Title
Adolescent language minority students' vocabulary growth: Exploring heterogeneity with multilevel analysis

Permalink
https://escholarship.org/uc/item/0105g4kv

Author
Hwang, Jin Kyoung

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Adolescent language minority students' vocabulary growth:
Exploring heterogeneity with multilevel analysis

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

Jin Kyoung Hwang

Dissertation Committee:
Assistant Professor Joshua Lawrence, Chair
Associate Professor Penelope Collins
Associate Professor Jeannette Mancilla-Martinez
Professor Jamal Abedi

2015

DEDICATION

To

My loves, Samuel and JJ

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I would like to thank my committee chair and academic advisor, Josh Lawrence. Throughout the doctoral program, you have supported me in pursuing my research interests and believed in me in times when I had doubts about myself as a researcher. You were always available to share your knowledge and gave me advice when I needed it the most. You were a wonderful mentor.

I would like to thank my committee members, Penelope Collins, Jeannette Mancilla-Martinez, Jamal Abedi. I was blessed to have worked with great scholars in the field. To Penelope Collins, thank you for being my co-advisor and guiding me through the program. I truly appreciate your continued support and encouragement in every stage of writing this dissertation. To Jeannette Mancilla-Martinez, I was very lucky to have you on my committee. Thank you for giving me insights on my dissertation work. To Jamal Abedi, thank you for serving on my committee and giving me valuable feedback in a timely manner.

I want to thank my close colleagues and friends: To Arena Lam, Cathy Tran, Kat Schenke, and Ken Lee (a.k.a. the Kakao group and A-team), thank you for being there for me and keeping me sane. Without your support and collaboration, I would not have made it through the program. To Youngmin Park and Soobin Yim, I am so blessed to have known you. Thank you for providing laughter and emotional support. I would also like to thank my current and former LIIG members, Alex Lin, Joanna Yau, Rachel Stumpf, Karen Taylor, and Grace Lin. Thank you all for giving me constructive feedback and support on my work during lab meetings. To Huy Chung, I am grateful that I got to know you during my graduate years. Thank you for encouraging me to do my best whenever I was feeling down.

Finally, I would like to thank my family for providing endless support. Thank you to my husband, JJ, and son, Samuel, for understanding me and sacrificing your time so that I could come this far. Many thanks to my parents who always believed in me and kept me in their prayers. I would also like to thank my father- and mother-in-law for their support.

Jin Kyoung Hwang

Ph.D. Candidate
3200 Education, School of Education
University of California, Irvine
Irvine, CA 92697-5500
Email: jkhwang1@uci.edu

EDUCATION

University of California, Irvine   2010-2015
Ph.D., Education, Specialization in Language, Literacy, and Technology
  Dissertation: Adolescent language minority students' vocabulary growth: Exploring
       heterogeneity with multilevel analysis.
  Committee members: Joshua F. Lawrence (chair), Penelope Collins, Jeannette Mancilla-
       Martinez, Jamal Abedi

Harvard Graduate School of Education   2008-2009
Ed.M., Focus in Language and Literacy

Sookmyung Women's University   2002-2007
B.A., English Language and Literature with Honors

University of Toronto   2005-2006
Exchange student, took courses in English and Linguistics

PUBLICATIONS

Articles in Peer-Reviewed Journals

Hwang, J. K., Lawrence, J. F., Mo, E., & Snow, C. E. (2015). Differential effects of a
    systematic vocabulary intervention on adolescent language minority students with varying
    levels of English proficiency. *International Journal of Bilingualism, 19*(3), 314-322. doi:
    10.1177/1367006914521698

Zheng, B., Warschauer, M., **Hwang, J. K.**, & Collins, P. (2014). Laptop use and science
    achievement among at-risk students. *Journal of Science Education and Technology, 23*(4),
    591-603. doi: 10.1007/s10956-014-9489-5

Collins, P., **Hwang, J. K.**, Zheng, B., & Warschauer, M. (2013). Writing with laptops: A quasi-
    experimental study. *Writing and Pedagogy, 5*(2), 203-230. doi: 10.1558/wap.v5i2.203

Bruckner, T. A., Domina, T., **Hwang, J. K.**, Gerlinger, J., Nguyen, T., Carpenter, C., &
    Wakefield, S. (2014). State-level education standards for substance use prevention programs
    in schools: A systematic content analysis. *Journal of Adolescent Health*. *54*(4), 467-473. doi:
    10.1016/j.jadohealth.2013.07.020

Book Chapters

Park, Y., Warschauer, M., Collins, P., **Hwang, J. K.**, & Vogel, C. (2013). Building awareness of language structures with visual-syntactic text formatting. In K. E. Pytash, & R. E. Ferdig (Eds.), *Exploring technology for writing and writing instruction* (pp. 21-36). Hershey, PA: IGI Global.

Manuscripts under Review

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (*under review*). Differential vocabulary growth trajectories among adolescent language minority students: A two-year longitudinal study.

