same is equally true for the use of glossing to aid readers in comprehension; an abundance of glossing research has found

benefits to bilingual learners (Lofgren, 2022), yet as reported by my participants, it seems remarkably uncommon as a classroom practice.

Combating Bullying

Bullying remains a common problem in American public schools (Rettew & Pawlowski, 2022), yet antibullying campaigns meant to curb abuse are ineffective among older adolescents (Yeager et al., 2015). Most participants in my study described bullying as a regular part of the high school experience, with one participant discussing its prevalence even though he was not subjected to it. Because power differentials often drive bullying behavior (Peker, 2000), using a critical approach to teach children of all ages to understand the nature of hierarchical relationships seems paramount.

Raise Critical Language Awareness. Bullying often arises from an imbalance of power between the perpetrator and the victim (Peker, 2000). Consequently, successful campaigns to stop bullying must begin with teaching students and teachers to understand how and why power hierarchies operate in society through teaching sociocultural perspectives on diversity. Raising critical language awareness (CLA) among students and teachers may be a promising way to combat the bullying of bilingual learner students. Critical language awareness refers to awareness of the political nature of language ideology and the understanding that language practices are socially constructed and used to reinforce power hierarchies (Fairclough, 2013). Race and language are often linked through a language ideology that assigns accents associated with White speakers more prestige than those spoken by others (Kubota, 2004; Lindemann, 2003; Pennycook, 1998). Furthermore, high school students sometimes conflate immigration status with deviant behavior (Seirk, 2019). Teaching students and teachers to consider how language ideology operates in organizational arenas such as school while combating the “illegal

immigrant” narrative through the sharing of immigrant stories may help address the underlying reasons that bilingual learners are often targeted for bullying and may plant seeds of understanding and empathy towards others.

Teachers are the primary factor in Latinx education (Gonzalez, 2010; Villenas & Deyhle, 1999). Because implicit bias often rears itself in teachers through standards that are not always high for bilingual learner students, teachers must learn to question assumptions they hold about Latinx bilingual learners and work to challenge stereotypes (McLaughlin & Pettit, 2014). One way for teachers to achieve this is through educating themselves about the social identities and lived experiences of the bilingual learners in their classrooms. The participants in my study considered their teachers to be kind, yet some learners revealed that they are often left to fend for themselves in the classroom. As McGrail (2017) poignantly argues,

by putting value to only certain parts in student performance and abilities, we are likely to miss out on the other competencies and performance indicators or ways of being and seeing that also constitute our students’ academic and social identities. As a result, we may end up representing only one image of our students as learners and individuals, the part which does not necessarily reflect the whole human beings that they truly are (p. 1)

When teachers learn about their students’ lived experiences and social identities, they are better prepared to consider the whole student and are less likely to allow biases to impact their teaching practices.

Teach through Critical Race Theory and Intersectionality Frameworks. Bullying occurred for one Black participant because of her race; another described bias she faced by saying she is treated as though she were Black. When attempting to teach students about racism,

teachers often define racism very narrowly, ignoring the systemic nature in favor of presenting it as individual prejudice (Wills, 2019). However, combating racism requires teaching students the systemic nature of racism (Farag, 2021; Gaynor, 2021; Peet et al., 2020) as well as its history (Farag, 2021; Gaynor, 2021; Sheth, 2018) and to critically examine race as a social construct (Farag, 2021; Peek et al. 2020).

For example, students may be challenged to consider what they mean by “normal,” a term that sometimes appeared when I asked students if they believed they were treated or viewed in any special way at school by teachers, students, or anyone else. Farag (2021) argues that teachers need a framework to guide them in teaching about racism. Recent examples of how teachers in varied contexts have successfully achieved an understanding of racism among their students through the use of a framework can be seen in Farag’s work and in Gaynor (2021), Sheth (2018), and Peet et al. (2021), all of whom relied on critical race theory (CRT), intersectionality, or both to help students to critically examine racism through their respective content areas. Students have examined the impact of racism, how social identities influence the lives of marginalized people, and how power and privilege operate to maintain a status quo that perpetuates social inequity. The opportunity gap between students of color and White students is a product of systemic racism, and as Gloria Ladson-Billing (2009) explains, the gap is a product of a long history that must be taught as well.

Through CRT and intersectionality, students may also learn the role that implicit bias plays in their perceptions of those who are different from them, including those of different sexual orientations, ability status, socioeconomic status, bilingual learner status, and accent. Students should not have to wait until they attend university to learn the secret that we all possess implicit biases that continually drive our behavior and to learn the importance of

surfacing, reflecting on, and questioning such biases. Students should not have to wait until they attend university to learn that treating people who are different with contempt—without even understanding why—has devastating consequences for individuals and for society. Students must learn to question why some students engage in bullying and why the problem seems so relentless despite the damage it causes. CRT and intersectionality frameworks may encourage students who face discrimination based on social identities and their intersections to take pride in their identities—like Sophia, whose narrative demonstrates the power of shifting the locus of control to oneself—and may help prevent the bullying of students like Paper Lamp, Stephanie, and Jenny, who avoid speaking English and interacting with others at school because of the intersections of bilingual learner status, presence of a nonnative accent, and gender—or of Apollo, whose social identity intersection spans a bilingual learner status, nonnative accent, and a communication disorder.

Recommendations for Further Research

More research is needed on the acquisition of individual lexical sequences through reading and writing to determine whether the findings of my study are supported in other contexts. For instance, a comparative study using the same lexical sequences for both interventions may reveal better than my own study whether or not a writing intervention leads to greater gains than reading. Moreover, research on the process of a learner's movement in depth of lexical sequence knowledge may lead to a better understanding of why some lexical sequences were more readily acquired than others or whether this finding is unique to my study.

Further research is also needed to explore several issues facing bilingual learner students that my findings brought to light. First, it would be useful to explore whether the findings of my 12 participants hold true for a larger group of students in other schools. A multiple regression

analysis of bilingual learners using a larger sample from multiple schools may shed light on the extent to which school environment may account for the finding that female and White bilingual learners were more likely to report bullying, and if found to be a widespread occurrence, this would necessitate examining potential reasons for this finding. Repetition studies comparing reading and writing interventions among bilingual learners with a greater range of L1s than my own participants and with greater variation in demographics are needed to better determine the impact of L1 on interventions such as mine, and given the finding that bilingual learners who identified English as their first language had more positive learning experiences, a comparison study between native English speakers and bilingual learners might better tease out potential reasons. Moreover, despite ever increasing research emphasizing the negative impact on literacy brought about by the pandemic, further research examining narratives of students for whom the quarantine helped establish positive literacy habits might reveal factors that could lead to recommendations for teaching and learning applicable even in post-pandemic times. Further research is needed to determine which combination of study aspects that my participants found useful (e.g., glossing, the scheduled presentation of readings, the text interest) are most impactful for bilingual learners and why.

Finally, my study highlights the important place of integrative mixed methods methodology—the full integration of both qualitative and quantitative strands—in conducting educational studies, particularly in studies attempting to uncover factors affecting learning. The quantitative strand allowed me to compare the success of two interventions using reading and writing to help bilingual learners acquire depth of knowledge of lexical sequences, but the inclusion of the qualitative analysis provided much more. For example, without a quantitative understanding of written responses to readings (as qualitative data) and the subsequent

quantitative analysis of the lexical sequences themselves, I may not have considered how the individual lexical sequences played a role in learning.

Likewise, without the delving aspect integral to semi-structured interviews, I would be left with many more questions than quantitative measures alone could answer. For example, a quantitative survey alone could not have uncovered the contradictory ways that my participants defined “work” and the true nature of their intersubjective experiences with their teachers. I would have large gaps in understanding the role of reading interest in the success of the interventions. Without semi-structured interview data, I could not have discovered the nuanced ways that the participants’ social identities—disability, gender, bilingual learner status, and L1—impacted their learning despite the absence of impact from these factors on the quantitative intervention results. These factors were uncovered only through the appearance of bullying in representational and intersubjective arenas of studies—something that the quantitative analysis alone could not have shown.

References

- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*(2), 249-306. <https://doi.org/10.1111/j.1467-9922.2009.00507.x>
- AbuSeileek, A. F. (2011). Hypermedia annotation presentation: The effect of location and type on the EFL learners' achievement in reading comprehension and vocabulary acquisition. *Computers & Education, 57*(1), 1281-1291. <https://doi.org/10.1111/jcal.12004>
- Achebe, C. (1959). *Things fall apart*. Heinemann.
- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*(2), 81-92. <https://doi.org/10.1016/j.esp.2011.08.004>
- Adlof, S. M., & Patten, H. (2017). Nonword repetition and vocabulary knowledge as predictors of children's phonological and semantic word learning. *Journal of Speech, Language, and Hearing Research, 60*, 682-693. https://10.1044/2016_JSLHR-L-15-0441
- Agarwal, P. K. (2019). Retrieval practice & Bloom's taxonomy: Do students need fact knowledge before higher order learning? *Journal of Educational Psychology, 111*(2), 189-209. <https://doi.org/10.1037/edu0000282>
- Aguilar, I., Nelson, S., & Manuel Nino, J. (2016). Swimming between: An examination of the inherent complexity within social justice. *The Teacher Educator, 51*, 250-265. <https://doi.org/10.1080/08878730.2016.1177331>
- Agustín Llach, M. P. (2016). Does L1 make a difference? Evidence from the receptive vocabulary size of Spanish and German EFL primary school learners. *The Language Learning Journal, 44*(2), 214-235. <https://doi.org/10.1080/09571736.2012.737358>

- Agustín Llach, M. P., & Terrazas, M. (2012). Vocabulary knowledge development and gender differences in a second language. *Estudios de lingüística inglesa aplicada*, 12, 45-75.
- Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology*, 106, 419-434. <https://doi.org/10.1037/a0035692>
- Ahour, T., & Abdi, M. (2015). The relationship between EFL learners' multiple intelligences and vocabulary learning strategies use with a focus on gender. *Theory and Practice in Language Studies*, 5(4), 800-809. <https://doi.org/10.17507/tpls.0504.16>
- Ajideh, P., Rahimpour, M., Amini, D., & Farrokhi, F. (2013). Motivational strategies, task effectiveness and incidental acquisition of second language vocabulary. *Journal of Language Teaching and Research*, 4(5), 1044-1052. <https://doi.org/10.4304/jltr.4.5.1044-1052>
- Alavi, S., & Keyvanshekouh, A. (2012). Using the MoodleReader as an extensive reading tool and its effects on Iranian EFL students' incidental vocabulary learning. *English Language Teaching*, 5(6), 135-145. <https://10.5539/elt.v5n6p135>
- Alavinia, P., Rahimi, H. (2019). Task types effects and task involvement load on vocabulary learning of EFL learners. *International Journal of Instruction*, 12(1) 1501-1516. <https://doi.org/10.29333/iji.2019.12196a>
- Albrechtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Processes and development*. Palgrave Macmillan. <https://doi.org/10.1057/9780230593404>
- Alcón, E. (2007). Incidental focus on form, noticing and vocabulary learning in the EFL classroom. *International Journal of English Studies*, 7(2), 41-20.

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.
- Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment, (part 1). *Language Testing*, 18, 213-236. <https://doi.org/10.1017/S0261444800014464>
- Alharbi, B. (2018). The impact of glossed texts on reading comprehension among tertiary Saudi students. *English Language Teaching*, 11(3), 153-161.
<http://doi.org/10.5539/elt.v11n3p153>
- Alharbi, M. A. (2016). Using different types of dictionaries for improving EFL reading comprehension and vocabulary learning. *JALT CALL Journal*, 12(2), 123-149. (EJ1112430). ERIC. <http://files.eric.ed.gov/fulltext/EJ1112430.pdf>
- Alharthi, T. (2015). Reasons for vocabulary attrition: Revisiting the state of the art. *International Education Studies*, 8(10), 86-95. <https://doi.org/10.5539/ies.v8n10p86>
- Alsaif, A., & Masrai, A. (2019). Extensive reading and incidental vocabulary acquisition: The case of a predominant language classroom input. *International Journal of Education & Literacy Studies*, 7(2), 39-45. <https://doi.org/10.7575/AIAC.IJELS.V.7N.2P.39>
- Amankwaa, L. (2016). Creating protocols for trustworthiness in qualitative research. *Journal of Cultural Diversity*, 23(3), 121-127.
- Amini, D., & Iravani, Z. (2021). Generic differentiation in the relationship between L2 vocabulary knowledge and writing performance. *Journal of Modern Research in English Language Studies*, 8(1), 97-117. <https://doi.org/10.30479/jmrels.2020.14009.1723>
- Amirian, S. M. R., & Behshad, A. (2016). The effect of native model writing on the vocabulary richness of Iranian EFL learners' written text in a four-stage writing task. *Theory & Practice in Language Studies*, 6(8), 1713-1721. <http://doi.org/10.17507/tpls.0608.28>

- Anderson, R. C., & Freebody, P. (1979). *Vocabulary knowledge* (Technical Report 136). U. S. Department of Health Education & Welfare, National Institute of Education.
- Appel, R., & Murray, L. (2020). L1 differences in L2 English academic writing: A lexical bundles analysis. *Journal of English for Academic Purposes, 46*, 1-5.
[https://doi.org/ 10.1016/j.jeap.2020.100873](https://doi.org/10.1016/j.jeap.2020.100873)
- Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. *Review of Educational Research, 86*(1), 163-206. <https://doi.org/10.3102/0034654315582066>
- Azari, F., Abdullah, F. S., Heng, C. S., & Hoon, T. B. (2012). *Effects of glosses on vocabulary gain and retention among tertiary level EFL learners*. (ED533228). ERIC.
<http://files.eric.ed.gov/fulltext/ED533228.pdf>
- Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon, 5*, 436-461. <https://doi.org/10.1075/ml.5.3.10baa>
- Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of Second Language Writing, 18*, 191-208.
<https://doi.org/10.1016/j.jslw.2009.05.003>
- Baddeley, A. & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47-89. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Baddeley, A. D. & Hitch, G. (2017). Is the levels of processing effect language-limited? *Journal of Memory and Language, 92*, 1-13. <https://doi.org/10.1016/j.jml.2016.05.001>
- Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex, 112*, 91-106. <https://doi.org/10.1016/j.cortex.2018.05.015>

- Baddeley, A., Gahtercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158-173. <https://doi.org/10.1037/0033-295X.105.1.158>
- Bardel, C., Lindqvist, C., & Laufer, B. (2013). Foreword. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge, and use: New perspectives on assessment and corpus analysis*. European Second Language Association.
- Barsalou, L. W. (1992). Frames, concepts, and conceptual fields. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 21-74). Lawrence Erlbaum Associates.
- Basu, S. (2018). Age-of-arrival effects on the education of immigrant children: A sibling study. *Journal of Family and Economic Issues*, *39*, 474-493. <https://doi.org/10.2139/ssrn.1720573>
- Bayat, A., & Oveidi, M. (2015). The effect of cognates on vocabulary retention of Irania bilingual EFL students. *Cumhuriyet University Faculty of Science Journal (CSJ)*, *36*(3), 1211-1215. <https://doi.org/10.1080/09571736.2012.708051>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.
- Bereiter, C., Scardamalia, C., & Hewitt, J. (1997). Postmodernism, knowledge building, and elementary science. *The Elementary School Journal*, *97*(4), 1-15 *The Elementary School Journal*, *97*(4), 324-340. <https://doi.org/10.1086/461869>
- Bergman, M. M. (2008). The straw men of the qualitative-quantitative divide and their influence on mixed methods research. In M. M. Bergman (Ed.), *Advances in mixed methods research* (pp. 11-21). SAGE.

- Biber, D. (2009). A corpus-driven approach to formulaic language in English multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, 14(3), 275-311. <https://doi.org/10.1075/ijcl.14.3.08bib>
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263-286. <https://doi.org/10.1016/j.esp.2006.08.003>
- Biber, D., Conrad, S., Cortes, V. (2004). "If you look at ...": Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405. <https://doi.org/10.1093/applin/25.3.371>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Biddle, C., & Shaftt, K. A. (2015). Axiology and anomaly in the practice of mixed methods work: Pragmatism, valuation, and the transformative paradigm. *Journal of Mixed Methods Research*, 9(4), 320-334. <https://doi.org/10.1177/1558689814533157>
- Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 68(4), 706-755. <https://doi.org/10.2307/416851>
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56(1), 9-49. <https://doi.org/10.1111/j.1467-9922.2006.00353.x>
- Blanchett, W. J., Klingner, J. K., & Harris, B. (2009). The intersection of race, culture, language, and disability: Implications for urban education. *Urban Education* 44(4), 389-409. <https://doi.org/10.1177/0042085909338686>
- Blau., F. D., Kahn, L. M., Comey, M., Eng., A., Meyerhofer, P., & Willén, A. (2020). Culture and gender allocation of tasks: Source country characteristics and the division of non-

- market work among US immigrants. *Rev Econ Household*, 18, 907-958.
<https://doi.org/10.1007/s11150-020-09501-2>
- Blom, E., Paradis, J., & Sorenson Duncan, T. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular -s in child L2 English. *Language Learning*, 20(10), 1-30. <https://doi.org/10.1111/j.1467-9922.2012.00715.x>
- Blum, S., & Levenston, E. (1979). Lexical simplification in second language acquisition. *Studies in Second Language Acquisition*, 2, 85-94.
- Blumenthal-Dramé, A. (2012). *Entrenchment in usage based theories: What corpus do and do not reveal about the mind*. De Gruyter Mouton.
- Blumenthal-Dramé, A. (2017). Entrenchment from a psycholinguistic and neurolinguistics. In H. J. Schmid (Ed.), *Entrenchment of the psychology of language learning*. American Psychological Association.
- Bogdan, R. & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theories and methods* (5th ed.). Pearson.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589-597.
<https://doi.org/10.1080/2159676X.2019.1628806>
- Braun, V., & Clarke, V. (2020). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*.
<https://doi.org/10.1080/14780887.2020.1769238>

- Breeze, R. (2013). Lexical bundles across four legal genres. *International Journal of Corpus Linguistics*, 18(2), 229-253. <https://doi.org/10.1075/ijcl.18.2.03bre>
- Briggs, C. L. (2000) Interviewing, power/knowledge, and social inequality. In N. Y. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 495-506). SAGE.
- Brown, D. (2008). Using a modified version of the vocabulary knowledge scale to aid vocabulary development. *The Language Teacher*, 32(12), 15-16.
- Brown, N. (2018). "Listen to your gut": a reflexive approach to data analysis. In A. P. Costa (Ed.), *Abstracts Book of the 3rd World Conference on Qualitative Research, Vol. 2* (pp. 231-232). World Conference on Qualitative Research.
- Bruffee, K. (1986). Social construction, language, and the authority of knowledge: A bibliographical essay. *College English*, 48, 773-790. <https://doi.org/10.2307/376723>
- Brusnighan, S. M., & Folk, J. R. (2012). Combining contextual and morphemic cues is beneficial during incidental vocabulary acquisition: Semantic transparency in novel compound word processing. *Reading Research Quarterly*, 47(2), 172-190. <https://doi.org/10.1002/RRQ.015>
- Burgess, A. (1962). *A clockwork orange*. Heinemann.
- Burke, R., Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133. <https://doi.org/10.1177/1558689806298224>
- Byrd, P., & Coxhead A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5, 31-64.