Hwang, J. K., Lawrence, J. F., Snow, C. E., & Collins, P. (*under review*). Vocabulary and reading performances of reclassified fluent English proficient students.

Hwang, J. K., Collins, P., Zheng, B., & Warschauer, M. (*under review*). Exploring the writing patterns of elementary school students as a function of their English proficiency.

Manuscripts in Preparation

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (*in preparation*). Investigating the effects of Word Generation on adolescent language minority students: A longitudinal follow-up study.

Lawrence, J. F., Hagen, A. M., **Hwang, J. K.**, & Lervåg, A. (*in preparation*). Using multiple measures of vocabulary to predict reading comprehension.

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (*in preparation*). Reading growth trajectories of adolescent language minority students.

Kulesz, P., Francis, D., Lawrence, J. F., **Hwang, J. K.**, & Snow, C. E. (*in preparation*). The effects of target word characteristics on difficulty of vocabulary test items.

AWARDS

Chancellor's Club for Excellence Fellowship ($10,000)	2015
University of California, Irvine	
Doctoral Dissertation Grant ($2,000)	2014-2015
The International Research Foundation for English Language Education	
Korean Honor Fellowship ($1,000)	2014
Embassy of the Republic of Korea	
National Academy of Education/Spencer Dissertation Fellowship Finalist	2014
Sookmyung Leadership Fellowship	2003
Sookmyung Academic Fellowship	2003, 2004

SELECTED CONFERENCE PRESENTATIONS

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (2016, April*). Differential vocabulary growth trajectories among adolescent language minority students: A two-year longitudinal study*. Paper to be presented at the meeting of the American Education Research Association. Washington, D. C.

Hwang, J. K., & Lawrence, J. F. (2016, April). *Unpacking reading performances of reclassified fluent English proficient students*. Paper to be presented at the meeting of the American Education Research Association. Washington, D. C.

Hwang, J. K., & Lawrence, J. F. (2015, December). *Differential vocabulary growth trajectories among adolescent language minority students: A two-year longitudinal study.* Paper presented at the meeting of the Literacy Research Association. Carlsbad, CA.

Lawrence, J. F., **Hwang, J. K.**, Dean, P., & Lawless, R. (2015, December*). Exploring the interface of language development and semantic knowledge with depth measures of academic vocabulary*. Paper presented at the meeting of the Literacy Research Association. Carlsbad, CA.

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (2015, July). *Differential treatment effects of an academic vocabulary intervention on adolescent language minority students: A two-year longitudinal study.* Poster presented at the meeting of the Society for Scientific Studies of Reading. Kailua Kona, HI.

Kulesz, P., **Hwang, J. K.**, Lawrence, J. F., & Francis, D. J. (2014, July). *Effects of word and reader characteristics on difficulty of vocabulary test items*. Paper presented at the meeting of the Society for Scientific Studies of Reading. Santa Fe, NM.

Hwang, J. K., & Lawrence, J. F. (2014, July). *Reading growth trajectories of adolescent language minority students with varying levels of English proficiency*. Paper presented at the meeting of the Society for the Scientific Study of Reading. Santa Fe, NM.

Hwang, J. K., Collins, P., Zheng, B., & Warschauer, M. (2014, April). *Exploring the writing patterns of elementary school students as a function of their English proficiency*. Paper presented at the meeting of the American Education Research Association. Philadelphia, PA.

Lawrence, J. F., Snow, C. E., Francis, D., Paré-Blagoev, E. J., & **Hwang, J. K.** (2014, March). *Effectiveness of a school-level intervention for academic language in ameliorating the consequences of low initial achievement*. Paper presented at the spring meeting of the Society for Research on Educational Effectiveness. Washington, D. C.

Hwang, J. K., Tran, C., & Conley, A.M. (2013, April). *How do levels of English proficiency relate to mathematics achievement and motivation?* Poster presented at the meeting of the Society for Research in Child Development. Seattle, WA.

Hwang, J. K., Lawrence, J. F., Mo, E., & Snow, C. E. (2013, April). Differential effects of a systematic vocabulary intervention on adolescent language minority students with varying levels of English proficiency. In Palincsar, A. S. (Chair), *Investigations of levers to enhance the oral language and reading comprehension of English language learners.* Symposium conducted at the meeting of the American Education Research Association. San Francisco, CA.

Hwang, J. K., Lawrence, J. F., Mo, E., & Hurley, P. (2012, November). Heterogeneous effects of Word Generation on students with differing levels of English proficiency. In Lawrence, J. (Chair), *Academic vocabulary instruction across the content areas: Results from a randomized trial of the Word Generation Program.* Symposium conducted at the meeting of the Literacy Research Association. San Diego, CA.

Hwang, J. K., Collins, P., Warschauer, M., Farkas, G., & Zheng, B. (2012, July). *Exploring the writing patterns of elementary school students as a function of their proficiency in English.* Poster presented at the meeting of the Society for the Scientific Study of Reading. Montreal, Canada.

Mo, E., Lawrence, J., & **Hwang, J. K.** (2012, April). *Systematic academic vocabulary development in middle schools.* Paper presented at the meeting of the American Education Research Association. Vancouver, Canada.

RESEARCH EXPERIENCE

Graduate Student Researcher and Project Researcher,   2012-Present
Joshua F. Lawrence, David Francis
   Word Generation: An Efficacy Trial (PI: Catherine Snow, Harvard University)
   Funding agency: Institute of Education Sciences
   - Multilevel modeling and item response theory analysis
   - Presented at national and international conferences
   - Published and preparing manuscripts for peer-reviewed journals

Project Researcher, *Jamal Abedi*   Summer, 2014
   Formative Assessment in Mathematics (PI: Jamal Abedi, UC Davis)
   Funding agency: National Science Foundation

Project Researcher, *Jamal Abedi*   2013-2014
   Impact of the WRITE Program on English Learner Achievement and Teacher Instructional Practice (PI: Eric Haas, WestEd)
   Funding agency: Institute of Education Sciences

Project Researcher, *Penelope Collins*   2011-2012
   Interactive Science and Technology Instruction for English Learners (PI: Mark Warschauer, UC Irvine)
   Funding Agency: National Science Foundation RAPID
   - Multivariate analysis of variance
   - Presented at conferences and prepared manuscripts for peer-reviewed journals

Graduate Student Researcher, *Thurston Domina*   2011-2012
   Human Capital Interventions across Childhood and Adolescence (PD: Greg J. Duncan, UC Irvine)
   Sub-project II: State Standards for Health Education Curricula Involving Alcohol, Tobacco, and Other Drugs (PI: Christopher Carpenter, Vanderbilt University)
   Funding agency: National Institute of Child Health and Human Development P01

Research Assistant, *Paola Uccelli*   2009

TEACHING EXPERIENCE

Instructor
Teaching Models of English, Sookmyung Women's University, Korea   Summer, 2011

Teaching Assistant

ED 40 Theories of Development and Learning, UC Irvine	Spring, 2015; Fall, 2015
ED 205 Critical Assessment of Teaching Practice and Learning, UC Irvine	Summer, 2015
ED151 & PSY BEH 192V, Language and Literacy, UC Irvine	Fall, 2014; Winter, 2015
ED 206, Designing Learning Environments, UC Irvine	Summer, 2013
ED107, Child Development in Education, UC Irvine	Spring, 2011
ED124, Multicultural Education, UC Irvine	Winter, 2011

INVITED TALKS

"Literacy Development in English Learners"　　　　　　　　June, 2013
　　English Learner Reading and Research Group Workshop, Claremont Graduate University

"Academic English and English Language Learners"　　　　February, 2011
　　ED 124 Multicultural Education in K-12 Schools, UC Irvine

RESEARCH SKILLS

Research Methods: Ordinary Least Squares & Logistic Regression, Analysis of Variance, Item Response Theory, Structural Equation Modeling, Multilevel Modeling

Statistical Software: STATA, SPSS, EQS, flexMIRT

Language: Native in Korean

PROFESSIONAL ORGANIZATION

American Educational Research Association

International Literacy Association

Literacy Research Association

Society for the Scientific Study of Reading

SERVICE

Peer Reviewer, International Journal of Bilingual Education and Bilingualism	
Proposal Reviewer, Annual Meeting of the Literacy Research Association	
Student Representative, Admissions Committee, UC Irvine	2012-2013
Literacy Assessor, Newport Beach Public Library Literacy Services	2009-2010

ABSTRACT OF THE DISSERTATION

Adolescent Language Minority Students' Vocabulary Growth:

Exploring Heterogeneity With Multilevel Analysis

By

Jin Kyoung Hwang

Doctor of Philosophy in Education

University of California, Irvine, 2015

Assistant Professor Joshua F. Lawrence, Chair

Three studies within this dissertation aim to provide insight on differential vocabulary

growth trajectories of adolescent language minority students across two years and examine how

they respond to a research-based academic vocabulary intervention, Word Generation. In all

three studies, the language minority student sample included initially fluent English proficient

(IFEP), redesignated fluent English proficient (RFEP), and limited English proficient (LEP)

students.

In Study 1, I investigated general vocabulary and academic vocabulary growth

trajectories of sixth- to eighth-grade English-only and language minority students using an

individual growth modeling analysis. Students were assessed at four time points on a

standardized measure of general vocabulary and a researcher-developed academic vocabulary

test. On both vocabulary measures, IFEP students slightly outperformed English-only students

on average, and English-only students scored higher than RFEP and LEP students at baseline.

Although there were differences in the general vocabulary growth trajectories across groups,

there were greater group differences in academic vocabulary growth. English-only students did not improve as much across the study.