- Cai, J. L. (2016). An exploratory study on an integrated genre-based approach for the instruction of academic lexical phrases. *Journal of English for Academic Purposes*, 24, 58-74.
<https://doi.org/10.1016/j.jeap.2016.09.002>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
<https://doi.org/10.1037/h0046016>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs*. Rand McNally & Company.
- Canagarajah, S. (2013). *Translingual practice: Global Englishes and cosmopolitan relations*. Routledge. <https://doi.org/10.1002/tesq.115>
- Cao, Z. (2013). The effects of tasks on the learning of lexical bundles by Chinese EFL learners. *Theory and Practice in Language Studies*, 3(6), 957-962.
<https://doi.org/10.4304/tpls.3.6.957-962>
- Cappelle, B., & Grabar, N. (2016). Towards an n-grammar of English. In *Constructionist approaches to second language acquisition and foreign language teaching* (pp. 271-302). John Benjamins.
- Cappelle, B., Depraetere, I., & Lesuisse, M. (2019). The necessity modals *have to*, *must*, *need to* and *should*: Using n-grams to help identify common and distinct semantic and pragmatic aspects. *Constructions and Frames* 12(2), 220-243. <https://doi.org/10.1075/cf.00029.cap>
- Caro, K. & Rosado Mendinueta, N. (2017). Lexis, lexical competence, and lexical knowledge: A review. *Journal of Language Teaching and Research*, 8(2), 205-213.
<https://doi.org/10.17507/jltr.0802.01>

- Carroll, S. E. (2001). *Input and evidence: The raw material of second language acquisition*. John Benjamins Publishing Company.
- Castillo-Montoya, M. (2016). Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report*, 21(5), 811-830.
- Castro-Olivo, S., Preciado, J., Le, Loan, Marciante, M., & Garcia, M. (2017). The effects of culturally adapted version of First Steps to Success for Latino bilingual learners: Preliminary pilot study. *Psychology in the Schools*, 55, 36-49.
<https://doi.org/10.1002/pits.22092>
- Çekiç, A., & Bakla, A. (2019). The effects of spacing patterns on incidental L2 vocabulary learning through reading with electronic glosses. *Instructional Science*, 47, 353–371.
<https://doi.org/10.1007/s11251-019-09483-4>
- Center for Disease Control and Prevention (2020, November 24). *Frequently asked questions about COVID-19 vaccination*. U. S. Department of Health and Human Services.
<https://www.cdc.gov/coronavirus/2019-ncov/vaccines/faq.html>
- Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693-713.
<https://doi.org/10.1093/applin/amq031>
- Cheng, Y., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention. *Reading in a Foreign Language*, 21(2), 119-142.
- Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, 38(6), 967-978. <https://doi.org/10.1037/0012-1649.38.6.967>

- Choi, Y. H. (2017). Roles of receptive and productive vocabulary knowledge in L2 writing through the mediation of L2 reading ability. *English Teaching*, 72(1), 3-24.
<http://doi.org/10.15858/engtea.72.1.201703.3>
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Praeger.
- Clark-Gareca, B., Short, D., Lukes, M., & Sharp-Ross, M. (2017). Long-term English learners: Current research, policy, and practice. *TESOL Journal*, 11(1).
<https://doi.org/10.1002/tesj.452>
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 225-237). Cambridge University Press.
- Coe, R. (2002, September 12-14). *It's the effect size, stupid: What effect size is and why it is important* [Paper presentation]. Annual Conference of the British Educational Research Association, University of Exeter, England.
<http://www.leeds.ac.uk/educol/documents/00002182.htm>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Columbo, M., Tigert, J. M., & Montecillo Leider, C. (2018). Positioning teachers, positioning learners: Why we should stop using the term English learners. *TESOL Journal*, 1-5.
<https://doi.org/10.1002/tesj.432>
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72-89. <https://doi.org/10.1093/applin/amm022>

- Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives*, (pp. 279-300). Lawrence Erlbaum.
- Cook, V. (2016) Where is the native speaker now? *TESOL Quarterly*, 50(1), 186-189.
<https://doi.org/10.1002/tesq.286>
- Cooper, H., & Hedges, L. V. (1994). Research synthesis as a scientific enterprise. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 3-14). Russell Sage Foundation.
- Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47(4), 671-718. <https://doi.org/10.1111/0023-8333.00025>
- Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17(4), 391-406.
<https://doi.org/10.1016/j.linged.2007.02.001>
- Coxhead, A. & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16(7), 129-147.
<https://doi.org/10.1016/j.jslw.2007.07.002>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
<https://doi.org/10.2307/3587951>
- Craik, F. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
[https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

- Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *The Journal of Educational Research*, 36, 206–217. <https://doi.org/10.1080/00220671.1942.10881160>
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistics sophistication. *Journal of Research in Reading*, 35(2), 115-135. <https://doi.org/10.1111/j.1467-9817.2010.01449.x>
- Cruse, D. A. (1996). *Lexical semantics*. Cambridge University Press.
- Cumming, A. (2017). Writing development and instruction for bilingual learners. In C. A. MacArthur & S. Graham (Eds.), *Handbook of writing research* (2nd ed., pp. 364-376). The Guilford Press.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33(6), 934-945. <https://doi.org/10.1037/0012-1649.33.6.934>
- Cunningham, A. E., Stanovich, K. E., & Maul, A. (2011). Of correlations and causes: The use of multiple regression modeling in literacy research. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 50-69). The Guilford Press.
- Curry, L. A., Nembhard, I. M., & Bradley, E. H. (2009). Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation*, 119(10), 1442-1452. <https://doi.org/10.1161/CIRCULATIONAHA.107.742775>
- Custodio, B., & O'Loughlin, J. B. (2020). Students with interrupted formal education: Understanding who they are. *American Educator*, 44(1), 9-11.
- Dabaghi, A., & Rafiee, M. (2012). Incidental vocabulary learning and the development of receptive and productive vocabulary: How gloss types work. *Contemporary Linguistics*, 38(74), 175-187.

- Dakhi, S., & Nur Fitria, T. (2019). The principles and the teaching of English vocabulary: A review. *Journal of English Teaching*, 5(1), 15-25. <https://doi.org/10.33541/jet.v5i1.956>
- Danilović, J. (2010). Language transfer, cognates and productive derivational knowledge of Serbian EFL learners. *Наслеђе*, 7(15/2), 77-83.
- Darling-Hammond, L., Bae, S., Cook-Harvey, C. M., Lam, L., Mercer, C., Podolsky, A., & Stosich, E. L. (2016). *Pathways to new accountability through every student succeeds act*. The Learning Policy Institute.
- Daskalovska, N. (2014). Incidental vocabulary acquisition from reading an authentic text. *Reading Matrix: An International Online Journal*, 14(2), 201-216.
- Daulton, F. E. (2010). High-frequency English-based loanword cognates for EFL in Japan. *Asian Journal of English Language Teaching*, 65-80.
- Davis, M., & Morley, J. (2015). Phrasal intertextuality: the responses of academics from different disciplines to students' re-use of phrases. *Journal of Second Language Writing*, 28, 20-35. <https://doi.org/10.1016/j.jslw.2015.02.004>
- De Smet, H. (2017). Entrenchment effects in language change. In H. J. Schmid (Ed.), *Language and the human lifespan series. Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 75-99). De Gruyter Mouton. <https://doi.org/10.1037/15969-005>
- Deans, T. (2009). Richard Rorty's social hope and community literacy. *Community Literacy*, 3(2), 3-18.
- DeCapua, A., & Marshall, W. (2011). Reaching Els at risk: Instruction for students with limited or interrupted formal education. *Preventing School Failure: Alternative Education for Children and Youth*, 55(1), 35-41. <https://doi.org/10.1080/10459880903291680>

- DeCuir-Gunby, J. T., & Schutz, P. A. (2017). *Developing a mixed methods proposal: A practical guide for beginning researchers*. SAGE. <https://doi.org/10.4135/9781483399980>
- Denzin, N. K. (2000). The practices and politics of interpretation. In N. Y. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 897-922). SAGE.
- Denzin, N. K. (2010). Moments, mixed methods, and paradigm dialogs. *Qualitative Inquiry*, 16(6), 419–427. <https://doi.org/10.1177/1077800410364608>
- Denzin, N. K., & Lincoln, Y.S. (2000). Methods of collecting and analyzing empirical materials. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 633-644). Sage.
- Derrida, J. (1967/2016). Of grammatology. In R. Rivkin & M. Ryan (Eds.), *Literary theory: An anthology* (pp. 300-331). Blackwell.
- Dewey, J. (1916/1967). *Democracy and education: An introduction to the philosophy of education*. Free Press.
- Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford University Press.
- Dillon, D. R., & O'Brien, D. G. (2019). Pragmatism [not just] practicality as theoretical framework in literacy research. In D. E. Alvermann, N. J. Urrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed., pp. 582-600). Routledge.
- Dillon, D. R., O'Brien, D. G., & Heilman, E. E. (2000). Literacy research in the next millennium: From paradigms to pragmatism and practicality. *Reading Research Quarterly*, 35(1), 10-26. <https://doi.org/10.1598/RRQ.35.1.2>

- Dörnyei, Z. (2018). *Research methods in applied linguistics*. Oxford University Press.
- Douven, I. (2017). Abduction. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Summer 2017 ed.) Stanford University. <https://plato.stanford.edu/entries/abduction/>
- Dressman, M., & McCarthy, S. J. (2011). Towards a pragmatics of epistemology, methodology, and social theory. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 441-464). Guildford Press.
- Duan, S. (2018). Effects of enhancement techniques on L2 incidental vocabulary learning. *English Language Teaching*, 11(3), 88-101. <http://doi.org/10.5539/elt.v11n3p88>
- Dube, S. R. (2018). How childhood experiences contribute to the education-health link. *The Conversation*. Article 89069.
- Duff, D., Tomblin, B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew effect. *Journal of Speech Language and Hearing Research*, 58(3), 853-864. https://doi.org/10.1044/2015_JSLHR-L-13-0310
- Duff, P. A. (2008). *Case study research in applied linguistics*. Lawrence Erlbaum.
- Dunbar, Jr., C., Rodriguez, D., & Parker, L. (2003). Race, subjectivity, and the interview process. In J. A. Holstein & J. F. Gubrium (Eds.), *Inside interviewing: New lenses, new concerns* (pp. 131-152). SAGE.
- Durso, F. T., & Shore, W. J. (1991). Partial knowledge of word meanings. *Journal of Experimental Psychology: General*, 120(2), 190-202. <https://doi.org/10.1037/0096-3445.120.2.190>
- Ebrahimi, F., Namaziandost, E., Ziafar, M., & Ibna Seraj, P. M. (2021). The effect of teaching formulaic expressions through contrastive lexical approach on Iranian pre-intermediate

- EFL learners' writing skill. *Journal of Psycholinguistic Research*, 50, 1087-1105.
<https://doi.org/10.1007/s10936-021-09778-z>
- Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language and Teaching Research*, 16(2), 227-252. <https://doi.org/10.1177/136216881143377>
- Ehsanzadeh, S. J. (2012). Depth versus breadth of lexical repertoire: Assessing their roles in EFL students' incidental vocabulary acquisition. *TESL Canada Journal*, 29(2), 24-41.
<https://doi.org/10.18806/tesl.v29i2.1098>
- Eisenkraemer, R. E., Jaeger, A., & Stein, L. M. (2013). A systematic review of the testing effect in learning. *Paidéia*, 23(56), 397-406. <https://doi.org/10.1590/1982-43272356201314>
- El-Dakhs, D. A. S. (2017). The effect of the explicit instruction of formulaic sequences in pre-writing vocabulary activities on foreign language writing. *International Journal of Applied Linguistics and English Literature*, 6(4), 21-31.
<https://doi.org/10.7575/aiac.ijalel.v.6n.4p.21>
- El-Dakhs, D. A., ElHajj, H., & Al-Haqban, J. N. (2018). The effect of word type on the L1 support for L2 vocabulary learning: The case of Arab EFL learners. *International Journal of English Linguistics*, 8(4), 25-38. <https://doi.org/10.5539/ijel.v8n4p25>
- Elementary and Secondary Education Act, Title III Part A, 20 § 3102 (2001).
<https://www2.ed.gov/policy/elsec/leg/esea02/pg40.html#sec3102>
- Elhami, G., & Mahmood, K. (2015). The effect of making weblogs with respect to gender on elementary Iranian EFL learners' vocabulary development. *International Journal of Educational Investigations*, 2(1), 226-238.

- Ellis, L., Hershberger, S., Field, E., Wersinger, S., Pellis, S., Geary, D., Palmer, C., Hoyenga, K., Hetsroni, A., & Karadi, K. (2008). *Sex differences*. Psychology Press.
- Ellis, N. & Wulff, S. (2014). Usage-based approaches in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp.75-93). Routledge.
- Ellis, N. (2011). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In M. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 139-158). Blackwell Publishing.
- Ellis, N. C. (2003). Incidentals that build fluency: Optimal word processing and its implications for vocabulary acquisition. *New Zealand Studies in Applied Linguistics*, 9, 1-18.
- Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 164–194. <https://doi.org/10.1093/applin/aml015>
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second-language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 41(3), 375-396. <https://doi.org/10.1002/j.1545-7249.2008.tb00137.x>
- Erkuş, E. C. & Purutçuoğlu, V. (2021). Outlier detection and quasi-periodicity optimization algorithm: Frequency domain based outlier detection (FOD). *European Journal of Operational Research*, 291(2), 560-574. <https://doi.org/10.1016/j.ejor.2020.01.014>
- Every Student Succeeds Act, 20 U.S.C. § 6301 (2015).
<https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf>
- Fairclough, N. (2013). *Critical language awareness*. Routledge.
<https://doi.org/10.4324/9781315845661>

- Far, M. T. (2016). The effects of text type, text length and text difficulty on vocabulary retention through glossing. *The Journal of Language Learning and Teaching*, 6(1), 90-104.
- Farag, A. (2021). The fear of multiple truths: On teaching about racism in a predominantly White school. *Phi Delta Kappan*, 102(5), 18-23.
- Fardad, Z., Koosha, M., & Shafiee, S. (2015). Relationship between EFL learners' multiple intelligence scores, gender, and their vocabulary knowledge. *Research Journal of Recent Sciences*, 4(12), 29-36. <https://doi.org/10.17507/tpls.0504.16>
- Fatemi, A. H., & Asghari, A. (2016). Attribution theory, personality traits, and gender differences among EFL learners. *Theory and Practice in Language Studies*, 6(3), 518-524. <https://doi.org/10.17507/tpls.0603.09>
- Fathi, M., & Sarkhosh, M. (2019). A study of the impact of input modification by using glosses on vocabulary acquisition and retention. *International Journal of English Language & Translation Studies*, 7(1), 55-67.
- Feagin, J. R. (2004). Toward an integrated theory of systemic racism. In M. Krysan & A. E. Lewis (Eds.), *The changing terrain of race and ethnicity* (pp. 203-223). Sage.
- Feak, C. B., & Swales, J. M. (2009). *Telling a research story: Writing a literature review*. The University of Michigan.
- Feilzer, M. Y. (2009). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4(1) 6–16. <https://doi.org/10.1177/1558689809349691>
- Fernald A., Marchman, V. A., & Weisleder, A. (2012). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16(2), 234-248. <https://doi.org/10.1111/desc.12019>

- Fernández Fontecha, A. (2015). Individual and contextual variables in EFL receptive vocabulary: Analysis of gender effect and motivation of the CLIL primary graders. *Odisea, 15*, 57-72. <https://doi.org/10.25115/odisea.v0i15.285>
- Fine, M., Weis, L., Weseen, S., & Wong, L. (2003). For whom? Qualitative research, representations, and social responsibilities. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research* (2nd ed., pp. 167-207). SAGE.
- Fish, R. E. (2019). The racialized construction of exceptionality: Experimental evidence of race/ethnicity effects on teachers' interventions. *Social Science Research, 62*, 317-334. <https://doi.org/10.1016/j.ssresearch.2016.08.007>
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*, 39-50. https://doi.org/10.1207/S15326985EP3501_5
- Flanagin, A., Frey, T., & Christiansen, S. L. (2021). Updated guidance on the reporting of race and ethnicity in medical and science journals. *Journal of the American Medical Association, 326*(7), 621-627. <https://doi.org/10.1001/jama.2021.13304>
- Flowerdew, J. (2011). Corpora in language teaching. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 327-350). Wiley Blackwell.
- Forum on Child and Family Statistics (2021). Youth enrolled in school and working: Percentage of youth ages 16–19 who are enrolled in school and working by age, gender, and race and Hispanic origin, 2000–2021. Retrieved from <https://www.childstats.gov/americaschildren/tables/ed5b.asp>
- Fox, N. J., & Alldred, P. (2018). Mixed methods, materialism and the micropolitics of the research-assemblage. *International Journal of Social Research Methodology, 21*(2), 191-204. <https://doi.org/10.1080/13645579.2017.1350015>

- Freeman, M., deMarrias, K., Preissle, J., Routston, K., & St. Pierre, E. A. (2007). Standards of evidence in qualitative research: An incitement to discourse. *Educational Researcher*, 36(1) 25-32. <https://doi.org/10.3102/0013189X06298009>
- Frishkoff, G. A., Collins-Thompson, K., Perfetti, C. A., & Callan, J. (2008). Measuring incremental changes in word knowledge: Experimental validation and implications for learning and assessment. *Behavior Research Methods*, 40(4), 907-925. <https://doi.org/10.3758/BRM.40.4.90>
- Fry, R. (2008). *The role of schools in the Bilingual learner achievement gap* (ED502050). Pew Hispanic Center. <http://files.eric.ed.gov/fulltext/ED502050.pdf>
- Garcia, E. E., Lawton, K., & Diniz de Figueiredo, E. H. (2010). *The education of bilingual learners in Arizona: A legacy of persisting achievement gaps in a restrictive language policy climate*. The Civil Rights Project. <https://www.civilrightsproject.ucla.edu/research/k-12-education/language-minority-students/the-education-of-english-language-learners-in-arizona-a-legacy-of-persisting-achievement-gaps-in-a-restrictive-language-policy-climate/garcia-az-ell-gaps-2010.pdf>
- Garner, A. S., Shonkoff, J. P., Siegel, B. S., Dobbins, M. I., Earls, M. F., McGuinn, L., Pascoe, J., & Wood, D. L. (2012). Early childhood adversity, toxic stress, and the role of the pediatrician: Translating developmental science into lifelong health. *Pediatrics*, 129(1), 224-231. <https://doi.org/10.1542/peds.2011-2662>
- Gass, S. M., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 182-206). Routledge.