In Study 2, I examined general vocabulary, academic vocabulary, and reading comprehension growth trajectories of adolescent RFEP students using an individual growth modeling analysis. Students completed up to four waves of reading-related measures during a two-year time period. Findings indicate that students' scores on the vocabulary and reading measures were positively correlated with their years since redesignation and students showed growth over time on average on all outcomes. The rate of growth did not differ by years since redesignation.

In Study 3, I investigated the longitudinal effects of an academic language intervention, Word Generation, on adolescent English-only and language minority students' word learning. Thirteen middle schools from an urban district in California were randomized to treatment and control conditions. Using individual growth modeling, I found main effect of treatment on students' academic vocabulary knowledge. In addition, students in the treatment condition maintained the improvement in their vocabulary knowledge in the follow-up year.

Findings from this dissertation underscore that language minority students represent a heterogeneous group of students with varying configurations of English language proficiency. They also indicate that both English-only and language minority students can benefit from an academic language intervention.

CHAPTER 1

Introduction

Language minority students are a diverse group of students who differ on various dimensions including their proficiency in English. They are generally classified based on their mastery of the English language, however, the classification criteria that each state uses vary (Abedi, 2008; Ragan & Lesaux, 2006). Language minority students who enter school with "full" English proficiency are referred to as initially fluent English proficient (IFEP). Those with limited English proficiency are classified as limited English proficient (LEP) students[1]. LEPs are required to receive language support from their schools. Redesignated fluent English proficient (RFEP) students are those who used to be LEPs but attained English language proficiency after some English language development instruction and were redesignated as being fluent. RFEP students no longer receive English development services from schools.

Language minority students are often treated as a homogeneous group in the research literature. Even when researchers do focus on language proficiency status, more attention has been paid to one subgroup of language minority students—LEPs—than others. Language minority students who are classified as being English proficient are the majority within the language minority student population in the United States, however, we still do not have a clear understanding of the literacy development of these students and how to best support them.

This three-study dissertation examines the vocabulary growth trajectories of subgroups of adolescent language minority students in their middle school years. There are two goals of this dissertation. The first goal is to provide insight on differential vocabulary growth trajectories across two years among subgroups of adolescent language minority students. This is to detail

---

[1] LEP students are also referred to as English language learners in the research literature.

heterogeneity within the language minority student population to help researchers and educators understand the relationship between students' English proficiency designation and their vocabulary outcomes. The second goal is to examine how adolescent language minority students respond to a research-based academic vocabulary intervention, Word Generation, and whether the treatment effects maintain or atrophy in the subsequent follow-up year. Below, I discuss briefly how the three studies meet these two goals.

Overview of the studies

Study 1. Differential vocabulary growth trajectories among adolescent language minority students: A two-year longitudinal study

In the first study of my dissertation, I examine general and academic vocabulary growth trajectories of subgroups of language minority students (i.e., IFEPs, RFEPs, and LEPs) across two years. Although I have fit models that show vocabulary knowledge to be unidimensional in nature, it is still possible for students to have different learning trajectories as a function of students' language proficiency status and word types. A general vocabulary measure (Vocabulary subtest of Gates-MacGinitie Reading Test [MacGinitie, MacGinitie, Maria, & Dreyer, 2000]) included a wide range of English words, whereas the words in the researcher-developed academic vocabulary measure tended to be mid-frequency words that are often encountered in texts across content areas (Coxhead, 2000). To address my research questions, I employed multilevel models for change (Singer & Willet, 2003) using four waves of students' responses ($N = 3,161$) in general and academic vocabulary measures. The results from Study 1 describe differential student vocabulary growth trajectories by language subgroups in their middle school years.

Study 2. Vocabulary and reading performances of redesignated fluent English proficient students

In the second study of my dissertation, I examine the literacy development of one subgroup of language minority students—the RFEPs. The goal of the second study is to describe differences in students' vocabulary and reading growth over time based on their years since redesignation. By definition, RFEP students are assumed to be proficient in English and not expected to receive any English language services. However, it is not clear in the research literature whether RFEP students' years since redesignation is or is not correlated with their literacy outcomes. In order to investigate students' vocabulary and reading growth over time, I employed multilevel models for change using four waves of RFEP students' data ($N = 1,226$). Study 2 describes heterogeneity within the RFEP students in regards to their performance in reading-related outcomes.

Study 3. Investigating the effects of Word Generation on adolescent language minority students: A longitudinal follow-up study

The third study examines the longitudinal treatment effect of the Word Generation program on adolescent language minority students. In this study, I first examined whether adolescent English-only and language minority students ($N = 5,052$) responded differently to Word Generation in the instructional year (2010-2011; Year 1). Additionally, I tested whether students who showed improvements in their academic vocabulary were able to maintain that vocabulary knowledge by the end of the follow-up year (2011-2012; Year 2). To answer my research questions, I employed multilevel models for change using four waves of students' scores on researcher-developed academic vocabulary test. The results from Study 3 provide

information about whether Word Generation is appropriate for teaching academic vocabulary words to students from diverse linguistic backgrounds and English proficiency levels.