- Gaynor, T. S. (2021). Coming to terms: Teaching systemic racism and (the myth of) white supremacy. *Journal of Public Affairs Education*, 28(2), 211-225.
<https://doi.org/10.1080/15236803.2021.1994326>
- Georgia Department of Education (2021). <https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/ACCESS-for-ELLs.aspx>
- Ghorbani, M. R., & Rahmandoost, M. (2012). Higher task-induced involvement load enhances students' EFL vocabulary learning. *Journal of Language Teaching and Research*, 3(6), 1202-1207. <https://doi.org/10.4304/jltr.3.6.1202-1207>
- Gibson, C. (2016). Bridging Bilingual learner achievement gaps through effective vocabulary development strategies. *English Language Teaching*, 9(9), 134-138.
<https://doi.org/10.5539/elt.v9n9p134>
- Gil, N. N., & Caro, E. M. (2019). Lexical bundles in learner and expert academic writing. *Bellaterra Journal of Teaching & Learning Language & Literature*, 12(1), 65-90.
<https://doi.org/10.5565/rev/jtl3.794>
- Gilmore, A., & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. *English for Specific Purposes*, 51, 1-17.
<https://doi.org/10.1016/j.esp.2018.02.002>
- Gluszek, A., Anna-Kaisa, A., Dovidio, J. F. (2010). Social psychological orientations and accent strength. *Journal of Language and Social Psychology*, 30(1), 28-45.
- Gobet, F. (2017). Entrenchment, Gestalt formation, and chunking. In H. J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (p. 245–267). De Gruyter Mouton. <https://doi.org/10.1037/15969-012>

- Godfroid, A., Boer, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition, 35* (3), 483-517.
- Gohar, M. J., Rahmanian, M., & Soeimani, H. (2018). Technique feature analysis or involvement load hypothesis: Estimating their predictive power in vocabulary learning. *Journal of Psycholinguistic Research, 47*, 859–869. <https://doi.org/10.1007/s10936-018-9568-5>
- Golaghaei, N., Yamini, M., (2014). Impact of metalinguistic English vocabulary knowledge and lexical inferencing on EFL learners' lexical knowledge considering the cross-linguistic issue of L1 lexicalization. *Journal of Teaching Language Skills, 33*(2), 27-53. <https://doi.org/10.22099/jtls.2014.2403>
- Goldenberg, C., Gallimore, R., & Reese, L. (2005). Using mixed methods to explore Latino children's literacy development: Discovering successful pathways in children's development. In T. S. Weisner (Ed.), *Mixed methods in the study of childhood and family life* (pp. 21-46). University of Chicago Press.
- González, M. C. (2017). The contribution of lexical diversity to college-level writing. *TESOL Journal, 8*(4), 899-919. <https://doi.org/10.1002/tesj.342>
- Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady & T. Huckin (Eds.) *Second language vocabulary acquisition* (pp. 98-122). Cambridge University Press.
- Graham, S., & Herbert, M. (2010). *Writing to read: Evidence for how writing can improve reading*. Carnegie Corporation.
- Granger, S. (2014). A lexical bundle approach to comparing languages: Stems in English and French. *Languages in Contrast, 14*(1), 58-72. <https://doi.org/10.1075/lic.14.1.04gra>

- Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27-49). John Benjamins Publishing Company.
- Grantmakers for Education (2013). *Educating bilingual learners: Grantmaking strategies for closing America's other achievement gap*. Kellogg Foundation.
http://www.edfunders.org/sites/default/files/Educating%20English%20Language%20Learners_April%202013.pdf
- Graves, J. M., Machelprang, J. L., Barbosa-Leiker, C., Miller, M. E., & Li, A. Y. (2017) Quality of life among working and non-working adolescents. *Quality of Life Research*, 26, 107-120. <https://doi.org/10.1007/s11136-016-1376-5>
- Grenberg, J. (2005). *Kant and the ethic of humility: A story of dependence, corruption, and virtue*. Cambridge University Press.
- Greving, S., Lenhard, W., & Richter, T. (2022). The testing effect in university teaching: Using multiple-choice testing to promote retention of highly retrievable information. *Teaching of Psychology*. Advance Online Publication.
<https://doi.org/10.1177%2F00986283211061204>
- Griers, S. (2008). Phraseology and linguistic theory. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 3-25). John Benjamins.
- Gu, P. W. (2002). Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELC Journal*, 33(1), 35-54. <https://doi.org/10.1177/003368820203300102>
- Guthrie, J. t., Wigfield, Al, & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. L. Christenson, A. L., Reschly, & C. Wylie (Eds.),

- Handbook of research on student engagement* (pp. 601-634). Springer.
https://doi.org/10.1007/978-1-4614-2018-7_29
- Hall, J. N. (2013). Pragmatism, evidence, and mixed methods evaluation. *New Directions for Evaluation, 138*, 15-26. <https://doi.org/10.1002/ev.20054>
- Hansen Edwards, J. G. (2017). Defining ‘native speaker’ in multilingual settings: English as a native language in Asia. *Journal of Multilingual and Multicultural Development, 38*(9), 757-771. <https://doi.org/10.1080/01434632.2016.1257627>
- Harklau, L., & Pinnow, R. (2009). Adolescent second-language writing. In L. Christenbury, R. Bomer, & P. Smagorinsky (Eds.), *Handbook of adolescent literacy research* (pp. 126–139). Guilford Press.
- Harrison, J., & Lakin, J. (2018). Mainstream teachers’ implicit beliefs about bilingual 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>
- Hatch, (1992). *Discourse and language education*. Cambridge University Press.
- Hawley, C. E., Cardoso, E., & McMahon, B. T. (2013). Adolescence to adulthood to STEM education and career development: The experience of students at the intersection of underrepresented minority status and disability. *Journal of Vocational Rehabilitation, 39*(3), 193-204. <https://doi.org/10.3233/JVR-130655>
- Hazrat, M. (2015). The effects of task type and task involvement load on vocabulary learning. *Journal of Education, 20*(2), 79-92.
- Heid, U. (2008). Computational phraseology. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 337-360). John Benjamins Publications.

- Heidari-Shahreza, M. A., & Tavakoli, M. (2016). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition by Iranian EFL learners. *The Language Learning Journal*, 44(1), 17-32. <https://doi.org/10.1080/09571736.2012.708051>
- Heidari-Shahreza, M. A., Moinzadeh, A., & Hossein, B. (2014). A comparative investigation into effects of L1 lexicalization and cultural loadedness on incidental vocabulary acquisition and retention. *International Journal of Research Studies in Language Learning*, 3(5), 83-96. <https://doi.org/10.5861/ijrsl.2014.698>
- Henderson, H. (2018). Difficult questions of difficult questions: The role of the researcher and transcription styles. *International Journal of Qualitative Studies in Education*, 31(2), 143-157. <https://doi.org/10.1080/09518398.2017.1379615>
- Henriksen, B. (2013). Research on L2 learners' collocational competence and development: A progress report. In C. Bardel, C. Linquist, & B. Laufer (Eds.), *L2 vocabulary and use: New perspectives on assessment and corpus analysis* (pp. 29-56). EuroSLA.
- Herdina P. (2017). Entrenchment, embeddedness, and entanglement: A dynamic complexity view. In H. J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 387-407). De Gruyter Mouton. <https://doi.org/10.1037/15969-018>
- Hernandez, D. J., Denton, N. A., Macartney, S. E. (2010). Children of immigrants and the future of America. In E. L. Grigorenko & R. Takanishi (Eds.), *Immigration, diversity, and education* (pp. 7-25). Routledge.
- Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. Guilford Press.

- Hilden, K., & Pressley, M. (2011). Verbal protocols of reading. In M. H. Mallette & N. K. Duke (Eds.), *Literacy research methodologies* (2nd ed., pp. 427-440). The Guilford Press.
- Hilgard, E. R. (1956). *Theories of learning* (2nd ed.). Appleton-Century-Crofts.
- Hilpert, M., & Diessel, H. (2017). Entrenchment in construction grammar. In H. J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 57–74). De Gruyter Mouton.
- Hitchcock, J. H., & Onwuegbuzie, A. J. (2019). Developing mixed methods crossover analysis approaches. *Journal of Mixed Methods Research*, 14(1), 63-83.
<https://doi.org/10.1177/1558689819841782>
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368-1378.
- Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, 73(2), 418-433. <https://doi.org/10.1111/1467-8624.00415>
- Hoffman, B. (2015). *Motivation for learning and performance*. Elsevier Academic Press.
- Holstein, J. A., & Gubrium, J. F. (2000). *The self we live by: Narrative identity in a postmodern world*. Oxford University Press.
- Hopkins, J. S., Loftus, P. (2020, November 30). Moderna's covid-19 vaccine: When could it be ready and everything else you need to know. *Wall Street Journal*.
<https://www.wsj.com/articles/modernas-covid-19-vaccine-when-could-it-be-ready-and-everything-else-you-need-to-know-11605548631>
- Hopp, H., Kieseier, T., Volgebacher, M., & Thoma, D. (2018). L1 effects in the early L3 acquisition of vocabulary and grammar. In A. Bonnet & P. Siemund (Eds.), *Foreign*

language education in multilingual classrooms. John Benjamins.

<https://doi.org/10.1075/hslid.7.14hop>

- Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207-223.
- Hos, R. (2016). The lives, aspirations, and needs of refugee and immigrant students with interrupted formal education (SIFE), in a secondary newcomer program. *Urban Education*, 55(7), 1021-1044. <https://doi.org/10.1177/0042085916666932>
- Hsu, H. (2018). Incidental professional vocabulary acquisition of EFL business learners: Effect of captioned video with glosses. *JALT CALL Journal*, 14(2) 119-142.
<https://doi.org/10.1093/elt/ccs067>
- Hsu, W. (2014). The most frequent opaque formulaic sequences in English-medium college textbooks. *System*, 47, 146-161. <https://doi.org/10.1016/j.system.2014.10.001>
- Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 13-23.
<https://doi.org/10.1016/j.system.2015.06.011>
- Hyland, K. (1998) *Hedging in scientific research articles*. John Benjamins Publishing
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4-21. <https://doi.org/doi:10.1016/j.esp.2007.06.001>
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21. <https://doi.org/10.1016/j.esp.2007.06.001>
- Hyland, K. (2011). Academic discourse. In K. Hyland & B. Paltridge (Eds.), *The Bloomsbury companion to discourse analysis* (pp. 171-184). Bloomsbury.

- Ishikawaa, Y. (2014). Gender differences in vocabulary use in essay writing by university students. *Procedia-Social and Behavioral Sciences*, 192, 593-500.
<https://doi.org/10.1016/j.sbspro.2015.06.078>
- James, W. (1890). *Principles of psychology*. Henry Holt and Company.
- James, W. (1995) *Pragmatism*. Dover. (Original work published 1907)
- Jesson, R., Mcnaughton, S., & Parr, J. M. (2011). Drawing on intertextuality in culturally diverse classrooms: Implications for transfer of literacy knowledge. *English Teaching: Practice and Critique*, 10(2), 65-77.
- Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *SSLA*, 24, 617-637. <https://doi.org/10.1017.S0272263102004047>
- Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(iii), 433-445.
<https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Jiménez-Castellanos, O. H., & García. D. (2017). School expenditures and academic achievement differences between high-ELL-performing and low-ELL-performing high schools. *Bilingual Research Journal*, 40(3), 318-330.
<https://doi.org/10.1080/15235882.2017.1342717>
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357-377.
- Johnson, D. B. (1972). Computer frequency control of vocabulary in language learning reading materials. *Instructional Science*, 1(1), 121-131. <https://doi.org/10.1007/BF00053972>
- Johnson, P. L., & Fay, L. C. (1950). The Johnson-Neyman technique: Its theory and application. *Psychometrika*, 15(4) 349–367. <https://doi.org/10.1007/BF02288864>

- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
<https://doi.org/10.3102/0013189X033007014>
- Jones, D. C. (2002). John Dewey and Peter Elbow: A pragmatist revision of social theory and practice. *Rhetoric Review*, 21(3), 264-281. https://doi.org/10.1207/S15327981RR2103_4
- Jung, J. (2016). Effects of glosses on learning of L2 grammar and vocabulary. *Language Teaching Research*, 20(1), 92-112. <https://doi.org/10.1177/1362168815571151>
- Kachru, B. B. (1982). Standard, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge University Press.
- Kangas S. E. (2017). 'That's where the rubber meets the road': the intersection of special education and dual language education. *Teachers College Record*, 119(7), 1-36.
- Kant, I. (2012). *Groundwork of the metaphysic of morals* (M. Gregor, Trans.). Cambridge University Press. (Original work published 1785).
- Kaplan, I., Stolk, Y., Valibhoy, M., Tucker, A., Baker, J. (2015). Cognitive assessment of refugee children: Effects of trauma and new language acquisition. *Transcultural Psychiatry*, 53(1), 81-109. <https://doi.org/10.1177/1363461515612933>
- Karakoç D. & Köse, G. D. (2017). The impact of vocabulary knowledge on reading, writing, and proficiency scores of EFL learners. *Journal of Language and Linguistic Studies*, 13(1), 352-378.
- Kaushanskaya M., Marian V., & Yoo, J. (2011). Gender differences in adult word learning. *Acta Psychologica*, 137, 24-35. <https://doi.org/10.1016/j.actpsy.2011.02.002>

- Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and Individual Differences, 27*, 82-89.
<https://doi.org/10.1016/j.lindif.2013.07.002>
- Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences, 8*(9), 1-17. <https://doi.org/10.3390/socsci8090255>
- Kayumova, S., Karsli, E., Alleksaht-Snider, M., & Buxton, C. (2015). Latina mothers and daughters: Ways of knowing, being, and becoming in the context of bilingual family science workshops. *Anthropology & Education Quarterly, 46*(3), 260–276.
<https://doi.org/10.1111/aeq.12106>
- Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing, 25*(1), 4-22.
<https://doi.org/10.1016/j.jslw.2014.05.005>
- Kelley, J. G., Lesaux, N. K., Kieffer, M. J., & Faller, S. E. (2010). Effective academic vocabulary instruction in the urban middle school. *The Reading Teacher, 64*(1), 5-14.
<https://doi.org/10.1598/RT.64.1.1>
- Kennison, S. M. (2013). *Introduction to language development*. SAGE.
- Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures design: A review. *British Journal of Mathematical and Statistical Psychology, 54*, 1-20.
<https://doi.org/10.1348/0007110011159357>
- Khan, M. A. (2019). New ways of using corpora for teaching vocabulary and writing in the ESL classroom. *ORTESOL Journal, 36*, 17-24.

- Khatib, M., & Faruji, L. F. (2012). The impact of using story maps as graphic organizers on development of vocabulary learning of EFL learners. *Journal of Language Teaching and Research, 3*(6), 1114-1121. <https://doi.org/10.4304/jltr.3.6.1114-1121>
- Khonamri, F., & Roostae, S. (2014). Does extensive reading combined with form-focused or meaning-focused activities affect lexical collocational knowledge of Iranian learners? *Theory and Practice in Language Studies, 4*(5), 1038-1044. <https://doi.org/10.4304/tpls.4.5.1038-1044>
- Khoshsim, H., & Eskandari, Z. (2017). Task effectiveness predictors: Technique feature analysis versus involvement load hypothesis. *Iranian Journal of English for Academic Purposes, 6*(2), 50-69.
- Kiliç, M. (2019). Vocabulary knowledge as a predictor of performance in writing and speaking: A case of Turkish EFL learners. *PASAA, 57*, 133-164.
- Kim, Y., & Taguchi, N. (2015). Promoting task-based pragmatics instruction in EFL classroom contexts: The role of task complexity. *Modern Language Journal, 99*(4), 656-677. <https://doi.org/10.1111/modl.12273>
- Kimmel, M., & Gordon, K. M. (2019) Gender. In M. Ryan (Ed.), *Core concepts in sociology*. John Wiley & Sons.
- Kivunja, C., & Kuvini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education, 6*(5), 26-41. <https://doi.org/10.5430/ijhe.v6n5p26>
- Kobayashi, K., & Little, A. (2020). Gender differences in the effects of explicit vocabulary learning strategies training. *Journal of Second Language Teaching & Research, 8*(1), 23-51.

- Koopman, C., & Garside, D. (2019). Transition, action and education: Redirecting pragmatist philosophy of education. *Journal of Philosophy of Education*, 53(4), 734-747.
<https://doi.org/10.1111/1467-9752.12399>
- Kopat, G., & Hlebec, V. (2020). Quality guidelines for mixed methods research in intervention studies: A conceptual model. *Advances in Methodology & Statistics / Metodoloski zvezki*, 17(2), 1-29.
- Koro-Ljungberg, M., Yendol-Hoppey, Smith, J. J., & Hayes, S. B. (2009). (E)pistemological awareness, instantiation of methods, and uninformed methodological ambiguity in qualitative research projects. *Educational Researcher*, 38(9), 687-699.
<https://doi.org/10.3102/0013189X09351980>
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Prentice-Hall.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal*, 73, 440-464.
- Kremmel, B. (2020). Measuring vocabulary learning progress. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 406-418). Routledge.
- Kubota, R. (2003). New approaches to race, class, and gender in second language writing. *Journal of Second Language Writing*, 12(1), 31-47. [https://doi.org/10.1016/S1060-3743\(02\)00125-X](https://doi.org/10.1016/S1060-3743(02)00125-X)
- Kubota, R. (2015). Race and language learning in multicultural Canada: Towards critical antiracism. *Journal of Multilingual & Multicultural Development*, 36(1), 3-12.
<https://doi.org/10.1080/01434632.2014.892497>
- Kweon, S., Kim, H. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, 20(2), 191-215.

- Kwon, S. (2009). Lexical richness in L2 writing: How much vocabulary do L2 learners need to use? *English Teaching*, 64(3), 155-174.
- Ladner, S. (2019). *Mixed methods: A short guide to applied mixed methods research*.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, 35(7), 3-12.
<https://doi.org/10.3102%2F0013189X035007003>
- Ladson-Billings, G. (2021). *Critical race theory in education: A scholar's journey*. Teachers' College Press.
- Lakoff, R. (1972). Language in context. *Language*, 48(4), 907-927.
- Lakoff, R. (2003). Language, gender, and politics: Putting "women" and "power" in the same sentence In J. Holmes & M. Meyerhoff (Eds.), *The handbook of language and gender* (pp. 160-178). <https://doi.org/10.1002/9780470756942>
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211-240. <https://doi.org/10.1037/0033-295X.104.2.211>
- Lane, C., Kim, M., & Schrynemakers, I. (2019). American acculturation and academic performance: Analyzing the relationship between years living in the United States and freshman composition grades at a diverse community college. *Community College Journal of Research and Practice*, 44(9), 677-689.
<https://doi.org/10.1080/10668926.2019.1633971>
- Lantolf, J. P. (1994). Sociocultural theory and second language learning: Introduction to the special issue. *The Modern Language Journal*, 78(4), 418-420.
<https://doi.org/10.1111/j.1540-4781.1994.tb02058.x>

- Larson, S. C. (2014) Exploring the roles of the generative vocabulary matrix and academic literacy engagement of ninth grade biology students. *Literacy Research and Instruction*, 53(4), 287-325. <https://doi.org/10.1080/19388071.2014.880974>
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady, & T. Huckin (Eds.) *Second Language Vocabulary Acquisition* (pp. 20-34). Cambridge University Press.
- Laufer, B., & Goldstein, Z. (2004) Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language learning*, 54(3), 399-436.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
<https://doi.org/10.1093/applin/22.1.1>
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, 365–391.
<https://doi.org/10.1111/0023-8333.00046>
- Laufer, B., & Rozovski-Roitblat, B (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence, and their combination. *Language Teaching Research*, 15(4), 391-411. <https://doi.org/10.1177/1362168811412019>
- Lease, S. H., Shuman, W. A., & Gage, A. N. (2020). Female and male coworkers: Masculinity, sexism, and interpersonal competence at work. *Psychology of Men & Masculinities*, 21(1), 139–147. <https://doi.org/10.1037/men0000218>
- Lee, S. (2020). Examining the roles of aptitude, motivation, strategy use, language processing experience, and gender in the development of the breadth and depth of EFL Learners' vocabulary knowledge. *SAGE Open*, 1-15. <https://doi.org/10.1177/2158244020977883>

- Lee, S. H. & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly*, 40(2), 295-320.
<https://doi.org/10.2307/40264524>
- Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System*, 31(4), 537-561. <https://doi.org/10.1016/j.system.2003.02.004>
- Lee, S. J. (2006). Additional complexities: Social class, ethnicity, generation, and Asian American student experiences. *Race, Ethnicity & Education*, 9(1), 17-28.
<https://doi.org/10.1080/13613320500490630>
- Lenders, O. (2008). Electronic glossing—Is it worth the effort? *Computer Assisted Language Learning*, 21(5), 457-481. <https://doi.org/10.1080/09588220802447933>
- Lessard-Clouston, M. (2012). Technical vocabulary use in English-Medium disciplinary writing: A native/non-native case study. *The Linguistics Journal*, 6(1), 127-150.
- Leung, C., Harris, R., & Rampton, B. (1997). The idealized native speaker, reified ethnicities, and classroom realities. *TESOL Quarterly*, 31(3), 543-560.
<https://doi.org/10.2307/3587837>
- Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: Single words and collocations. In C. Bardel, C. Linquist, & B. Laufer, (Eds.), *L2 vocabulary and use: New perspectives on assessment and corpus analysis* (pp. 127-147). European Second Language Association.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. LTP Teacher Training.
- Liggett, T. (2014). The mapping of a framework: Critical race theory and TESOL. *The Urban Review*, 46(1), 112–124. <https://doi.org/10.1007/s11256-013-0254-5>