Structure of the Dissertation

Each study within this dissertation will be laid out as a separate manuscript in turn. The three studies will have study-specific literature reviews, research questions, methods, discussions, and conclusions. I will end the dissertation by discussing the overarching conclusion that summarizes and includes the significance of the three studies and implications for future research.

References

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practice, 27*(3), 17-31. doi:10.1111/j.1745-3992.2008.00125.x

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

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading test technical report: Forms S and T*. Chicago, IL: The Riverside Publishing Company.

Ragan, A., & Lesaux, N. K. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives, 14*(20), 1-32. doi:10.14507/epaa.v14n20.2006

Singer, J., & Willet, J. (2003). *Applied longitudinal data analysis: Modeling change and even occurrence*. New York, NY: Oxford University Press.

CHAPTER 2

Study 1. Differential vocabulary growth trajectories among

adolescent language minority students: A two-year longitudinal study

There are a large and growing number of language minority (LM) students in American public schools. In 2011, over 60 million Americans over the age of five spoke a language other than English at home (Ryan, 2013) and that number is projected to increase in coming years (Shin & Ortman, 2011). LM students, like their English-speaking peers, may struggle with reading. In early grades, these students need to know how to phonologically represent a new language, understand a new system for representing sounds, and comprehend the decoded text. Fortunately, the vocabulary demands in early grade texts are low: most words in primary grade texts are high-frequency words that are semantically transparent. Texts in middle school are more complex. In secondary schools, children are expected to master abstract vocabulary words in reading (Leach, Scarborough, & Rescorla, 2003; Short & Fitzsimmons, 2007; Storch & Whitehurst, 2002). The Common Core State Standards (CCSS) may intensify the challenge LM students face with the increased emphasis on nonfiction texts, academic discussion, and argumentation (Bunch, Walqui, & Pearson, 2014; Common Core State Standards Initiative, 2010). We must understand the developmental trends in LM students' general language and academic language proficiency if we are to support them. In this study, we model students' general and academic vocabulary growth across the middle grades, a time when the proportion of academic words appearing in texts increases and knowledge of them is of critical importance for academic success (Corson, 1997; Nagy & Townsend, 2012; Townsend, Filippini, Collins, & Biancarosa, 2012).

Language designations

In this paper, we use the term *language minority student* to describe any school-aged student whose home language is not English (August & Shanahan, 2006). Within this extremely large and heterogeneous group, policy makers have identified three relevant subgroups (Ragan & Lesaux, 2006): initially fluent English proficient (IFEP), limited English proficient (LEP), or redesignated fluent English proficient (RFEP). IFEP students are those who enter school with enough English proficiency and who do not receive additional language support or assessment. LEP students do not have sufficient English proficiency to function in schools and so receive additional support from the schools they attend. RFEP students are those who were identified as LEP when they first entered school, but subsequently met proficiency criteria and were redesignated by their school (at which point they stopped receiving support).

State-level proficiency thresholds largely determine whether a student is identified as IFEP, RFEP, or LEP, as there is no specific guideline at the federal level for identifying and classifying LM students (Bailey & Kelly, 2013; Linquanti & Cook, 2013; Ragan & Lesaux, 2006). One empirical analysis of data from three states shows that the more stringent the classification criteria, the larger the difference between RFEP and LEP students is (Kim & Herman, 2009). Kim and Herman also found that in the state with the highest standards for redesignation, RFEP students actually outperform English-only (EO) students on some measures. Unfortunately, one consequence of a stringent policy may be that some students are never redesignated as being English proficient.

The data for the study presented here come from a large urban district in California. The state-level guidelines in California prescribed that LEP students be redesignated after obtaining either *early advanced* or *advanced* on the California English Language Development Test

(CELDT), which assesses students' overall English proficiency, as well as *proficient* or

advanced on the English language arts subtest of the California Standards Test (CST,

www.cde.ca.gov). LEP students take the CELDT annually until they are redesignated[2]. Along

with these test results, LEP students' grade point average and teacher interviews are also

considered in redesignation decisions. Though these state-level guidelines set a high bar for LEP

students to be reclassified, the ultimate decision is made by personnel in each district.

California's stringent reclassification criteria may be one of the reasons why the majority

(approximately 60%) of California secondary school LEP students are long-term LEP students

(Olsen, 2010; Parrish, Perez, Merickel, & Linquanti, 2006).

General vocabulary knowledge

It is well acknowledged that there is a strong relationship between second language (L2)

vocabulary knowledge and reading comprehension (August, Carlo, Dressler, & Snow, 2005;

Hoover & Gough, 1990; Lesaux, Lipka, & Siegel, 2006; Proctor, Carlo, August, & Snow, 2005;

Rydland, Aukrust, & Fulland, 2012). Although vocabulary knowledge is one of the foundational

skills for all students who *read to learn* new concepts across different content areas (Chall &

Jacobs, 2003), national statistics and research studies show that language minority students

(particularly those with limited English proficiency) tend to lag behind their EO counterparts in

English vocabulary knowledge (August & Shanahan, 2006; National Center for Education

Statistics, 2012). The few longitudinal studies that have investigated preadolescent and

adolescent EO and LM students' vocabulary growth confirm this (Jean & Geva, 2009; Lawrence,

---

[2] Up until the 2013-2014 school year, all students in California took the CST for the first time at the end of second grade. Therefore, LEP students could be reclassified as early as the middle of third grade when the second grade CST scores were available. With the implementation of Smarter Balanced Assessments (www.smarterbalanced.org) during the 2013-2014 academic year, the state of California no longer administers CSTs to their students. Thus, the redesigntion process does not use CST results anymore.