- Limes-Taylor Henderson, K., & Esposito, J. (2017). Using others in the nicest way possible: On colonial and academic practice(s), and an ethic of humility. *Qualitative Inquiry*, 25, 9-10. <https://doi.org/10.1177/1077800417743528>
- Lin, I. C., & Kawai, G. (2016). Identifying and activating receptive vocabulary by an online vocabulary survey and an online writing task. In S. Papadima-Sophocleous, L. Bradley & S. Thouësny (Eds.), *CALL communities and culture – short papers from EUROCALL 2016* (pp. 271-276). <https://doi.org/10.14705/rpnet.2016.eurocall2016.574>
- Liu, D. (2012). The most frequently used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, 31, 25-35.
- Liu, F., Chen, Y. (2019). Investigating the intrinsic motivation and native language effect in second language learning. *Sino-US English Teaching*, 16(5), 197-202. <https://doi.org/10.17265/1539-8072/2019.05.002>
- Liu, S. (2017). An experimental research on effects of types of glossing on incidental vocabulary acquisition through reading. *Journal of Language Teaching & Research*, 8(4), 782-793. <http://doi.org/10.17507/j>
- Liu, S. (2018). Review on incidental vocabulary acquisition through reading. *Theory & Practice in Language Studies*, 8(6), 702-711. <http://doi.org/10.17507/tppls.0806.22>
- Lofgren, A. (2022). Unraveling contradictions: Which glosses facilitate reading comprehension among ELLs, and why? *Journal of Language Teaching and Research*, 13(1), 1-11. <https://doi.org/10.17507/jltr.1301.01>
- Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences*, 379(1), 259-278. <https://doi.org/10.1111/j.1749-6632.1981.tb42014.x>

- Lorde, A. (1984). *Sister outsider*. The Crossing Press.
- Lu, C., Rodriguez, S. L., & Bukoski, B. (2021). Exploring how gay Latinx men cope in college using emotion regulation. *Journal of Hispanic Higher Education*, 20(4), 438-452.
<https://doi.org/10.1177/1538192719888755>
- Luke, A. (2019). Regrounding critical literacy: Representation, facts, and reality. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed., pp. 349-361). Routledge.
- Ma, Q. (2013). Matching vocabulary learning process with learning outcome in L2 academic writing: An exploratory case study. *Linguistics and Education*, 24(2), 237-246.
- Maghsoudi, M., & Shamaei, F. (2016). The impact of Iranian EFL learners' gender on their vocabulary retention. *International Academic Journal of Humanities*, 3(2), 101-106.
- Maher, J. M., Markey, J. C., & Ebert-May, D. (2017). The other half of the story: Effect size analysis in quantitative research. *Life Sciences Education*, 12(3).
<https://doi.org/10.1187/cbe.13-04-008>
- Mallette, M. H., Duke, N. K., Strachan, S. L., Waldron, C. H., & Watanabe, L. M. (2013). Synergy in literacy research methodology. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 91-128). International Reading Association.
- Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*, 102(3), 701-711. <https://doi.org/10.1037/a0019135>
- Manihuruk, D. (2020). The correlation between EFL students' vocabulary knowledge and reading comprehension: A case study at the English Education Department of Universitas

- Kristen Indonesia. *Journal of English Teaching*, 6(1), 86-95.
<https://doi.org/10.33541/jet.v6i1.1264>
- Manuel, N. N. (2017). Evaluating vocabulary learning strategies (VLS): Gender differences, the most and least used (VLS) among Angolan EFL Students at the Faculty of Arts. *International Journal of Scientific Research in Education*, 10(5), 483-504.
- Maroun, J. (2018). The predictive power of out-of-school community and family level demographic factors on district level student performance on the New Jersey PARCC in Algebra I and Grade 10 English Language Arts/Literacy [Dissertation, Seton Hall University]. ProQuest Dissertations Publishing.
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299-320. <https://doi.org/10.1093/applin/ams010>
- Masrai, A., & Milton, J. (2015). The impact of L1 lexical organisation in L2 vocabulary acquisition. *Language in Focus Journal*, 1(1), 15-34. <https://doi.org/10.1515/lifij-sal-2015-0002>
- McBreen, M., & Savage, R. (2021). The impact of motivational reading instruction on the reading achievement and motivation of students: A systematic review and meta-analysis. *Educational Psychology Review*, 33, 1125-1163.
- McDevitt, J. A. (2021). Educating and empowering teen activists in public libraries: A case study of the impact of reading on young adult social justice actions [Master's thesis, University of Alberta]. Education and Research Archive. <https://doi.org/10.7939/r3-ndwc-vw83>
- McGrail, E. (2017). Seeking the comprehensive gestalt in student identities: A means to social justice. *Ubiquity: The Journal of Literature, Literacy, and the Arts, Research Strand*, 4(2), 1-5.

- McHugh, M. L. (2015). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276-282.
- McKeown, M. G., Deane, P. D., Scott, J. A., Krovetz, R., & Lawless, R. R. (2017). *Vocabulary assessment to support instruction: Building rich word-learning experiences*. The Guilford Press.
- McKim, C. A. (2017). The value of mixed methods research: A mixed methods study. *Journal of Mixed Methods Research*, 11(2), 202-222. <https://doi.org/10.1177/1558689815607096>
- McLaughlin, H. J., & Pettit, S. K. (2014). “What kind of possibilities do we have?”: Educators’ complex images of Latino immigrant students and families. *Journal of Contemporary Research in Education*, 2, 115-127.
- McLaughlin, K., & Coderre, S. (2015). The potential and conditional benefits of retrieval practice on learning. *Advances in Health Science Education*, 20, 321-324. <https://doi.org/10.1007/s10459-014-9529-1>
- Meara, P. (1990). A note on passive vocabulary. *Second Language Research*, 6(2), 150-154. <https://doi.org/10.1177/02676583600204>
- Meisel, J. M. (2011). *First and second language acquisition: parallels and differences*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511862694>
- Mendez, J. J., Bauman, S., & Guillory, R. M. (2012). Bullying of Mexican immigrant students^[1]_{SEP} by Mexican American students: An examination of intracultural bullying. *Hispanic Journal of Behavioral Sciences*, 34(2), 279–304. <https://doi.org/10.1177/0739986311435970>

- Mertens (2007) Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research*, 1(3), 212-225.
<https://doi.org/10.1177/1558689807302811>
- Mertens et al.'s (2010) Mertens, D. M., Bledsoe, K. L., Sullivan, M., & Wilson, A. (2010). Utilization of mixed methods for transformative purposes. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavior research* (2nd ed.; pp. 193-214). SAGE.
- Merton, R. (1968). *Social theory and social structure*. The Free Press.
- Milton, J. (2013) Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 57-78). European Second Language Association.
- Milton, J. (2022). Vocabulary knowledge and educational attainment. In H. Mohebbi & C. Coombe (Eds.), *Research questions in language education and applied linguistics* (pp. 267-271). Springer. https://doi.org/10.1007/978-3-030-79143-8_48
- Min, H. T. (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>
- Montero-SaizAja, A. (2021). Gender-based differences in EFL learners' language learning strategies and productive vocabulary. *Theory and Practice of Second Language Acquisition*, 7(2), 83-107. <https://doi.org/10.31261/TAPSLA.8594>

- Monzó, L. D., & Rueda, R. (2009). Passing for English fluent: Latino immigrant children masking language proficiency. *Anthropology & Education Quarterly*, 40(1), 20-40. <https://doi.org/10.1111/j.1548-1492.2009.01026.x>
- Moore, B. A., & Klingner, J. K. (2014). Considering the needs of bilingual learner populations: An examination of population validity of reading intervention research. *Journal of Learning Disabilities*, 47(5), 391-408. <https://doi.org/10.1177/0022219412466702>
- Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry*, 20(8), 1045-1053. <https://doi.org/10.1177/1077800413513733>
- Morgan, D. L. (2020). Pragmatism as a basis for grounded theory. *The Qualitative Report*, 25(1), 64-73.
- Morse, A. (2005). *A look at immigrant youth: Prospects and promising practices*. National Conference of State Legislatures (NCSL). <https://www.ncsl.org/research/immigration/a-look-at-immigrant-youth-prospects-and-promisin.aspx>
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavior research* (2nd ed., pp. 189-208). SAGE.
- Mortimer, J. T. (2010). The benefits and risks of adolescent employment. *Prev Res*, 17(2), 8-11.
- Motulsky, H. (2021). *QuickCalcs* [Computer software]. GraphPad. <https://www.graphpad.com/quickcalcs/randomize1.cfm>
- Muncie, J. (2002). Process writing and vocabulary development: Comparing lexical frequency profiles across drafts. *System*, 30(2), 225-235. [https://doi.org/10.1016/S0346-251X\(02\)00006-4](https://doi.org/10.1016/S0346-251X(02)00006-4)

- Na, W. (2016). Gender differences in the use of English vocabulary learning strategies in Chinese senior high schools. *Studies in Literature and Language* 12(4), 58-62.
<https://doi.org/10.3968/8225>
- Nagy (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). Guilford Press.
- Nagy, W. E., & Herman, P. (1985). Incidental vs. instructional approaches to increasing reading vocabulary. *Educational Perspectives*, 23, 16-21.
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 269-284). Erlbaum.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237-270.
<https://doi.org/10.3102/00028312024002237> 237-270
- 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>
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, P., & Chung, T. (2011). Teaching and testing vocabulary. In M. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 543-559). Blackwell Publishing.
<https://doi.org/10.1002/9781444315783.ch28>
- Nation, P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). Longman.
- Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle.

- Nesselhauf, N (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242.
<https://doi.org/10.1093/applin/24.2.223>
- Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008) Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, 45(3), 211-236. <https://doi.org/10.1080/01638530802073712>
- Nielson, M. B., Tengen, T., Idsoe, T., Matthiesen, S. B., Magerøy, N. (2015). Post-traumatic stress disorder as a consequence of bullying at work and at school. A literature review and meta-analysis. *Aggression and Violent Behavior*, 21, 17-24.
<https://doi.org/10.1016/j.avb.2015.01.001>
- Núñez, A. M. (2013). Advancing an intersectionality framework in higher education: Power and Latino postsecondary opportunity. In M. B. Paulsen (Ed.), *Handbook of theory and research in higher education* (Chapter 2). Springer.
- Núñez, A. M. (2014). Employing multilevel intersectionality in educational research: Latino identities, contexts, and college access. *Educational Researcher*, 43(2), 85-92.
<https://doi.org/10.3102/0013189X14522320>
- Nussle, J. (Nov 20, 2008). *Update of statistical area definitions and guidance on their uses* (OMB Bulletin No. 09-01). Office of Management and Budget.
<http://www.whitehouse.gov/sites/default/files/omb/assets/omb/bulletins/fy2009/09-01.pdf>
- Nuttall (1982). *Teaching reading skills in a foreign language*. Heinemann.
- Ogodo, J. (2022). Uncharted territories: COVID-19 and other 2020 events that changed lives forever. *The Journal of Multicultural Affairs*, 7(1). Article 4.

- Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing, 26*, 45-46. <https://doi.org/10.1007/s11145-012-9392-5>
- Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluence and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing, 19*, 218-233. <https://doi.org/10.1016/j.jslw.2010.10.003>
- Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report, 12*(2), 281-316.
- Onwuegbuzie, A. J., & Hitchcock, J. H. (2015). Advanced mixed analysis approaches. In S. Nagy e-Biber & R. B. Johnson (Eds.), *The Oxford handbook of multimethod and* <https://doi.org/10.1093/oxfordhb/9780199933624.013.19>
- Onwuegbuzie, A. J., & Mallett, M. H. (2011). Mixed research techniques in literacy research. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 301-330). The Guilford Press.
- Ortega, L. (2014). *Understanding second language acquisition*. Routledge.
- Osborne, J. W., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation, 8*(2), 1-5. Article 2. <https://doi.org/10.7275/r222-hv23>
- Osborne, J. W., & Waters, E. (2003). Multiple regression assumptions (ED470205). ERIC. <https://files.eric.ed.gov/fulltext/ED470205.pdf>
- Oshima, T. C., & Neel, J. H. (2003). *Mind your quants and quals in multiple regression*. Georgia Educational Research Association.

- Ozturk, Y., & Durmusoglu, Kose, G. (2016). Turkish and native English writers' use of lexical bundles. *Journal of Language and Linguistics Studies*, 12(1), 149-165.
- Paiz, J. M. (2019) Queering practice: LGBTQ+ diversity and inclusion in English language teaching. *Journal of Language, Identity & Education*, 18(4), 266-275
<https://doi.org/10.1080/15348458.2019.1629933>
- Palmer, J., Fam, D., Smith, T., & Kilham, S. (2014) Ethics in fieldwork: reflections on the unexpected. *Qualitative Report*, 19(28), 1-13.
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71. <https://doi.org/10.1016/j.jeap.2015.11.003>
- Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, 18(3), 391-417. <https://doi.org/10.1075/ijcl.18.3.06paq>
- Paribakht, T. S. (2005). The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language. *Language Learning*, 55(4), 701-748. <https://doi.org/10.1111/j.0023-8333.2005.00321.x>
- Paribakht, T. S., & Wesche, M. (2016). L1 influences in L2 lexical inferencing. In L. Yu & T. Odlin (Eds.), *New perspectives on transfer in second language learning* (pp. 76-103). Multilingual Matters.
- Paribakht, T. S., & Wesche, M. B. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 174-200). Cambridge University Press.
- Parry, K. (1997). Vocabulary and comprehension: Two portraits. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 55-68). Cambridge University Press.

- Peel, L. K. (2020). A beginner's guide to applied educational research using thematic analysis. *Practice Assessment, Research, and Evaluation*, 25. Article 2.
<https://doi.org/10.7275/ryr5-k983>
- Peet, M. E., Vela, M. B., Chin, M. (2020). Practical lessons for teaching about race and racism: Successfully leading free, frank, and fearless discussions. *Academic Medicine*, 95(12S), S139-S144. <https://doi.org/10.1097/ACM.00000000000003710>
- Peirce, C. S. (1933). *The collected papers of Charles Sanders Peirce* (C. Hartshorne & P. Weiss, Eds.). Harvard University Press.
- Peker, H. (2016). *Bullying victimization, feared second language self, and second language identity: Reconceptualizing second language motivational self-system* [Doctoral Dissertation, University of Central Florida]. Electronic Theses and Dissertations, 2004-2019. 5083.
- Peker, H. (2020). The effect of cyberbullying and traditional bullying on bilingual learners' national and oriented identities. *Bartın University Journal of Faculty of Education*, 9(1), 185-199.
- Pellicer-Sánchez, A. & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language*, 22(1), 31-55.
- Pereyra, N. (2015). Extensive reading in enhancing lexical chunks acquisition. *Reading Matrix: An International Online Journal*, 15(2), 218-234.
- Pichette, F., de Serres, L., & Lafontaine, M. (2012). Sentence reading and writing for second language vocabulary acquisition. *Applied Linguistics*, 33(1), 66-82.
<https://doi.org/10.1093/applin/amr037>

- Pilar, M. & Llach, A. (2007). Lexical errors as writing quality predictors. *Studia Linguistica*, 61(1), 1-19. <https://doi.org/10.1111/j.1467-9582.2007.00127>
- Pittaway, E., & Bartolomei, L. (2001). Refugees, race, and gender: The multiple discrimination against refugee women. *Refuge*, 9(6), 21-32. <https://doi.org/10.25071/1920-7336.21236>
- Pitts, M., White, H., Krashen, S. (1989). Acquiring the second language vocabulary through reading: A replication of the *Clockwork Orange* study using second language acquirers. *Reading in a Foreign Language*, 5(2), 271-275. <https://doi.org/10.125/67017>
- Plug, I., Stommel, W., Lucassen, P. L., olde Hartman, T. C., Dulmen, S. v., & Das, E. (2021). Do women and men use language differently in spoken face-to-face interaction? A scoping review. *Review of Communication Research*, 9, 43-79. <https://doi.org/10.12840/ISSN.2255-4165.026>
- Poland, B. D. (2000). Transcription quality. In J A. Holstein & J. F. Gubrium (Eds.), *Inside interviewing: New lenses, new concerns* (pp. 267-287). SAGE.
- Polat, N., & Mahalingappa, L.J. (2010). Gender differences in identity and acculturation patterns and L2 accent attainment. *Journal of Language, Identity, and Education*, 9, 17-35. <https://doi.org/10.1080/15348450903476832>
- Polio, C., & Williams, J. (2011). Teaching and testing writing. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 486-516). Wiley-Blackwell.
- Ponniah, J. (2011). Incidental acquisition of vocabulary by reading. *The Reading Matrix*, 11(2), 135-139.
- Poon-McBrayer, K. F. (2016). Complexities of shared ethnicity, immigrant education, and disabilities: Reconceptualizing multicultural special education. *International Journal of Special Education*, 31(3).

- Provis, C. (2009). On the definition of work. *Labour & Industry: A journal of the social and economic relations of work*, 20(2), 123-137.
<https://doi.org/10.1080/10301763.2009.10669396>
- Pourakbari, A. A. & Biria, R. (2015). Efficacy of task-induced involvement in incidental lexical development of Iranian senior EFL students. *English Language Teaching*, 8(5), 122-131.
<https://doi.org/10.5539/elt.v8n5p122>
- Pratt-Johnson, Y. (2015). Stressors experienced by immigrant and other non-native English-speaking students in the U.S. schools and their families. *Journal of Social Distress and Homelessness*, 24(3), 140-150.
- Purwanti Hery, I. S., Abas, I. H., Ilmu Ekonomi, S. T., & Al Sumait, A. (2020). The role of gender in vocabulary acquisition through nursery rhymes among young EFL learners in Indonesia. *Journal of Critical Reviews*, 7(15), 5959-5969.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282-307.
- Qian, D. D., & Lin, L. H. F. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 66-80). Routledge.
- Rashidi, N. & Khosravi, N. (2010) Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension of Iranian EFL learners. *Journal of Pan-Pacific Association of Applied Linguistics*, 14(1), 81-108.
- Ravitch, S. M., & Mittenfelner, N. C. (2021). *Qualitative research: Bridging the conceptual, theoretical, and methodological* (2nd ed.).
- Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

- Reichenbach, L., & Wesolkowska, M. (2008). All that is seen and unseen: The physical environment as informant. In M. Cefkin & M. Cotton (Eds.), *Proceedings of the Third Annual Ethnographic Practices in Industry (EPIC) Conference* (pp. 160-174).
- Restrepo Ramos, F. D. (2015). Incidental vocabulary learning in second language acquisition: A literature review. *PROFILE Issues in Teachers' Professional Development*, 17(1), 157-166. <https://doi.org/10.15446/profile.v17n1.43957>
- Rettew, D., & Pawlowski, S. (2022). Bullying: An update. *Child and Adolescent Psychiatric Clinics of North American*, 31(1), 1-9. <https://doi.org/10.1016/j.chc.2021.09.001>
- Reynolds, B. L. (2014). A mixed-methods approach to investigating first- and second-language incidental vocabulary acquisition through the reading of fiction. *Reading Research Quarterly*, 50(1), 111-127. <https://doi.org/10.1002/rrq.88>
- Reynolds, B. L. (2015). The effects of word form variation and frequency on second language incidental vocabulary acquisition through reading. *Applied Linguistics Review*, 6(4), 467-497. <https://doi.org/10.1515/applirev-2015-0021>
- Reynolds, B. L. (2016). The effects of target word properties on the incidental acquisition of vocabulary through reading. *TESL-EJ*, 20(3).
- Riazi, M. (2017). *Mixed methods research in language teaching and learning*. Equinox.
- Rigney, D. (2010). *The Matthew effect: How advantage begets further advantage*. Columbia University Press.
- Rintala, H., & Nokelainen, P. (2019). Vocational education and learners' experienced workplace curriculum. *Vocations and Learning*, 13, 113-130. <https://doi.org/10.1007/s12186-019-09229-w>

- Roche, T. & Harrington, M. (2013). Recognition vocabulary skill as a predictor of academic English performance and academic achievement in English. *Language Testing in Asia*, 3(12), 133-144. <https://doi.org/10.1186/2229-0443-3-12>
- Roessingh, H., Kouer, P., & Watt, D. (2005). Developing cognitive academic language proficiency: The journey. *TESL Canada Journal* 23(1), 1-27. <https://doi.org/10.18806/tesl.v23i1.75>
- Romer, D., & Jamieson, K. H. (2020). Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. *Social Science & Medicine*, 263, 1-8. <https://doi.org/10.1016/j.socscimed.2020.113356>
- Roohani, A., & Akbarpou, S. (2016). The effect of using songs on young Iranian EFL learners' vocabulary learning: A case of gender. *Journal of Modern Research in English Language Studies*, 3(3) 59-82.
- Ross, P. T., Hart-Johnson, T., Santen, S. A., & Bibler Zaidi, N. L. (2020). Considerations for using race and ethnicity as quantitative variables in medical education research. *Perspectives on Medical Education*, 9, 318-323. <https://doi.org/10.1007/s40037-020-00602-3>
- Roulston, K. (2013) *Reflective interviewing: A guide to theory & practice*. SAGE.
- Roxas, K., & Roy, L. (2012). "That's how we roll": A case study of a recently arrived refugee student in an urban high school. *Urban Review*, 44, 468-486. <https://doi.org/10.1007/s11256-012-0203-8>
- Roy, L. A. (2015). Borders and intersections of possibilities: Multilingual repertoires of refugee families in the Southwest U.S. *Multicultural Perspectives*, 17(2), 61-68. <https://doi.org/10.1080/15210960.2014.984718>

- Roy, L. A. (2018). *Teaching while white: Addressing the intersection of race and immigration in the classroom*. Rowman & Littlefield.
- Ruiz-Marin, H., & Bybee, R. W. (2022). The cognitive principles of learning underlying the 5E model of instruction. *International Journal of STEM Education*, 9, 1-9. Article 21.
<https://doi.org/10.1186/s40594-022-00337-z>
- Ryen, A. (2000). Cross-cultural interviewing. In J. A. Holstein & J. F. Gubrium (Eds.), *Inside interviewing: New lenses, new concern* (pp. 429-448). SAGE.
- Salam, U. (2021). Students' difficulties in learning vocabularies. *English Community Journal*, 5(1), 46-53. <https://doi.org/10.32502/ecj.v5i1.3327>
- Salazar, D. (2014). *Lexical bundles in the native and non-native scientific writing: Applying a corpus-based study to language teaching*. John Benjamins Publishing Company.
- Sapir, E. (1949). *Selected writings in language, culture and personality*. University of California Press.
- Saragi, T., Nation, P., & Meister, G. (1978). Vocabulary learning and reading. *System*, 6, 70-78
- Sarbazi, M., Khany, R., Shoja, L. (2021). The predictive power of vocabulary, syntax, and metacognitive strategies for L2 reading comprehension. *Southern African Linguistics and Applied Language Studies*, 39(3), 244-258.
<https://doi.org/10.2989/16073614.2021.1939076>
- Schmid H. J. (Ed.). (2017). *Entrenchment of the psychology of language learning*. American Psychological Association; De Gruyter Mouton. <https://doi.org/10.1037/15969-000>
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave.

- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951. <https://doi.org/10.1111/lang.12077/full>
- Schmitt, N., Dörnyei, Z., Adolphs, S., & Durow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 55-86). John Benjamin Publishing.
- Schoonen, R., van Gelderen, A., Stoel, R., & Hulstijn, J. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning*, 16(1), 31-79. <https://doi.org/10.1111/j.1467-9922.2010.00590.x>
- Schoonenboom, J., & Johnson, B. (2017). How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2), 107-131. <https://doi.org/10.1007/s11577-017-0454-1>
- Schuele, C. M., & Justice, L. M. (2006). The importance of effect sizes in the interpretation of research. *The ASHA Leader*, 11(10), 14-27. <https://doi.org/10.1044/leader.FTR4.11102006.14>
- Schumann, K., & Ross, M. (2010). Why women apologize more than men: Gender differences in thresholds for perceiving offensive behavior. *Psychological Science*, 21(11), 1649-1655. <https://doi.org/10.1177/0956797610384150>
- Schutz, P. A., Chambless, C. B., & DeCuir, J. T. (2004). Multimethods research. In K. deMarras & S. D. Lapan (Eds.), *Foundations for research: Methods of inquiry in education and the social sciences* (pp. 267-282). Lawrence Erlbaum Associates.
- Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 189-213). SAGE.

- Seirk, J. (2019). White girls and the “other” in the new Latino diaspora: The conflation of deviance and immigration and the role of language. *Theory in Action, 12*(2), 46-70.
<https://doi.org/10.3798/tia.1937-0237.1911>
- Severino, C., & Deifell, E. (2011). Empowering L2 tutoring: A case study of a second language writer’s vocabulary learning. *The Writing Center Journal, 31*(1), 25-54.
- Shadikah, A. A., Fauziati, E., Supriyadi, S. (2017). The effect of vocabulary learning strategies on vocabulary mastery based on gender differences. In U. Utomo et al. (Eds.), *Proceeding of the 2nd International Conference on Art, Language, and Culture* (pp. 495-505). ASSEHR.
- Sharkey, J., & Layzer, C., (2000). Whose definition of success? Identifying factors that affect bilingual learners' access to academic success and resources. *TESOL Quarterly, 34*(2), 352-368.
- Sharma, A., Borah, S. B. (2020). Covid-19 and domestic violence: An indirect path to social and economic crisis. *Journal of Family Violence, 2020*. <https://doi.org/10.1007/s10896-020-00188-8>
- Sheth, M. J. (2018). Grappling with racism as foundational practice of science teaching. *Science Education, 103*(1), 37-60.
- Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and variation. *Journal of English for Academic Purposes, 49*, 1-13. <https://doi.org/10.1016/j.jeap.2020.100946>
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics 31*(4), 487-512.
<https://doi.org/10.1093/applin/amp058>

- Sinclair, J. (2008). The phrase, the whole phrase, and nothing but the phrase. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 407-410). John Benjamins Publications.
- Singer, D., Avedon, J., Hering, R., McCann, A., & Sacks, C. (1977). Sex differences in the vocabulary of college students. *The Journal of Sex Research, 13*(4), 267-273.
<https://doi.org/10.1080/00224497709550984>
- Singleton, D. (2006). Lexical transfer: Interlexical or intralexical? In J. Arabski (Ed.), *Cross linguistic influences in the second language lexicon* (pp. 13-143). Multilingual Matters.
<https://doi.org/10.21832/9781853598579-014>
- Smith, B., & Sparkes, A. C. (2020). Qualitative research. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (4th ed., pp. 999-1019). John Wiley & Sons.
<https://doi.org/10.1002/9781119568124.ch49>
- Smolen, P., Zhang, Y., & Byrne, J. H. (2016). The right time to learn: Mechanisms and optimization of spaced learning. *National Review of Neuroscience, 17*(2), 77-88.
<https://10.1038/nrn.2015.18>
- Snoder, P. (2017). Improving English learners' productive collocation knowledge: The effects of involvement load, spacing, and intentionality. *TESL Canada Journal/Revue TESL du Canada, 34*(3), 140-164. <https://doi.org/10.18806/tesl.v34i3.1277>
- Sobel, H. S., Wiseheart, M., & Kapler, I. (2010). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology 25*(5), 763-767.
<https://doi.org/10.1002/acp.1747>

- Solati-Dehkordi, Seyed, A., & Salehi, H. (2016). Impact of explicit vocabulary instruction on writing achievement of upper-intermediate EFL learners. *International Education Studies*, 9(4), 141-154. <https://doi.org/10.5539/ies.v9n4p141>
- Soleimani, H., & Rahmanian, M. (2015). Visiting involvement load hypothesis and vocabulary acquisition in similar task types. *Theory and Practice in Language Studies*, 5(9), 1883-1889. <https://doi.org/10.17507/tpls.0509.16>
- Spagnola, M. & Fiese, B. H. (2007). Family routines and rituals: A context for development in the lives of young children. *Infants & Young Children*, 20(4), 284–299.
- Staff, J., Yetter, A. M., Cundiff, K., Ramirez, N., Vuolo, M., & Mortimer, J. T. (2020). Is adolescent employment still a risk factor for high school dropout? *Journal of Research on Adolescence*, 30(2), 406-422.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56(1), 72-110.
<https://doi.org/10.3102/00346543056001072>
- Stake, R. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443-466). SAGE.
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes* 12, 214-225. <https://doi.org/10.1016/j.jeap.2013.05.002>
- Stefanowitsch, A., & Flach, S. (2017). The corpus-based perspective on entrenchment. In H. J. Schmid (Ed.), *Language and the human lifespan series. Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 101–127). De Gruyter Mouton. <https://doi.org/10.1037/15969-006>

- Stevens, J. P. (2007). *Intermediate statistics: A modern approach* (3rd ed.). Routledge.
- Stewart, J., Batty, A. O., & Bovee, N. (2012). Comparing multidimensional and continuum models of vocabulary acquisition: An empirical examination of the vocabulary knowledge scale. *TESOL Quarterly*, *46*(4), 695-721. <https://doi.org/10.1002/tesq.35>
- Stockdale, M.S., Bell, M.P., Crosby, F. & Berdahl, J. (2020), From me too to what now: advancing scholarship on sex harassment issue 1: a persistent problem. *Equality, Diversity and Inclusion*, *39*(1), 1-4. <https://doi.org/10.1108/EDI-02-2020-296>
- Sugarman, J., & Geary, C. (2018). *English learners Georgia: Demographics, outcomes, and state accountability policies*. Migration Policy Institute.
https://www.migrationpolicy.org/sites/default/files/publications/EL-factsheet2018-Georgia_Final.pdf
- Sullivan, G. B. (2015). Wittgenstein's later philosophy and "pictures" of mixed methods research in psychology: A critical investigation of theories and accounts of methodological plurality. *Theory & Psychology*, *25*(4), 473-493.
<https://doi.org/10.1177/0959354315586092>
- Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. <https://doi.org/10.4300/JGME-D-12-00156.1>
- Suter, W. N. (2012). *Introduction to educational research: A critical thinking approach* (2nd ed.). SAGE.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition*. Newbury House.

- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Swales, J. M. (2004). Foreword. In D. Belcher & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connections*. University of Michigan.
- Tabrizi, A. R., & Abbasi, S. (2016). The effect of oral summary of short stories on Iranian intermediate EFL learners' vocabulary learning: With a focus on gender. *International Journal of English Linguistics*, 6(6), 129-137. <https://doi.org/10.5539/ijel.v6n6p129>
- Tan, D. N. L., & Mante-Estaci, M. J. (2021). Reader-text connection: Reporting the engagement of high school students with culturally relevant texts. *TEFLIN*, 32(2), 342-361
<https://doi.org/10.15639/teflinjournal.v32i2/342-361>
- Tan, D., Pandian, A., & Jaganathan, P. (2016). Lexical testing and the reliability of the modified vocabulary knowledge scale. *Advances in Language and Literary Studies*, 7(5), 91-96.
<https://doi.org/10.7575/aiac.all.v.7n.5p.91>
- Tashakkori, A., & Creswell, J. W. (2007). The new era of mixed methods. *Journal of Mixed Methods Research*, 1(1), 3-7. <https://doi.org/10.1177/2345678906293042>
- Tashakkori, A., & Teddlie, C. (2008). Quality of inferences in mixed methods research: Calling for an integrative framework. In M. M. Bergman (Ed.), *Advances in mixed methods research* (pp. 11-21). SAGE. <https://doi.org/10.4135/9780857024329.d10>
- Teddlie, C. & Tashakkori, A. (2012). Common “core” characteristics of mixed methods research: A review of critical issues and call for greater convergence. *American Behavioral Scientist* 56(6), 744-788. <https://doi.org/10.1177/0002764211433795>

- Teddlie, C., & Johnson, R. B. (2009a). Methodological thought before the 20th century. In C. Teddlie & A. Tashakkori (Eds.), *Foundations of mixed methods: Integrating qualitative and quantitative approaches in the social and behavioral sciences* (pp. 40-61). SAGE.
- Teddlie, C., & Johnson, R. B. (2009b). Methodological thought since the 20th century. In C. Teddlie & A. Tashakkori (Eds.), *Foundations of mixed methods: Integrating qualitative and quantitative approaches in the social and behavioral sciences* (pp. 62-82). SAGE.
- Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, 1(1), 77-77. <https://doi.org/10.1177/2345678906292430>
- Tekman, E. A. F., & Daloğlu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals*, 39(2), 220-243. <https://doi.org/10.1111/j.1944-9720.2006.tb02263.x>
- Teng, F. (2016). The effects of word exposure frequency on incidental vocabulary of the depth of vocabulary knowledge. *GEMA Online Journal of Language Studies*, 16(3), 53-70. <https://doi.org/10.17576/gema-2016-1603-4>
- The Five Graces Group—Beckner, C., Blythe, R., Bybee, J., Christiansen, M., Croft, W., Ellis, N., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system. *Language Learning*, 59(s1), 1-23. <https://doi.org/10.1111/j.1467-9922.2009.00533.x>
- Thorndike, E. (1921). *The teachers word book*. Teachers College Press.
- Tinker Sachs, G. (2001). Transforming extensive reading lessons. *New Horizons in Education*, 43, 78-90.
- Torgerson, C. J., & Torgerson, D. (2001). The need for randomised controlled trials in educational research. *British Journal of Educational Studies* 49(3), 316-328.

- Trainor, A., Murray, A., & Kim, H. (2016). English learners with disabilities in high school: Population characteristics, transition programs, and postschool outcomes. *Remedial & Special Education, 37*(3), 146-158. <https://doi.org/10.1177/0741932515626797>
- Tremblay, D. (1966). Laws of learning general and specialized vocabularies: Summary. *Proceedings of the 74th Annual Convention of American Psychological Association* (pp. 229-230). <https://doi.org/10.1037/e469452008-115>
- Türker, E. (2019). Idiom acquisition by second language learners: The influence of cross-linguistic similarity and context. *The Language Learning Journal, 47*(2), 133-144. <https://10.1080/09571736.2016.1221441>
- Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 153-172). <https://doi.org/10.1075/LLLT.9.09UND>
- United States Census Bureau (2012). *Patterns of metropolitan and micropolitan population change: 2000 to 2010* (Report No. C2010SR-01). <https://www.census.gov/library/publications/2012/dec/c2010sr-01.html>
- U.S. Department of Education (2017). Profiles of English Learners (ELs). Office of English Language Acquisition, National Clearinghouse for English Language Acquisition. http://www.ncela.us/files/fast_facts/05-19-2017/ProfilesOfELs_FastFacts.pdf.
- U.S. News and World Reports (2023). Best high schools. <https://www.usnews.com>
- Uprichard, E., & Dawney, L. (2019). Data diffraction: Challenging data integration in mixed methods research. *Journal of Mixed Methods Research, 13*(1), 19-32. <https://doi.org/10.1177/1558689816674650>

- Urdaniz, R. P., & Skoufaki, S. (2019). Spanish L1 EFL learners' recognition knowledge of English academic vocabulary: The role of cognateness, word frequency and length. *Applied Linguistics Review*. Advance online publication. <http://doi.org/10.1515/applirev-2018-0109>
- Van Benthuysen, R. (2010). Japanese ESL learners' acquisition of English loanword cognates: Difficulties due to semantic, phonological, and orthographic factors. *Journal of Bunkyo Gakuin University, 10*, 255-263.
- van den Bergh, H., Rijlaarsdam, G., & van Steendam, E. (2016). Writing process theory: A functional dynamic approach. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (p. 57–71). The Guilford Press.
- Van den Broek, G., Takashima, A., Wiklund-Hörnqvist, L., Wirebring, L. K., Segers, E., Verhoeven, L., & Nyberg, L. (2016). Neurocognitive mechanisms of the “testing effect”: A review. *Trends in Neuroscience and Education, 5*, 52-66.
<https://doi.org/10.1016/j.tine.2016.05.001>
- Van Manen, M. (2016). *Researching the lived experience: Human science for an action sensitive pedagogy*. Routledge.
- VanPatten, V. & Williams, J. (Eds.), (2014). *Theories in second language acquisition: An introduction* (2nd ed., pp.75-93). Routledge.
- Varandi, S. B., & Faezi, F. (2013). The role of dimensions of lexical repertoire in reading comprehension and incidental vocabulary acquisition. *Theory and Practice in Language Studies, 3*(11), 2078-2085. <https://doi.org/10.4304/tpls.3.11.2078-2085>

- Varrick, D. (2016). Enhancing English academic vocabulary acquisition and retention in intensive English programs with the involvement load hypothesis. *Theory and Practice in Language Studies*, 6(12), 2237-2244. <https://doi.org/10.17507/tpls.0612.01>
- Vela, V. (2015). Using glosses for incidental vocabulary acquisition. *Procedia - Social and Behavioral Sciences*, 199, 305-310. <https://doi.org/10.1016/j.sbspro.2015.07.551>
- Vidal, K. (2012). Maintenance of lower-frequency vocabulary: A forgotten issue? *IUP Journal of English Studies*, 7(1), 54-63.
- Vlach, H. A., & Sandhofer, C. M. (2012). Distributed learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development*, 83(4), 1137-1144. <https://doi.org/10.1111/j.1467-8624.2012.01781.x>
- Vogt, W. P. (2011). *Dictionary of statistics and methodology: A non-technical guide for the social sciences* (4nd ed.). Sage Publications.
- Waitoller, F. R., & Fozlescki, E. B. (2013). Working in boundary practices: Identity development and learning in partnerships for inclusive education. *Teaching and Teacher Education*, 31, 35-45. <https://doi.org/10.1016/j.taa.2012.11.006>
- Walker, A., Shafer, J., & Iiams, M. (2004). "Not in my classroom": Teacher attitudes towards bilingual learners in the mainstream classroom. *National Association for Bilingual Education Journal of Research and Practice*, 2, 130-160.
- Walters, J., & Bozkurt, N. (2009). The effect of keeping vocabulary notebooks on vocabulary acquisition. *Language Teaching Research*, 13(4), 403-423.
<https://doi.org/10.1177/1362168809341509>

- Wang, B. (2016). Gender difference in effect of positive emotion on consolidation of memory for definitions of English vocabulary. *International Journal of Bilingualism*, 21(3), 306-316. <https://doi.org/10.1177/1367006915625456>
- 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, 130-163.
- Weaver, K. (2018). Pragmatic Paradigm. In B. B. Frey (Ed.), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation* (pp. 1286-1288). SAGE Reference. <https://doi.org/10.4135/9781506326139.n534>
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65. <https://doi.org/10.1093/applin/aml048>
- Webb, S. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *The Modern Language Journal*, 104(4), 715-738. <https://doi.org/10.1111/modl.12671>
- Webb, S. (1997). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27, 33-52. <https://doi.org/10.1017/S0272263105050023>
- Weis, M., Heikamp, T., & Trommsdorff, G. (2013). Gender differences in school achievement: The role of self-regulation. *Frontiers in Psychology*, 4, Article 442. <https://doi.org/10.3389/fpsyg.2013.00442>
- Wesche, M. B. & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: depth versus breadth. *Canadian Modern Language Review*, 53(1), 13-40. <https://doi.org/10.3138/cmlr.53.1.13>
- Wesche, M., & Paribakht, T. S. (2010). *Lexical inferencing in a first or second language: Cross-linguistic dimensions*. Multilingual Matters.