2012; Mancilla-Martinez & Lesaux, 2011a, 2011b). These studies also show that LM students may have different vocabulary growth trajectories than their EO peers during their elementary and middle school years. Lawrence (2012) examined the English vocabulary learning trajectories of sixth- and seventh-grade EO ($n = 210$) and LM ($n = 68$) students from an urban school district in the Northeast. LEP students were not included in his LM student sample. Using four waves of Group Reading Assessment and Diagnostic Evaluation data (Williams, 2001), he found that all students improved during the school year but experienced roughly a two-month summer setback in their vocabulary knowledge. Moreover, LM students regressed in vocabulary knowledge more severely than their EO peers during summer, but had steeper vocabulary learning trajectories during the school year. This pattern held even when well-known predictors of vocabulary (i.e., time spent on independent reading) and summer setback (i.e., socioeconomic status) were controlled for.

Mancilla-Martinez and Lesaux (2011b) modeled growth rates of Spanish-speaking LM students' ($N = 173$) word reading and vocabulary knowledge in English and Spanish from ages 4.5 to 11. In this study, the researchers did not classify LM learners into different subgroups. Their LM student sample included any student whose primary home language was Spanish. They assessed students' word reading and vocabulary knowledge with the Letter-Word Identification subtest and Expressive Vocabulary subtest from the Woodcock Language Proficiency Battery-Revised (Woodcock, 1991), respectively. They found that LM students' word reading growth rate and ability were on par with the national monolingual sample used to norm the test, but their vocabulary knowledge was consistently below the national norm.

Academic vocabulary knowledge

Some components of L2 vocabulary develop rapidly, while others, such as knowledge of abstract or academic terms, grow slowly (August & Shanahan, 2006; Collier, 1989; Cummins, 2008; Goldenberg, 2010; Mancilla-Martinez & Lesaux, 2011a). *General academic words* have been hypothesized to be particularly important for academic reading (Nagy & Townsend, 2012; Snow & Uccelli, 2009). They are the glue words of academic texts. Coxhead (2000) created the Academic Word List using both word frequency thresholds (i.e., the most frequent words in English were excluded) and dispersion thresholds (i.e., only words which appeared regularly across academic disciplines were included). Researchers have argued that because these words appear across academic disciplines with some regularity, knowledge of them is particularly useful for LEPs (Baumann & Graves, 2010; Coxhead, 2000; Nagy & Townsend, 2012).

Lawrence and his colleagues (Lawrence, Capotosto, Branum-Martin, White, & Snow, 2012) collected four waves of academic vocabulary data from treatment and comparison students participating in a quasi-experimental study of an academic language intervention, Word Generation. Using 11 anchor items embedded within each wave of testing, they created an item response theory (IRT) model that formed a time-varying level-1 outcome. Baseline scores indicated that LEP students had lower academic vocabulary scores than both EO students and English-proficient LM students. Interestingly, LM students who were not identified as LEP outperformed EO students at baseline. Although there were differences in response to program participation by language status, the growth trajectories of all groups of students in the comparison schools were the same; in other words, there was no difference in the rates that different groups of middle school students learned academic words in "business-as-usual" schools.

In a recent cross-sectional study, Uccelli, Galloway, Barr, Meneses, and Dobbs (2015) found that there was only a small difference between English-proficient and former English language learners on the Vocabulary Association Test (see Lesaux, Kieffer, Faller, and Kelley, 2010) across grade levels (+.2 SD favoring English proficient students in fourth grade, + .39 SD favoring former LEPs in fifth grade and no difference in sixth grade). In their study, Lesaux et al. (2010) found that although LM control students had lower academic vocabulary pretest scores than EO control students on average, they improved roughly one point by the end of the study, whereas the EO control students did not improve. Because of state-by-state redesignation policies, it is not possible to draw definite conclusions about baseline differences in academic vocabulary for language groups across studies. However, these studies suggest that redesignated students learn academic words as quickly as, or more quickly than, EO students even though their baseline scores might be lower.