- White, R. (2015). The expanding literate lexicon in L1 secondary school writing: A mixed methods inquiry. *New Zealand Studies in Applied Linguistics*, 12, 5-19.
- WIDA (2020). *ACCESS for Els interpretive guide for score reports, grades K-12: Understanding student scores*.
<https://wida.wisc.edu/sites/default/files/resource/Interpretive-Guide.pdf>
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321-331. <https://doi.org/10.1016/j.2012.09.007>
- Wills, J. S. (2019). "Daniel was racist": Individualizing racism when teaching about the Civil Rights Movement. *Theory & Research in Social Education*, 47(3), 396-425.
<https://doi.org/10.1080/00933104.2019.1583620>
- Won, Y. (2019). The prediction of writing scores using vocabulary features in ESL university students' essays. *Modern English Education*, 11, 31-40.
<https://doi.org/10.18095/meeso.2019.20.4.31>
- Wood, D. (2020). Classifying and identifying formulaic language. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 30-45). Routledge.
- Wray, A. (2000). Formulaic sentences in second language teaching: Principle and practice. *Applied Linguistics*, 21(4), 463-489. <https://doi.org/10.1093/applin/21.4.463>
- Wray, A. (2002). *Formulaic sequences and the lexicon*. Cambridge University Press.
- Wright, T. S., & Cervetii, G. N. (2016). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, 52(2), 203-226. <https://doi.org/10.1002/rrq.163>

- Wright, T. S., & Cervetti, G. N. (2016). A systematic review of the research on vocabulary that impacts text comprehension. *Reading Research Quarterly, 52*(2), 203-226.
<https://doi.org/10.1002/rrq.163>
- Wu, H. (2018). The effects of field independent/field dependent cognitive styles on incidental vocabulary acquisition under reading task. *Theory and Practice in Language Studies, 8*(7), 813-822. <https://doi.org/10.17507/tpls.0807.12>
- Wu, S., Dixon, L. Q., Sun, H., & Zhang, P. (2021). Breadth or depth: The role of vocabulary in Chinese English language beginning writers' development. *International Journal of Bilingual Education and Bilingualism, 24*(9), 1356-1372.
<https://doi.org/10.1080/13670050.2019.1572066>
- Xu, W., & Zammit, K. (2020). Applying thematic analysis to education: A hybrid approach to interpreting data in practitioner research. *International Journal of Qualitative Methods, 19*, 1-9. <https://doi.org/10.1177/1609406920918810>
- Xu, X. (2010). The effects of glosses on incidental vocabulary acquisition in reading. *Journal of Language Teaching and Research, 1*(2), 117-120. <https://doi.org/10.4304/jltr.1.2.117-120>
- Yang, Y., Shintani, N., Li, Shaofeng, & Zhang, Y. (2017). The effectiveness of post-reading word-focused activities and their associations with working memory. *System, 70*, 38-49.
<https://doi.org/10.1016/j.system.2017.09.012>
- Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta analysis. *Journal of Applied Developmental Psychology, 37*, 36-51.
<https://doi.org/10.1016/j.appdev.2014.11.005>

- Yeatman, J. D., McCloy, D. R., Caffarra, S., Clarke, M. D., Ender, S., Gijibels, L., Joo, S. J., Kubota, E. C., Kuhl, P. K., Larson, E., O'Brien, G., Peterson, E. R., Takada, M. E., & Taulu, S. (2022). Reading instruction causes changes in category-selective visual cortex. *bioRxiv* [preprint] <https://doi.org/10.1101/2022.02.02.477919>
- Yoshii, M. (2013). Effects of gloss types on vocabulary learning through reading: Comparison of single translation and multiple-choice gloss types. *CALICO Journal*, 203-229. <https://doi.org/10.1558/cj.v30i0.203-229>
- Yu, B. B. (2016). Bilingualism as conceptualized and bilingualism as lived: A critical examination of the monolingual socialization of a child with autism in a bilingual family. *Journal of Autism & Developmental Disorders*, 46(2), 424-435. <https://doi.org/10.1007/s10803-015-2625-0>
- Yusuf, M. A., & Sim, T. S., & A., S. (2014). Students' proficiency and textual computer gloss use in facilitating vocabulary knowledge. *English Language Teaching*, 7(11), 99-107. <https://doi.org/10.5539/elt.v7n11p99>
- Zaff, J. F., Margolius, M., Varga, S. M., Lynch, A. D., Tang, C., Donlan, A. E. (2021). English learners and high school graduation: A pattern-centered approach to understanding within-group variations. *Journal of Education for Students Placed at Risk*, 26(1), 1-19. <https://doi.org/10.1080/10824669.2020.1780597>
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review*, 57(4), 541-72. <https://doi.org/10.3138/cmlr.57.4.541>

- Zarei, A. A., & Hasani, S. (2011). The effects of glossing conventions on L2 vocabulary recognition and production. *Journal of Teaching Language Skills*, 30(2) 209-233.
<https://doi.org/10.22099/jtls.2012.390>
- Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research*, 23(2), 123-153.
<https://doi.org/10.1177/0267658307076543>
- Zhai, W. (2020). Literature review: Theoretical and empirical research on age effects in SLA: Focusing on critical period hypothesis. In H. Yan, (Ed.), *2nd International Conference on Educational Reform, Management Science and Sociology* (pp. 82-86). Clausius Scientific Press. <https://doi.org/10.23977/ERMSS2020017>
- Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research*, 26(4), 696-725. <https://doi.org/10.1177/1362168820913998>
- Zhao, Y. G., Biales, C., & Olszewski, A. (2016). Exploring learner factors in second language (L2) incidental vocabulary acquisition through reading. *Reading in a Foreign Language*, 28(2) 224-245.

APPENDICES

Appendix A: Literature Review Descriptive Tables

Vocabulary Acquisition Study Methodologies, Designs, and Methods

Study	Methodology	Design	Treatment Duration	Delayed Testing
AbuSeileek (2011)	Quant	True Exp	9 months	none
Ahmed et al. (2014)	Quant	Quasi Exp	4 years	N/A
Alavi & Keyvanshekouh (2012)	Quant	True Exp	3 months	none
Alavinia & Rahimi (2019)	Quant	True Exp	1 day	1 month
Alharbi (2016)	Quant	True Exp	4 months	none
Alharbi (2018)	Mixed	Quasi Exp	1 day	none
Alsaif & Masrai (2019)	Mixed	Quasi Exp	2 months	none
Amirian & Behshad (2016)	Quant	Quasi Exp	1 day	2 months
Arnaud & Savignon (1997)	Quant	Quasi Exp	5 years +	N/A
Azari et al. (2012)	Quant	True Exp	2 months	none
Bai (2016)	Quant	Quasi Exp	3 weeks	IV
Bao (2015)	Quant	True Exp	1 day	none
Cao (2013)	Quant	Quasi Exp	1 day	1 week
Çekiç & Bakla (2019)	Quant	Quasi Exp	2 months	none
Cheng & Good (2009)	Quant	True Exp	1 day	2 weeks
Choi (2017)	Quant	Quasi Exp	1 day	none
Cunningham & Stanovich (1997)	Quant	Quasi Exp	3 years	N/A
Dabaghi Varnosfadrani & Rafiee (2012)	Quant	True Exp	1 day	none
Daskalovska (2014)	Quant	Quasi Exp	2 days	none
Dobbs & Kearns (2016)	Quant	Quasi Exp	1 day	1 week
Duan (2018)	Quant	Quasi Exp	1 day	1 week
Duff et al. (2015)	Quant	Quasi Exp	6 years	N/A
Eckerth & Tavakoli (2012)	Quant	True Exp	3 months	3 weeks
Ehsanzadeh (2012)	Quant	Quasi Exp	2 weeks	1 month
El-Dakhs et al. (2017)	Quant	Quasi Exp	1 month	none
Far (2016)	Quant	Quasi Exp	1 day	none
Fathi & Sarkhosh (2019)	Quant	Quasi Exp	1 day	3 weeks
Frishkoff et al. (2008)	Quant	True Exp	3 weeks	1 week
Ghorbani & Rahmandoost (2012)	Quant	Quasi Exp	2 months	none
Gohar et al. (2018)	Quant	Quasi Exp	1 day	1 week
Hazrat (2015)	Quant	Quasi Exp	1 day	3 weeks
Heidari-Shahreza & Tavakoli (2016)	Quant	Quasi Exp	1 day	2 weeks
Horst et al. (1998)	Quant	Quasi Exp	3 months	N/A
Huang (2015)	Quant	Quasi Exp	N/A	N/A
Hulstijn & Laufer (2001)	Quant	Quasi Exp	2 days	2 weeks

Study	Methodology	Design	Treatment Duration	Delayed Testing
Joe (2010)	Mixed	Quasi Exp	3 months	N/A
Jung (2016)	Quant	Quasi Exp	3 weeks	1 week
Kelley et al. (2010)	Quant	Quasi	3 months	none
Khan (2019)	Qual	Case	3 months	N/A
Khatib & Faruji (2012)	Quant	True Exp	1 month	none
Khoshsima & Eskandari (2017)	Quant	Quasi Exp	1 day	none
Kim (2011)	Quant	Quasi Exp	2 weeks	2 weeks
Kweon & Kim (2008)	Quant	True	1 month	1 month
Larson (2014)	Mixed	Quasi Exp	4 months	none
Laufer & Rozovski-Roitblat (2011)	Mixed	Quasi Exp	4 months	none
Lee (2003)	Qual	Case	1 day	1 month
Lee & Muncie (2006)	Quant	Quasi Exp	1 day	2 weeks
Lee et al. (2016)	Quant	True Exp	1 day	2 weeks
Lessard-Clouston (2012)	Quant	Quasi Exp	N/A	N/A
Levitvsky & Laufer (2013)	Quant	Quasi Exp	8 years	N/A
Lin & Kawai (2016)	Mixed	Case	NR	N/A
Liu (2017)	Quant	Quasi Exp	1 day	1 week
Ma (2013)	Qual	Case	3 months	N/A
Mancilla-Martinez & Lesaux (2010)	Quant	Quasi Exp	5 months	N/A
Min (2008)	Quant	Quasi Exp	1 month	3 months
Muncie (2002)	Quant	Quasi Exp	1 month	N/A
Nagy et al. (1987)	Quant	True Exp	1 week	N/A
Parry (1997)	Qual	Case	2 months	N/A
Pellicer-Sánchez & Schmitt (2010)	Quant	Quasi Exp	2 weeks	none
Pereyra (2015)	Quant	Quasi Exp	4 months	none
Pichette et al. (2012)	Quant	Quasi Exp	1 day	1 week
Ponniah (2011)	Quant	Quasi Exp	1 day	none
Pourakbari & Biria (2015)	Quant	True Exp	1 day	2 weeks
Reynolds (2014/2015/2016)	Mixed	Quasi Exp	2 weeks	none
Roessingh et al. (2005)	Quant	Quasi Exp	3 years	N/A
Samian et al. (2016)	Quant	True Exp	1 day	1 month
Schmitt et al. (2004)	Quant	Quasi Exp	2, 3 months	none
Schoonen et al. (2011)	Quant	Quasi Exp	2 years	N/A
Severino & Deifell (2011)	Qual	Case	3 months	N/A
Snoder (2017)	Quant	Quasi Exp	IV	3 weeks
Sobel et al. (2010)	Quant	Quasi Exp	IV	3 weeks
Solati-Dehkordi & Salehi (2016)	Quant	Descriptive	1 day	1 month
Soleimani & Rahmanian (2015)	Quant	Quasi Exp	2 months	2 weeks
Sun (2014)	Mixed	Quasi Exp	1 day	none
Tekman & Daloğlu (2006)	Quant	Quasi Exp	1 week	3 months

Study	Methodology	Design	Treatment Duration	Delayed Testing
Teng (2016)	Quant	Quasi Exp	3 weeks	3 months
Teng (2017)	Quant	Quasi Exp	3 weeks	1 ½ month
Touti & Maleki (2016)	Quant	Quasi Exp	2 months	none
Varandi & Faezi (2013)	Quant	Quasi Exp	1 day	none
Varol & Erçetin (2016)	Quant	True Exp	1 day	3 weeks
Varrick (2016)	Quant	Quasi Exp	1 day	2 weeks
Vela (2015)	Quant	True Exp	1 day	2 weeks
Walters & Bozkurt (2009)	Mixed	True Exp	1 month	none
Waring & Takaki (2003)	Quant	Quasi Exp	1 day	3 months
Webb (2007)	Quant	True Exp	1 day	none
White (2015)	Mixed	Quasi Exp	3 years	N/A
Wolsey (2010)	Qual	Case	6 months	N/A
Wu (2018)	Quant	Quasi-Exp	1 day	none
Xu (2010)	Quant	True Exp	1 day	1 week
Yang et al. (2017)	Quant	Quasi-Exp	1 day	2 weeks
Yoshii (2013)	Quant	True Exp	1 day	1 week
Yoshii (2014)	Quant	Quasi-Exp	1 day	2 weeks
Yusuf et al. (2014)	Quant	Quasi-Exp	1 day	3 weeks
Zahar et al. (2001)	Quant	Quasi-Exp	2 days	none
Zarei & Hasani (2011)	Mixed	Quasi-Exp	2 months	2 weeks
Zhao et al. (2016)	Quant	Quasi-Exp	1 day	none

Note: *Quant* refers to quantitative methodology, whereas *Qualitative* refers to qualitative methodologies. *Mixed* refers to mixed methods methodology self-described by the study's researcher(s). The design column refers to the specific study design. *True Exp* refers to true experimental design. *Quasi-Exp* refers to quasi-experimental design. *Case* refers to case study. The time under treatment duration refers to the time of the actual treatment and does not include posttesting. The delayed testing column notes whether the researcher(s) conducted a self-described delayed testing. The time period entry is the time that elapsed between the immediate posttest and the last delayed posttest.

Targeted Items and Text Information

Study	Target Items	Text Genre or Origin	Word Count
AbuSeileek (2011)	25 Words	Textbook	400
Ahmed et al. (2014)	- Words	N/A	N/A
Alavi & Keyvanshekouh (2012)	- Words	Graded Reader	NR
Alavinia & Rahimi (2019)	45 Words	N/A	N/A
Alharbi (2016)	25 Words	Textbook	695
Alharbi (2018)	25 Words	Graded Reader	505
Alsaif & Masrai (2019)	- Words	Novels	Varied
Amirian & Behshad (2016)	- Both	NES Essays	NR
Arnaud & Savignon (1997)	60 Both	N/A	N/A
Azari et al. (2012)	30 Words	Wikipedia Article	NR
Bai (2016)	20 Words	Article	1208
Bao (2015)	18 Words	N/A	N/A
Cao (2013)	10 Phrases	N/A	N/A
Çekiç & Bakla (2019)	20 Words	Expository	3,840
Cheng & Good (2009)	16 Words	Textbook	< 300
Choi (2017)	- Words	N/A	N/A
Cunningham & Stanovich (1997)	- Words	N/A	N/A
Dabaghi Varnosfadrani & Rafiee (2012)	20 Words	Multiple	NR
Daskalovska (2014)	51 Words	Novel	11,672
Dobbs & Kearns (2016)	25 Words	N/A	N/A
Duan (2018)	11 Words	Article	570
Duff et al. (2015)	- Words	N/A	N/A
Eckerth & Tavakoli (2012)	30 Words	Article	-
Ehsanzadeh (2012)	13 Words-nonce	Graded Reader	NR
El-Dakhs et al. (2017)	4 Both	N/A	N/A
Far (2016)	108 Words	Exp, Narrative	IV
Fathi & Sarkhosh (2019)	- Words	Textbook	NR
Frishkoff et al. (2008)	60 Words-nonce	N/A	N/A
Ghorbani & Rahmandoost (2012)	40 Words	Textbook	NR
Gohar et al. (2018)	10 Words	N/A	N/A
Hazrat (2015)	10 Words	Expository	N/R
Heidari-Shahreza & Tavakoli (2016)	10 Words	Homogeneous	1400
Horst et al. (1998)	45 Words	Novel	21,232
Huang (2015)	- Phrases	N/A	N/A
Hulstijn & Laufer (2001)	10 Words	Persuasive	621
Joe (2010)	74 Words	N/A	N/A
Jung (2016)	14 Words-nonce	Articles	685
Kelley et al. (2010)	10 Words	NR	NR
Khan (2019)	- Phrases	N/A	N/A
Khatib & Faruji (2012)	38 Words	Narrative	NR
Khoshsima & Eskandari (2017)	20 Words	Expository	551

Study	Target Items	Text Genre or Origin	Word Count
Kim (2011)	10 Words	Expository	NR
Kweon & Kim (2008)	367 Words	Novel	134,013
Larson (2014)	- Words	N/A	N/A
Laufer & Rozovski-Roitblat (2011)	60 Words	Textbook	5,200
Lee (2003)	36 Both	Expository	893
Lee & Muncie (2006)	42 Both	Expository	623
Lee et al. (2016)	30 Words	N/A	N/A
Lessard-Clouston (2012)	7 Words	N/A	N/A
Levitzsky & Laufer (2013)	140 Words	N/A	N/A
Lin & Kawai (2016)	- Words	NR	NR
Liu (2017)	18 Words	NR	NR
Ma (2013)	- Words	N/A	N/A
Mancilla-Martinez & Lesaux (2010)	100 Words	NR	NR
Min (2008)	50 Words	Articles	NR
Muncie (2002)	- Words	N/A	N/A
Nagy et al. (1987)	50 Words	Expo & Narrative	N/A
Parry (1997)	- Words	Expository	NR
Pellicer-Sánchez & Schmitt (2010)	34 Words-Ibo	Novel	67,000
Pereyra (2015)	- Phrases	Varied	N/A
Pichette et al. (2012)	16 Words	N/A	N/A
Ponniah (2011)	20 Words	Narrative	N/A
Pourakbari & Biria (2015)	10 Words	N/A	N/A
Reynolds (2014/2015/2016)	49 Words-nonce	Novel	37,568
Roessingh et al. (2005)	- Words	N/A	N/A
Samian et al. (2016)	41 Words	Narratives	NR
Schmitt et al. (2004)	20 Phrases	NR	NR
Schoonen et al. (2011)	- Words	N/A	N/A
Severino & Deifell (2011)	- Both	N/A	N/A
Snoder (2017)	28 Phrases	Nonfiction	NR
Sobel et al. (2010)	8 Words	N/A	N/A
Solati-Dekhordi & Salehi (2016)	96 Both	Expository	378
Soleimani & Rahmanian (2015)	10 Words	Expository	326
Sun (2014)	10 Words-nonce	Narrative	1,674
Tekman & Daloglu (2006)	36 Words	Narrative	2,400
Teng (2016)	36 Words-nonce	Graded Reader	300
Teng (2017)	24 Words	Textbook	800
Touti & Maleki (2016)	28 Words	Expository	N/A
Varandi & Faezi (2013)	50 Words-German	Not Reported	N/A
Varol & Erçetin (2016)	28 Words	Expository	2,622
Varrick (2016)	36 Words	Textbook	NR
Vela (2015)	10 Words	Narratives	NR
Walters & Bozkurt (2009)	72 Words	N/A	N/A

Study	Target Items	Text Genre or Origin	Word Count
Waring & Takaki (2003)	25 Words	Graded Reader	400
Webb (2007)	10 Words-nonce	Graded Reader	NR
White (2015)	- Words	N/A	N/A
Wolsey (2010)	N/A Words	N/A	N/A
Wu (2018)	10 Words	Not Reported	NR
Xu (2010)	18 Words	Not Reported	774
Yang et al. (2017)	8 Words	News/Article	827
Yoshii (2013)	15 Words	Expository	300
Yoshii (2014)	20 Words	Expository	300
Yusuf et al. (2014)	13 Words	Narrative	NR
Zahar et al. (2001)	30 Words	Narrative	2,283
Zarei & Hasani (2011)	30 Words	Textbook	NR
Zhao et al. (2016)	20 Words	Expository	906

Note: When targeted words include additional information, it includes the language of the term or whether words are nonce: non-existent words fabricated for the sake of the study. Non-English and nonce words are sometimes incorporated to ensure that participants do not have previous knowledge of the term. Word counts with N/A refer to studies not requiring students to read or writing text longer than one sentence, whereas NR refers to studies where word counts may be expected but are not reported.

Abbreviations for Pretests, Posttests, and Language Proficiency Tests

Abbreviation	Test
Local	Created by researchers or used locally
CEFR	Common European Framework Reference
CPE	Cambridge Proficiency Test
EPT	English Placement Test, University of Michigan
Gates-Mac	Gates-MacGinitie Reading Test
IELTS	International English Language Testing System
LKB	Lexical Knowledge Battery
Nelson	Nelson Placement Test
Oxford	Oxford Placement Test
SR	The State Ratings Test
TOEFL	Test of English as a Foreign Language
VKS	Vocabulary Knowledge Scale
VLT	Vocabulary Levels Test
Writing	Vocabulary used in Writing

Note: Local tests may include those created by researchers, tests only used within the institution where the study took place, and tests recognized only nationally and used only within the country where the study took place.