Although LEP students' vocabulary knowledge is lower than their non-LEP counterparts (National Center for Education Statistics, 2012), it is not clear how proficient adolescent LM students (i.e., IFEP and RFEP students) are faring on reading-related outcomes in comparison to their EO peers. Nor is it clear if these differences are maintained over time, or are replicated in measures of student's academic vocabulary. As mentioned above, there are studies that suggest proficient LM students may show more growth compared to their EO peers during upper-elementary or middle school years (Lawrence et al., 2012; Uccelli et al., 2015). However, there is also evidence showing that a large proportion of reclassified students are still experiencing difficulties in English language arts and mathematics and a considerable number are retained one grade or more (Slama, 2014). It is worth exploring whether these academic challenges are associated with deficits in students' academic vocabulary knowledge, given its centrality to

skilled content area-reading (Townsend et al., 2012). Thus, the research questions that will be addressed in this study are the following:

1. What are the general vocabulary growth trajectories of middle school students in each language designation?

2. What are the academic vocabulary growth trajectories of middle school students in each language designation?

Methods

Sample

The data for this study came from control school students in an IES funded randomized efficacy trial of Word Generation (www.wordgeneration.org). Students from six middle schools in a large urban district in California participated the study. Initially, there were 3,653 students in this longitudinal study who contributed to at least one wave of data collection. Of these students, there were 474 who did not complete the baseline assessments and 96 students who were missing language proficiency status data that were essential in our analysis; Thus, our analytical sample included 3,161 students. In Year 2, 1,099 eighth-grade students left the study (graduated middle school).

Table 2.1 presents demographic information of the participating students by language designation (first four rows) and for the whole analytic sample (last row). Across the sample 84% of students were eligible for free and reduced lunch, but there were differences across groups. Only 71% of EO students were eligible, while 92% of LEP students were eligible. A very small number of RFEP (2%) and IFEP (4%) students had individualized education plans, and there was a large number of students in the gifted and talented education program (40% across all students).

Table 2.1

Demographic Information of the Participants

	Eligible for FRL	Special Education	Gifted and Talented Education	Race				
				Asian	Hispanic	African-American	White	Other
EO (*n* = 948)	71%	17%	35%	33%	11%	26%	16%	14%
IFEP (*n* = 322)	80%	4%	51%	61%	22%	2%	5%	10%
RFEP (*n* = 1,235)	90%	2%	58%	77%	17%	0%	2%	4%
LEP (*n* = 659)	92%	20%	9%	52%	39%	1%	2%	7%
Total (*N* = 3,161)	84%	10%	40%	57%	20%	8%	6%	8%

Note. FRL = Free and reduced lunch, EO = English-only, IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient.

Table 2.2

Language Minority Students' Home Languages

	Cantonese	Spanish	Filipino	Vietnamese	Mandarin	Other
IFEP	35%	18%	3%	2%	1%	41%
RFEP	47%	15%	3%	3%	2%	30%
ELL	25%	32%	7%	3%	2%	31%
Total LM Students	40%	20%	4%	3%	2%	31%

Note. IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient, LM = Language minority.

The analytical sample consists of Asian (57%), Hispanic (20%), White (6%), and African-American (8%) students. Most EO students were Asian (33%) or African-American (26%). Students in the other language designations were mostly Asian and Hispanic. Table 2.2 displays home language information of the participating LM students. The majority of the LM students in our sample were from Cantonese-speaking homes. LM students from Spanish-speaking homes were the second largest group in our sample.

Procedure

To assess students' general and academic vocabulary, we administered a standardized vocabulary test and a researcher-developed academic vocabulary test. All assessments were administered four times across two consecutive academic years in a group setting: once at the beginning of the school year (September/October) and once at the end of the school year (May).

Measures

Time. TIME is a level-1 variable indicating the time since the start of the study at which students completed the assessments. The data were collected in the fall and spring of two consecutive years. We coded each wave in months (i.e., wave 1 = 0 month, wave 2 = 7 months, wave 3 = 12 months, wave 4 = 19 months).

Language status. The collaborating school district provided detailed information about participating language-minority students. Dummy variables for EO, IFEP, RFEP, and LEP students were used as student-level (level-2) predictors in our analyses. Although some of the LEP students' status changed during this course of the study (i.e., 229 LEP students in Year 1 were reclassified as RFEP in Year 2), we treated this variable as time-invariant, reflecting classification of students' English proficiency status at the beginning of the study (i.e.,

September Year 1). The analytical sample consists of 30% EO, 10% IFEP, 39% RFEP, and 21% LEP students.

General vocabulary. The Gates-MacGinitie Reading Test (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) is a group-administered assessment that includes reading and vocabulary subtests. As the vocabulary subtest assesses "words of general usefulness," (MacGinitie et al., 2000, p. 9) the score from this test was used as an indicator of students' general vocabulary knowledge. The level 6 Form T was administered to sixth grade students and the level 7/9 Form T was given to seventh and eighth grade students. This test was administered twice each year, in September/October and in May. The vocabulary subtest consists of 45 multiple-choice items that ask students to choose the synonym of a target word. Extended scale scores of the vocabulary subtest (GV) were used in the analysis as they allow for estimating growth over time on a single scale (MacGinitie et al., 2000). The internal reliability (Cronbach's α) for our analytical sample was .90 at the first wave. The mean and the standard deviation of the general vocabulary were 518.82 and 36.92, respectively, at the first wave.