Vocabulary Acquisition Study Participants

Study	#	L1	M/F	Context	Proficiency Level(s)	PA
AbuSeileek (2011)	80	Arabic	M	FL-Tertiary	≥ Intermediate	Local
Ahmed et al. (2014)	316	English	M,F	NES-Primary	-	-
Alavi & Keyvanshekouh (2012)	38	Persian	M,F	FL-Tertiary	Not Reported	OPT
Alavinia & Rahimi (2019)	165	Persian	M,F	FL-Tertiary	< Intermediate	OPT
Alharbi (2016)	140	Arabic	M	FL-Tertiary	Not Reported	TOEFL
Alharbi (2018)	72	Arabic	-	FL-Tertiary	≥ Intermediate	Local
Alsaif & Masrai (2019)	1	Arabic	M	FL-Tertiary	≥ Intermediate	CEFR
Amirian & Behshad (2016)	14	Persian	-	FL-Tertiary	≥ Intermediate	-
Arnaud & Savignon (1997)	236	French	M,F	FL-Adults	≥ Intermediate	-
Azari et al. (2012)	76	Persian	M,F	FL-Tertiary	< Intermediate	TOEFL
Bai (2016)	50	Chinese	-	FL-Tertiary	≥ Intermediate	VLT
Bao (2015)	153	Chinese	M,F	FL-Tertiary	< Intermediate	VLT
Cao (2013)	70	Chinese	-	FL-Tertiary	Not Reported	-
Çekiç & Bakla (2019)	77	Turkish	-	FL-Tertiary	≥ Intermediate	CEFR
Cheng & Good (2009)	135	Chinese	-	FL-Tertiary	Multiple	EPT
Choi (2017)	178	Korean	M,F	FL-Tertiary	Multiple	-
Cunningham & Stanovich (1997)	134	English	M,F	NES-Multi	Multiple	Local
Dabaghi Varnosfadrani & Rafiee (2012)	59	Persian	M,F	FL-Tertiary	Not Reported	VLT
Daskalovska (2014)	18	Greek	M,F	FL-Tertiary	≥ Intermediate	VLT
Dobbs & Kearns (2016)	167	English	-	NES-Second	-	-
Duan (2018)	89	Chinese	-	FL-Tertiary	≥ Intermediate	VKS
Duff et al. (2015)	488	English	M,F	NES-Multi	-	-
Eckerth & Tavakoli (2012)	30	Multiple	M,F	SL-Tertiary	≥ Intermediate	IELTS
Ehsanzadeh (2012)	33	Persian	M,F	FL-Tertiary	≥ Intermediate	TOEFL
El-Dakhs et al. (2017)	81	Arabic	M,F	FL-Tertiary	< Intermediate	IELTS
Far (2016)	41	Persian	F	FL-Tertiary	≥ Intermediate	TOEFL
Fathi & Sarkhosh (2019)	120	Multiple	M,F	FL-Second	< Intermediate	-
Frishkoff et al. (2008)	37	English	M,F	NES-Tertiary	-	-
Ghorbani & Rahmandoost (2012)	60	Persian	-	FL-Tertiary	< Intermediate	Local
Gohar et al. (2018)	90	Persian	-	FL-Tertiary	≥ Intermediate	TOEFL
Hazrat (2015)	39	Persian	-	FL-Second	Not Reported	IELTS
Heidari-Shahreza & Tavakoli (2016)	90	Persian	-	FL-Tertiary	≥ Intermediate	OPT
Horst et al. (1998)	34	Arabic	-	FL-Tertiary	≥ Intermediate	OPT
Huang (2015)	-	Chinese	M,F	FL-Tertiary	Not Reported	Local
Hulstijn & Laufer (2001)	177	Multiple	M,F	FL-Tertiary	≥ Intermediate	-
Joe (2010)	1	Turkish	M	FL-Tertiary	≥ Intermediate	VLT
Jung (2016)	52	Korean	M,F	FL-Tertiary	Not Reported	CPE
Kelley et al. (2010)	476	Multiple	M,F	NES, SL-Sec	-	-
Khan (2019)	-	Multiple	-	SL-Adults	Multiple	-
Khatib & Faruji (2012)	60	Persian	M,F	FL-Tertiary	≥ Intermediate	TOEFL
Khoshsima & Eskandari (2017)	76	Persian	M,F	FL-Tertiary	Homogeneous	OPT

Study	#	L1	M/F	Context	Proficiency Level(s)	PA
Kim (2011)	64	Multiple	M,F	SL-Tertiary	Multiple	TOEFL
Kweon & Kim (2008)	12	Korean	M,F	FL-Tertiary	≥ Intermediate	TOEFL
Larson (2014)	222	English	M,F	NES-Second	-	-
Laufer & Rozovski-Roitblat (2011)	20	Multiple	M,F	FL-Tertiary	≥ Intermediate	Local
Lee (2003)	63	Multiple	M,F	SL-Second	≥ Intermediate	Gates
Lee et al. (2016)	80	Korean	-	FL-Tertiary	≥ Intermediate	TOEIC
Lee & Muncie (2006)	48	Multiple	M,F	SL-Second	≥ Intermediate	Gates
Lessard-Clouston (2012)	12	Multiple	M,F	NES, SL-Sec	-	Local
Levitzsky & Laufer (2013)	290	Hebrew	M,F	FL-Multi	-	-
Lin & Kawai (2016)	-	-	-	-	<i>Not Reported</i>	-
Liu (2017)	60	Chinese	M,F	FL-Tertiary	Homogeneous	Local
Ma (2013)	2	Chinese	-	FL-Tertiary	≥ Intermediate	Local
Mancilla-Martinez & Lesaux (2010)	49	Multiple	M,F	SL-Second	-	-
Min (2008)	25	Chinese	M	FL-Second	≥ Intermediate	VKS
Muncie (2002)	25	Japanese	M,F	FL-Tertiary	≥ Intermediate	LFA
Nagy et al. (1987)	352	English	-	NES-Multi	-	-
Parry (1997)	2	Multiple	M,F	SL-Tertiary	<i>Not Reported</i>	Local
Pellicer-Sánchez & Schmitt (2010)	20	Spanish	M,F	FL-Tertiary	≥ Intermediate	<i>Est.</i>
Pereyra (2015)	7	Spanish	-	FL-Adults	≥ Intermediate	CEFR
Pichette et al (2012)	203	French	-	SL-Tertiary	≥ Intermediate	-
Ponniah (2011)	49	Tamil	-	SL-Tertiary	≥ Intermediate	-
Pourakbari & Biria (2015)	150	Persian	-	FL-Tertiary	<i>Not Reported</i>	OPT
Reynolds (2014/2015/2016)	52	Chinese	-	FL-Tertiary	≥ Intermediate	<i>Est.</i>
Roessingh et al. (2005)	47	Multiple	-	SL-Second	Intermediate	-
Samian et al. (2016)	49	Persian	M,F	FL-Tertiary	≥ Intermediate	CEFR
Schmitt et al. (2004)	94	Multiple	M,F	NES-Tertiary	≥ Intermediate	TOEFL
Schoonen et al. (2011)	389	Dutch	M,F	FL-Multiple	-	-
Severino & Deifell (2011)	1	Chinese	M	SL-Tertiary	≥ Intermediate	-
Snoder (2017)	59	Swedish	-	FL-Second	≥ Intermediate	CEFR
Sobel et al. (2010)	39	English	M,F	NES-Second	-	-
Solati-Dehkordi & Salehi (2016)	30	Persian	F	FL-Tertiary	≥ Intermediate	OPT
Soleimani & Rahmanian (2015)	70	Persian	-	FL-Tertiary	≥ Intermediate	Nelson
Sun (2014)	35	Chinese	-	FL-Tertiary	Multiple	-
Tekman & Daloğlu (2006)	99	Turkish	-	FL-Tertiary	≥ Intermediate	FCE
Teng (2016)	36	Chinese	M,F	FL-Tertiary	≥ Intermediate	-
Teng (2017)	77	Chinese	M,F	FL-Tertiary	<i>Not Reported</i>	-
Touti & Maleki (2016)	64	Persian	F	FL-Tertiary	≥ Intermediate	Local
Varandi & Faezi (2013)	30	Persian	-	FL-Tertiary	≥ Intermediate	Nelson
Varol & Erçetin (2016)	90	Turkish	-	FL-Tertiary	≥ Intermediate	OPT
Varrick (2016)	5	Multiple	-	SL-Adults	< Intermediate	-
Vela (2015)	120	Multiple	-	FL-Tertiary	< Intermediate	CEFR
Walters & Bozkurt (2009)	60	Turkish	M,F	FL-Tertiary	< Intermediate	-
Waring & Takaki (2003)	14	Japanese	F	FL-Tertiary	≥ Intermediate	VLT
Webb (2007)	121	Japanese	-	FL-Tertiary	≥ Intermediate	Local

Study	#	L1	M/F	Context	Proficiency Level(s)	PA
White (2015)	141	English	M,F	NES-Multi	-	-
Wolsey (2010)	NS	English	-	NES-Second	<i>Not Reported</i>	-
Wu (2018)	100	Chinese	M,F	FL-Tertiary	< Intermediate	-
Xu (2010)	122	Chinese	-	FL-Tertiary	Multiple	-
Yang et al. (2017)	81	Chinese	M,F	FL-Tertiary	≥ Intermediate	CEFR
Yoshii (2014)	39	Japanese	-	FL-Tertiary	<i>Not Reported</i>	-
Yoshii (2013)	31	Japanese	M,F	FL-Tertiary	< Intermediate	TOEIC
Yusuf et al. (2014)	99	Bahasa	-	FL-Tertiary	Multiple	Grades
Zahar et al. (2001)	144	French	M	SL-Second	Multiple	VLT
Zarei & Hasani (2011)	158	Persian	M,F	FL-Tertiary	< Intermediate	MTELP
Zhao et al. (2016)	129	Chinese	M,F	FL-Tertiary	Multiple	<i>Local</i>

Note: PA refers to proficiency assessment, whereas FL refers to participants learning English in foreign language contexts. SL refers to participants learning English in a second language context. NES refers to participants who are native English speakers, and Tertiary refers to post-secondary education and may include both graduate and undergraduate students. Please see Appendix A for an explanation of proficiency test abbreviations.

Vocabulary Acquisition Study Foci

Study	Study Focus	Topic
AbuSeileek (2011)	Reading	Gloss Location
Ahmed et al. (2014)	N/A	Development
Alavi & Keyvanshekouh (2012)	Reading	Task Type
Alavinia & Rahimi (2019)	Reading, Writing	Task Type
Alharbi (2016)	Reading	Gloss Type
Alharbi (2018)	Reading	Gloss Language
Alsaif & Masrai (2019)	Reading	Attitudes towards Reading
Amirian & Behshad (2016)	Reading, Writing	Reading→Writing
Arnaud & Savignon (1997)	N/A	Development
Azari et al. (2012)	Reading	Gloss Language
Bai (2016)	Reading, Writing	Task Type
Bao (2015)	Reading	Task Type
Cao (2013)	Reading, Writing	Task Type
Çekiç & Bakla (2019)	Reading	Dispersion
Cheng & Good (2009)	Reading	Gloss Language
Choi (2017)	N/A	Development
Cunningham & Stanovich (1997)	N/A	Development
Dabaghi Varnosfadrani & Rafiee (2012)	Reading	Gloss Language
Daskalovska (2014)	Reading	Authentic Text
Dobbs & Kearns (2016)	N/A	Development
Duan (2018)	Reading	Gloss Types
Duff et al. (2015)	N/A	Development
Eckerth & Tavakoli (2012)	Reading	Task Type, Frequency
Ehsanzadeh (2012)	Reading	Learner Proficiency
El-Dakhs et al. (2017)	Reading, Writing	Reading→Writing
Far (2016)	Reading	Genre, Text Length
Fathi & Sarkhosh (2019)	Reading	Multiple Factors
Frishkoff et al. (2008)	Reading	Dispersion
Ghorbani & Rahmandoost (2012)	Reading	Task Type
Gohar et al. (2018)	Reading, Writing	Task Type
Hazrat (2015)	Reading, Writing	Task Type
Heidari-Shahreza & Tavakoli (2016)	Reading	Frequency, Lexicalization
Horst et al. (1998)	Reading	Task Type
Huang (2015)	N/A	Development
Hulstijn & Laufer (2001)	Reading, Writing	Task Type
Joe (2010)	Reading	Frequency
Jung (2016)	Reading	L1 Gloss
Kelley et al. (2010)	N/A	Development
Khan (2019)	Reading, Writing	Use of Corpora
Khatib & Faruji (2012)	Reading, Writing	Story Maps
Khoshsima & Eskandari (2017)	Reading, Writing	Task Type
Kim (2011)	Reading, Writing	Task Type

Study	Study Focus	Topic
Kweon & Kim (2008)	Reading	Extensive Reading
Larson (2014)	Reading, Writing	Reading→Writing
Laufer & Rozovski-Roitblat (2011)	Reading	Task Type Frequency
Lee (2003)	Reading, Writing	Reading→Writing
Lee & Muncie (2006)	Reading, Writing	Reading→Writing
Lee et al. (2016)	Reading	Gloss Types
Lessard-Clouston (2012)	N/A	Development
Levitzsky & Laufer (2013)	N/A	Development
Lin & Kawai (2016)	Reading, Writing	Task Type
Liu (2017)	Reading	Gloss Language
Ma (2013)	Writing	Development
Mancilla-Martinez & Lesaux (2010)	N/A	Development
Min (2008)	Reading	Task Type
Muncie (2002)	Writing	Development
Nagy et al. (1987)	Reading	Context
Parry (1997)	Reading	Notebooks
Pellicer-Sánchez & Schmitt (2010)	Reading	Frequency
Pereyra (2015)	Reading, Writing	Extensive Reading
Pichette et al. (2012)	Reading, Writing	Task Type
Ponniah (2011)	Reading	Task Type
Pourakbari & Biria (2015)	Reading, Writing	Task Type
Reynolds (2014/2015/2016)	Reading	Learner's L1, NES or ELL
Roessingh et al. (2005)	N/A	Development
Samian et al. (2016)	Reading, Writing	Gloss, Writing
Schmitt et al. (2004)	Reading	Multiple Factors
Schoonen et al. (2011)	N/A	Development
Severino & Deifell (2011)	N/A	Development
Snoder (2017)	Reading	Dispersion
Sobel et al. (2010)	Reading	Dispersion
Solati-Dehkordi & Salehi (2016)	Reading, Writing	Reading→Writing
Soleimani & Rahmanian (2015)	Reading, Writing	Task Type
Sun (2014)	Reading	Frequency
Tekman & Daloglu (2006)	Reading	Frequency, Proficiency
Teng (2016)	Reading	Frequency
Teng (2017)	Reading	Task Type, Metacognition
Touti & Maleki (2016)	Reading, Writing	Task Type
Varandi & Faezi (2013)	Reading	Proficiency
Varol & Erçetin (2016)	Reading	Gloss Type
Varrick (2016)	Reading, Writing	Task Type
Vela (2015)	Reading	Gloss Type
Walters & Bozkurt (2009)	Reading, Writing	Vocabulary Notebook Use
Waring & Takaki (2003)	Reading	Frequency
Webb (2007)	Reading	Frequency of Occurrence
White (2015)	N/A	Development

Study	Study Focus	Topic
Wolsey (2010)	Writing	Task Type
Wu (2018)	Reading	Cognitive Style
Xu (2010)	Reading	Gloss Language
Yang et al. (2017)	Reading, Writing	Task Type, Working Mem
Yoshii (2013)	Reading	Gloss Type
Yoshii (2014)	Reading	Gloss, Review
Yusuf et al. (2014)	Reading	Proficiency + Gloss type
Zahar et al. (2001)	Reading	Multiple
Zarei & Hasani (2011)	Reading	Gloss Location
Zhao et al. (2016)	Reading	Motivation, Anxiety

Note: The specific topic refers to the mediating factors examined as having an impact on vocabulary acquisition. Frequency refers to studies examining the frequency of exposure to individual target vocabulary and includes dispersion when how words are dispersed throughout a text is examined. Multiple refers to studies examining more than two mediating factors on vocabulary acquisition through reading, writing, or both. Development refers to studies examining vocabulary development longitudinally. Reading→Writing refers to studies requiring students to writing after reading a text with vocabulary items naturally found or seeded within.

Appendix B: Questionnaire

Pseudonym: _____

1. A first language is the language you believe you know best and the language you are most comfortable speaking. What do you consider your first language?
2. Your gender can be the sex you were born (male or female), but it does not have to be. What do you consider your gender to be right now?
3. What month and year did you first come to school in the United States?
4. How old were you when started going to school in the United States?
5. What is your race?
6. What is your ethnic group?
7. In which country were you born? If you were born in the United States, in which country did your parents come from?
8. A refugee is someone who had to leave their home country because they were not safe there. Are you a refugee?
9. Have you ever been a refugee?
10. How old were you when you first came to the United States?
11. How old are you now?
12. How old were you when you started to go to school for the first time?
13. Have you ever had to miss school for a long period of time? If so, how many total months or years did you miss school?

Appendix C: Adapted Vocabulary Knowledge Scale

(Brown, 2008; Wesche & Paribakht, 1996)

A = I've never seen this before.
B = I've seen this before, but I don't know what it means.
C = I know what it means but I'm not sure how to use it.
D = I know what this word/phrase means and I can use it in a sentence.
E = Here is an example of how to use it or what it means

	A	B	C	D	Example (if D)
1. a number of					
2. a wide range of					
3. as a result					
4. as well as					
5. at the same time					
6. can be found in					
7. in addition to					
8. in an attempt to					
9. in terms of					
10. in the case of					
11. in the context of					
12. in the field of					
13. on the other hand					
14. referred to as					
15. the degree to which					
16. the extent to which					
17. the majority of					
18. the nature of					
19. through the use of					
20. to ensure that					

Appendix D: Reading Passages and Writing Prompt Template¹³

Black Lives Matter

Many people want America to change. One group is called Black Lives Matter (BLM). It includes Black, brown, and white people. They want Americans to know that it is hard to be Black in America. Black people see racism in all parts of life. They want to be treated the same by police ^{and} **as well as** everyone else. Racism causes ^{many kinds} **a wide range of** problems in the US and the world. BLM began in 2013. But in 2020, it became very popular. Millions of people went to the streets to protest. It spread around the world. It is one of the biggest social movements in American history.

I read the paragraph.

Question *

Write a few sentences about BLM. You can write what you think about BLM, you can write about why it is popular, or you can write something else about BLM. Use “as well as” and “a wide range of” in your answer.

Escribe algunas oraciones sobre BLM. Puede escribir lo que piensa sobre BLM, puede escribir sobre por qué es popular o puede escribir algo más sobre BLM. Use “as well as” y “a wide range of” en su respuesta.

Écrivez quelques phrases sur BLM. Vous pouvez écrire ce que vous pensez de BLM, vous pouvez écrire pourquoi il est populaire ou vous pouvez écrire autre chose sur BLM. Utilisez « a wide range of » et « as well as ».

Bạn có thể viết những gì bạn nghĩ về BLM, bạn có thể viết về lý do tại sao nó được yêu thích, hoặc bạn có thể viết điều gì đó khác về BLM. Sử dụng “a wide range of” và “as well as” trong câu trả lời của bạn. of » et « as well as » dans votre réponse.

¹³ See Appendix E for links to original sources for each article.

GEN Z

Gen Z are people born between 1997–2012. How are Gen Z different from their parents and grandparents? They read less and spend more time on their phones. how much **The extent to which** a person reads affects how many words they know. This is why Gen Z also know fewer words. But they write more. They have more stress and also **in addition to** depression, and they do not get enough sleep. They worry about their futures more. They have fewer babies when they are teenagers, and they drink less alcohol. They get into trouble less often. More of them go to college, but they are less ready for college.

Implicit Ideas

Las creencias implícitas son ideas que tienes sobre el mundo sin saber por qué las tienes.

We all have ideas about the world. But most of our ideas are implicit. Because of this, This means we do not think about where the ideas came from. **As a result,** sometimes we do not know why we do the things we do. Everyone has implicit ideas. Sometimes, they come from things we see on T.V. or in movies. They also come from our life experiences and our families. But implicit beliefs can sometimes be a problem. about **In terms of** our behavior, implicit beliefs can make us treat others unfairly. Thinking about where our ideas about the world came from can help us treat others fairly.

What is a Social Movement?

Un movimiento social es un grupo de personas y organizaciones que se unen para cambiar la sociedad.

Social movements happen when people want change. Most social movements want to help people. This can mean poor people, women, or minority races. Social movements start when people are unhappy. People go to the streets **in addition to** **how much** **The extent to which** a social movement is successful depends on three things. 1. People must care and want change. 2. The movement must start with something small to change. 3. The leaders must get many people to join. The internet can help.

Youth Activism

Un activista es una persona que hace algo para hacer del mundo un lugar mejor.

There is a new kind of activist in 2022. Youth activists are people age 15-24 who want to change the world. They care about laws, health, and social problems. They care about children and the poor. They care about the environment. **In terms of** location, youth activists use the internet to share their ideas. They use Instagram, Facebook, and TikTok. Youth activism is popular, but it mostly happens in schools and online. **As a result,** adults may not know about it.

Education and Refugees

Los refugiados son personas que deben abandonar su país debido a una guerra, persecución o desastre.

Half of all refugees are children. School is very important for refugee children. **how much** **The extent to which** refugee children have hope may depend on whether they go to school. Schools are an important way to help refugee children have better lives. School can help people learn to fix problems **and also** caused by war **in addition to** stopping new wars. When refugee children go to school in a new country, it helps everyone. The other students learn to be kind to those who are different. They also learn about problems in the world.

Immigration

Immigration means the movement of people to a new country to live, to go to school, or to work. All Americans who are not Native American have ancestors who were immigrants. In American history, most changes in culture and population **because** have happened **as a result** of immigration.

About **In terms of** exact numbers, the United States has more immigrants than any other country. America had 47 million immigrants in 2015. That is 20% of all Americans. Half of all immigrants are relatives of American citizens, and 13% were refugees. mostly had a good effect on the American economy.

The Universal Declaration of Human Rights

Because of **As a result of** World War II, the United Nations wrote the **Universal Declaration of Human Rights**. It is very important. It tells everyone in the world about the rights and freedoms that all people have. It says that everyone must have free speech. It says that all genders, races, languages, and religions are equal. It says there should be no slavery and no torture. It also says that everyone should have healthcare, enough food, and a home to live in. **In terms of compliance**, many countries do not do a good job following the rules.

Compliance means how well people follow a rule or a law.

Climate Change

Scientists agree that human actions cause climate change. Coal, oil and gas are the problem. Farms, factories, and cutting down trees are also problems. As the world gets hotter, deserts become bigger. This means large fires **in addition to** less land for farms. Snow and ice melt. Sunlight hits earth, and water in the air means big storms and floods. Climate change means many animals will die. There will not be enough water for everyone. Many people will become homeless. The **extent to which** climate change will cause problems depends on how hot the world gets. It also depends on how fast change happens.

Dragon Con

many different

Every year in September, a **wide range of** people travel to Atlanta. They are fans of video games, comic books, and science fiction. They visit Dragon Con. Many fans dress up as anime characters and video game characters *and* **as well as** characters from comics, movies, and television. They have a parade to show everyone their interesting costumes. Famous authors and actors come to talk and answer questions. Over 80,000 people come to Dragon Con each year.

English Language Learners

called

Students who speak another language are **referred to as** English language learners or ELLs. This is a new way to think about students learning English. ELL means all people learning English. Most American students speak one language. Saying “ELL” shows that ELLs are learning something *new*. Saying “ELL” shows people that ELLs do not know *less* than other students. Instead, ELLs learn *more* than other students. There *many* are **a number of** ways to say “someone learning English.” Look below.

Other Ways of Saying ELL	Meaning
<i>non-native English students</i>	They were not born here.
<i>emergent bilingual students</i>	They are beginning to learn a new language.
<i>limited English proficient students</i>	They don't speak English very well.
<i>English as an additional language students</i>	They are learning a new language.
<i>English as a second language</i>	They are learning English but speak another first language.

What is Social Justice?

Social justice means fairness in the world. There are two kinds of social justice. One is about how society treats some groups of people unfairly. The other is about laws that hurt some groups of people. Social justice advocates want change for ^{many types} a **wide range** of people. This means rich and poor, ^{and} Black and brown, men **as well as** women, in all countries. People interested in social justice ask, “Which problems are big?” and “Which people have big problems?”

Fundamental Attribution Error

The different way we think about our actions compared to other people’s actions is ^{called} **referred to as** fundamental attribution error. When other people do something we do not like, we often believe it is their personality to act that way. We do not usually think that a situation caused it. The attribution error affects us in ^{many} a **number of** ways. Here is one example. If you see someone driving too fast, you probably will not think that they have an emergency. Instead, you may think that the person is selfish or always bad at driving.

Who is Greta Thunberg?

Greta Thunberg is a young woman from Sweden *and also* **as well as** an activist.

When she was 15, she did not go to school for many days. Instead, she went to her government. She stood outside with a sign. The sign said, “school strike for climate.” Greta was unhappy because her government did not make changes to *many different* stop climate change. Soon, **a wide range of** students around the world stopped going to school. Greta gave an angry speech to leaders around the world. She said, “How dare you? You have stolen my dreams and my childhood.” Greta has changed many people’s ideas about climate change. She has also won many awards.

Amnesty International

23

Amnesty International (AI) is a large organization who fights for human rights *many* for everyone. They fight **a number of** human rights abuses throughout the world. For example, in some countries, people may go to jail because of their religion. People may go to jail for wanting to change their government. These people are *called* **referred to as** political prisoners. Amnesty International demands that governments let them have freedom.

Doctors without Borders

In 1971, a group of doctors and writers helped people suffering in the Nigerian Civil War.

called

Today, the group is **referred to as** Médecins Sans Frontières (Doctors without Borders). They

many

help people around the world. They go to **a number of** places where war and disease cause suffering. They help people of all races and religions. They give people medical treatment, clean water, and food. They do not take money from governments. They tell the world about the problems they see. Doctors Without borders has helped people with many problems. Here is a map that shows where they help:



Appendix E: Wikipedia Article Links

Links to Wikipedia Source Articles

Text Title	Wikipedia Article
Gen Z	https://en.wikipedia.org/wiki/Generation_Z
Implicit Ideas	https://en.wikipedia.org/wiki/Implicit_stereotype
Social Movements	https://en.wikipedia.org/wiki/Social_movement
Youth Activism	https://en.wikipedia.org/wiki/Youth_activism
Refugees	https://en.wikipedia.org/wiki/Refugee
Immigration	https://en.wikipedia.org/wiki/Immigration
Climate Change	https://en.wikipedia.org/wiki/Climate_change
Universal Declaration of Human Rights	https://en.wikipedia.org/wiki/Universal_Declaration_of_Human_Rights
Dragon Con	https://en.wikipedia.org/wiki/Dragon_Con
Fundamental Attribution Error	https://en.wikipedia.org/wiki/Fundamental_attribution_error
What is an ELL?	https://en.wikipedia.org/wiki/English-language_learner
Social Justice	https://en.wikipedia.org/wiki/Social_justice
Black Lives Matter	https://en.wikipedia.org/wiki/Black_Lives_Matter
Implicit Beliefs	https://en.wikipedia.org/wiki/Bias
Meritocracy and the Poor	https://en.wikipedia.org/wiki/Meritocracy
Who is Greta Thunberg?	https://en.wikipedia.org/wiki/Greta_Thunberg
Amnesty International	https://en.wikipedia.org/wiki/Amnesty_International
Doctors without Borders	https://en.wikipedia.org/wiki/M%C3%A9decins_Sans_Frontiers

Appendix F: Interview Guide

Interview Guide

Script: *Today, I am going to ask you some questions about the reading and writing I asked you to do over the past nine weeks. Then I will ask you lots of questions about you and your life. I will say everything in English, but please let me know if you don't understand something. I can translate anything for you. If it is easier for you, you can answer in ----. In that case, I will translate what you say using iTranslate Voice3.*

I am going to record our entire conversation but your real name will not be attached to it. I will delete the recording as soon as I transcribe our conversation to paper. If you get tired during the interview and want to stop, just let me know. You can stop answering questions or you can tell me a better time when we can continue. Later, I might contact you to ask you a few more questions while I am looking at your answers. I will pay you \$100 for your time today at the end of the interview. I want to remind you that your answers will be kept private. Do you have any questions?

1. First, tell me what you think about the lessons I asked you to complete.

Probing Question: You said, "----." Please explain that part some more.

Probing Question: You said, "----." In what way?

2. Did you have any problems understanding the directions?

Probing Question: You said, "----." Please explain that part some more.

Probing Question: You said, "----." In what way?

Procedures

The first set of questions is about what you did during the lessons.

3. What exactly did you do for the reading part of the lesson?

Probing Question: You said, "----." What did you do right after that?

Probing Question: You said, "----." What did you do just before that?

Probing Question: You said, "----." Please explain that part some more.

4. What exactly did you do when you wrote the paragraphs?

Probing Question: You said, "----." What did you do right after that?

Probing Question: You said, "----." What did you do just before that?

Probing Question: You said, "----." Please explain that part some more.

5. What exactly did you do when you saw words in bold?

Probing Question: You said, "----." What did you do right after that?

Probing Question: You said, “---.” What did you do just before that?

Probing Question: You said, “---.” Please explain that part some more.

6. Did you have any interruptions during reading or writing?

Follow up: How often did you have interruptions?

Follow up: How long had you been reading?

Follow up: How long had you been writing?

Follow up: How long were you away from the lesson?

Follow up: How long did you work on it when you came back?

Multiple Factors

These questions are about things that affect learning English.

7. What is hard about reading in English?

Probing Question: Why is that hard?

Probing Question: You said, “-----.” Please explain that part some more.

8. What is hard about writing in English?

Probing Question: Why is that hard?

Probing Question: You said, “-----.” Please explain that part some more.

9. What is easy about reading in English?

Probing Question: Why is that easy?

Probing Question: You said, “-----.” Please explain that part some more.

10. What is easy about writing in English?

Probing Question: Why is that easy?

Probing Question: You said, “-----.” Please explain that part some more.

11. What was the hardest thing about the lessons?

Probing Question: Why do you think that was difficult for you?

Probing Question: You said, “-----.” Please explain that part some more.

12. What was the easiest thing about the lessons?

Probing Question: Why do you think that was hard for you?

Probing Question: You said, “-----.” Please explain that part some more.

13. In lesson --, you wrote about ---. Tell me more about that.

Probing Question: You said, “---.” Please explain that part some more.

14. Is there anything else you would like to say about things that make learning English easy or hard?

Probing Question: You said, "-----." Please explain that part some more.

Educational Factors

The next set of questions are about how school can help you learn English or make it hard to learn English.

15. Tell me about your school.

Probing Question: You said, "-----." Please explain that part some more.

Probing Question: Tell me about your teachers.

Probing Question: Tell me about the students.

Probing Question: Tell me about how you learn English at school.

Probing Question: Tell me about your ESOL teachers.

16. Are there things at school that make learning English hard?

Follow up: What types of things? Tell me more.

Probing Question: You said, "-----." Please explain that part some more.

17. Are there things at school that can make learning English easy?

Follow up: What things? Tell me about things at school that help you learn English.

Probing Question: Tell me more about that.

Probing Question: You said, "---." Please explain that part some more.

18. Tell me about things at school that sometimes make learning English hard.

Probing Question: Why is that?

Probing Question: You said, "---." Please explain that part some more.

19. Sometimes there are differences in how people are treated at school. Do you think you are treated or viewed in any special way at school, by students, teachers, or anyone else?

Follow up: In what ways?

Follow up: Who does this?

Probing Question: You said, "---." Please explain that part some more.

Probing Question: Can you give me an example?

20. Let's go back and look at some of what you wrote. I noticed that in lesson (s)---- you wrote less (or more) than other lessons. Why do you think that is?

Probing Question: You said, "-----." Please explain that part some more.

Probing Questions: Why do you think it was --- (adjective used to describe)?

Responsibilities

The next set of questions are about responsibilities you have that aren't related to your education.

21. Do you have a job?

Follow up: What is your job?

Follow up: What do you do at work?

Follow up: How many hours do you work?

Follow up: Are your hours usually the same, and what are they?

22. Tell me what your average day is like. What happens when you get home from school?

Probing Question: What do you do after that?

Probing Question: What do you do before that?

23. Tell me your weekends.

Probing Question: What do you do after that?

Probing Question: What do you do before that?

Family

A person's family can sometimes have a lot to do with how a person thinks about school and learning, so the next questions are about family.

24. Tell me about your family.

Probing Question: Who do you live with?

Probing Question: Who do you spend time with?

25. Are there things at home that make learning English hard?

Follow up: What things?

Probing Question: You said Tell me more about --.

26. Are there things at home that sometimes make learning English easy?

Follow up: What things? Tell me more.

27. What types of things do your (mom/dad/guardian/etc.) believe about school?

Probing Question: You said --. Tell me more about that.

Past Experiences

This is the last set of questions. These questions are about your past. I want to remind you that anything you say is confidential.

28. How did you come to the United States?

Probing Question: Tell me more about that.

29. Tell me about the big ways that life is different in the United States.

Probing Question: Tell me more about that.

(for refugee students)

30. In your questionnaire, you said that you are a refugee student.

Probing Question: Tell me more about that.

Probing Question: What was school like in ... ?

(for students with limited formal schooling)

31. In your questionnaire, you said that you have gone to school for ... years. Tell me more about the times you were not in school.

Probing Question: What did you do during that time?

Probing Question: How old were you when you stopped going to school?

Probing Question: Why did you stop going to school?

Follow up Question: What was happening in your country when you stopped going to school?

Follow up Question: What was happening in your family when you stopped going to school?

Follow up Question: What was happening in your life when you stopped going to school?

Appendix G: Data Set

Group	L1	Gender	Read1	Read2	Read3	Write1	Write2	Write3
2.00	NonEng	Female	6.00	8.00	8.00	6.00	6.00	5.00
1.00	English	Female	8.00	10.00	10.00	10.00	10.00	12.00
2.00	NonEng	Female	6.00	7.00	9.00	3.00	6.00	4.00
1.00	English	Female	10.00	11.00	11.00	11.00	12.00	12.00
2.00	English	Female	10.00	12.00	12.00	8.00	11.00	10.00
1.00	English	Male	3.00	7.00	6.00	3.00	10.00	10.00
2.00	English	Male	2.00	12.00	10.00	4.00	11.00	10.00
2.00	NonEng	Female	8.00	10.00	10.00	10.00	10.00	11.00
1.00	English	Male	2.00	10.00	11.00	1.00	12.00	11.00
1.00	NonEng	Female	10.00	11.00	10.00	10.00	12.00	12.00
2.00	English	Female	8.00	8.00	10.00	7.00	8.00	8.00
2.00	NonEng	Female	8.00	12.00	12.00	8.00	12.00	11.00
2.00	English	Male	7.00	8.00	10.00	7.00	7.00	8.00
1.00	NonEng	Male	7.00	11.00	11.00	9.00	12.00	11.00
1.00	NonEng	Male	7.00	10.00	10.00	8.00	12.00	11.00
2.00	NonEng	Male	9.00	11.00	12.00	7.00	8.00	9.00
1.00	English	Male	6.00	9.00	10.00	3.00	11.00	12.00
1.00	NonEng	Male	9.00	7.00	8.00	6.00	12.00	12.00
2.00	English	Female	8.00	11.00	11.00	7.00	8.00	9.00
1.00	NonEng	Male	8.00	8.00	8.00	9.00	9.00	12.00
1.00	English	Male	9.00	12.00	7.00	8.00	11.00	12.00
2.00	English	Female	5.00	11.00	10.00	5.00	7.00	10.00
1.00	English	Male	6.00	9.00	8.00	7.00	6.00	9.00
1.00	NonEng	Male	7.00	8.00	6.00	9.00	5.00	9.00
1.00	English	Male	9.00	9.00	10.00	11.00	12.00	12.00
2.00	English	Male	6.00	12.00	12.00	6.00	10.00	10.00
2.00	NonEng	Female	5.00	9.00	11.00	9.00	10.00	9.00
2.00	NonEng	Female	6.00	10.00	10.00	3.00	9.00	7.00
1.00	NonEng	Female	7.00	6.00	7.00	5.00	6.00	11.00
1.00	English	Male	8.00	10.00	9.00	10.00	10.00	12.00
2.00	NonEng	Female	9.00	6.00	9.00	8.00	8.00	6.00
2.00	NonEng	Male	5.00	7.00	7.00	8.00	8.00	5.00
1.00	NonEng	Male	7.00	9.00	10.00	8.00	12.00	11.00
2.00	NonEng	Female	7.00	10.00	11.00	8.00	11.00	9.00
2.00	English	Male	9.00	11.00	11.00	9.00	11.00	9.00
2.00	NonEng	Female	3.00	10.00	9.00	6.00	11.00	10.00
1.00	English	Female	4.00	7.00	5.00	4.00	9.00	6.00
1.00	NonEng	Female	8.00	12.00	11.00	9.00	12.00	11.00
1.00	NonEng	Female	7.00	9.00	10.00	8.00	10.00	10.00
2.00	English	Female	7.00	8.00	10.00	8.00	9.00	9.00
2.00	NonEng	Female	.00	6.00	3.00	.00	7.00	6.00
1.00	NonEng	Female	4.00	11.00	12.00	7.00	11.00	11.00

1.00	English	Female	9.00	7.00	9.00	8.00	10.00	12.00
1.00	NonEng	Female	6.00	6.00	4.00	5.00	12.00	9.00
2.00	NonEng	Male	7.00	11.00	9.00	8.00	11.00	10.00
2.00	English	Female	7.00	10.00	12.00	6.00	7.00	9.00
1.00	NonEng	Male	4.00	6.00	4.00	3.00	11.00	7.00
2.00	NonEng	Male	7.00	11.00	10.00	8.00	8.00	11.00
2.00	NonEng	Female	11.00	11.00	11.00	8.00	11.00	10.00

Appendix H: Z Scores

ZScores1	ZScores2	ZScores3
-0.34843	-1.76254	-2.06952
0.4844	0.25179	0.2457
-1.59769	-1.76254	-2.53256
1.31724	0.75537	0.70874
0.4844	0.75537	0.2457
-1.59769	-1.25895	-1.60647
-2.0141	1.25895	0.2457
0.4844	0.25179	0.2457
-2.0141	0.25179	0.70874
1.31724	0.75537	0.2457
0.4844	-0.75537	0.2457
0.4844	1.25895	1.17178
0.06799	-0.75537	0.2457
0.06799	0.75537	0.70874
0.06799	0.25179	0.2457
0.90082	0.75537	1.17178
-0.34843	-0.25179	0.2457
0.90082	-1.25895	-0.68039
0.4844	0.75537	0.70874
0.4844	-0.75537	-0.68039
0.90082	1.25895	-1.14343
-0.76485	0.75537	0.2457
-0.34843	-0.25179	-0.68039
0.06799	-0.75537	-1.60647
0.90082	-0.25179	0.2457
-0.34843	1.25895	1.17178
-0.76485	-0.25179	0.70874
-0.34843	0.25179	0.2457
0.06799	-1.76254	-1.14343
0.4844	0.25179	-0.21735
0.90082	-1.76254	-0.21735
0.4844	-0.75537	-2.06952
0.06799	-0.25179	0.2457
0.06799	0.25179	0.70874
0.90082	0.75537	0.70874
-1.59769	0.25179	-0.21735
-1.18127	-1.25895	-2.06952
0.4844	1.25895	0.70874
0.06799	-0.25179	0.2457
0.06799	-0.75537	0.2457

-2.84694	-1.76254	-2.9956
-1.18127	0.75537	1.17178
0.90082	-1.25895	-0.21735
-0.34843	-1.76254	-2.53256
0.06799	0.75537	-0.21735
0.06799	0.25179	1.17178
-1.18127	-1.76254	-2.53256
0.06799	0.75537	0.2457
1.73366	0.75537	0.70874