Academic vocabulary. Students' academic vocabulary was assessed with a 50-item multiple-choice test developed by the Word Generation research team (See Appendix A for the academic vocabulary test forms that were used in this study). For each item, the target word was embedded in a short sentence, and students were asked to choose a synonym for the target word from among four choices. The target words were taken from the Academic Word List (Coxhead, 2000). The same academic vocabulary test was administered twice, in September/October and in May. A different test form was used in the first (waves 1 and 2) and second (wave 3 and 4) year. Out of 50 items, 20 were anchor items (items that appear in both test forms), and 30 were unique to each test form.

To make responses from the two forms of the assessment equivalent, we conducted test scaling using item response theory (IRT) analysis. Scoring with IRT models allows students' academic vocabulary scores from each wave to be estimated on a common metric. We ran a unidimensional three-parameter IRT model on the first and third wave of data constraining the anchor items to have same item parameters (i.e., slope and intercept). In this process, we dropped seven anchor items, two unique items from the Year 1 test form, and one unique item from the Year 2 test form based on the overall fit index (i.e., root mean square error of approximation, RMSEA; Brown & Cudeck, 1993) and local dependence statistics. The final model fit the data well (RMSEA = .03). Marginal reliability for the first wave was .91 and that of the third wave was .92. Once item parameters were obtained, we used them to score all the waves of our data. The scoring method we used was *expected a posteriori* (EAP). These scores were used as an indicator for academic vocabulary (ACV) in our analysis. The scaled scores had a mean of -.02 and standard deviation of .89 at the first wave for our sample.

Reading comprehension. The Gates-MacGinitie Reading Test (MacGinitie et al., 2000) measures students' reading comprehension skills. The level 6 Form T was administered to sixth grade students and the level 7/9 Form T was given to seventh and eighth grade students. Students were asked to read a passage and answer relevant comprehension questions (48 items total). Extended scale scores from the reading subtest (RC) obtained at the first wave were used as a covariate in our analysis to address both of our research questions. The Cronbach's α for our analytical sample was .91 in the first wave. The mean and standard deviation of our sample on reading comprehension was 524.40 and 37.79, respectively, at the first wave.

Grade-level cohort. To control for different grade levels in the analyses, student-level dummy variables were created for grade six (GRADE_6), seven (GRADE_7), and eight

(GRADE_8). There was a comparable representation of sixth (32%), seventh (35%), and eighth (33%) grade students in our sample.

Summer. How many summers (SUMMER) students experienced since the start of the study was also included in our analysis (e.g., wave 1 = 0, wave 2 = 0, wave 3 = 1, wave 4 = 1 for sixth and seventh grade cohorts). SUMMER wass a time-varying continuous level-1 variable, and its parameter indicated whether or not students experience summer setback with their outcomes.

Socioeconomic status (SES). Eligibility for receiving free or reduced lunch (FRL) was used as an indicator of students' SES. A student-level dummy variable was created to indicate students who received free and reduced lunch (FRL = 1) and those who did not (FRL = 0).

Ethnicity. Five student-level dummy variables (ASIAN, HISPANIC, BLACK, WHITE, and OTHER) were created to control for students' ethnicity. Asian students were used as the reference group in our analysis.

Special education status. A student-level dummy variable was created to indicate students who were receiving special education (SPED = 1) and those who were not (SPED = 0).

Gifted and talented education status. A student-level dummy variable was created to indicate students who were receiving gifted and talented education (GATE = 1) and those who were not (GATE = 0).

Data Analysis

To answer our research questions, we conducted multilevel models for change (Singer & Willet, 2003). These models allowed us to use all waves of data from each student to create a model of vocabulary growth over the course of two years. The data was prepared in a person-

period dataset such that each student had up to four rows of data. The hypothesized multilevel

model for change for the first research question was:

Level-1 (outcomes in four waves across two years):

$$\widehat{GV} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}TIME^2_{ij} + \pi_{3i}SUMMER_{ij} + \varepsilon_{ij} \tag{1}$$

Level-2 (student-level):

$$\pi_{0i} = \gamma_{00} + \gamma_{01}ACV_i + \gamma_{02}RC_i + \gamma_{03}IFEP + \gamma_{04}RFEP_i + \gamma_{05}LEP_i + \gamma_{06}GRADE7_i +$$

$$\gamma_{07}GRADE8_i + \gamma_{08}SES_i + \gamma_{09}HISPANIC_i + \gamma_{10}BLACK_i + \gamma_{11}WHITE_i +$$

$$\gamma_{12}OTHER_i + \gamma_{13}SPED_i + \gamma_{13}GATE_i + \zeta_{0i} \tag{2}$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11}IFEP_i + \gamma_{12}RFEP_i + \gamma_{13}LEP_i + \zeta_{1i} \tag{3}$$

$$\pi_{2i} = \gamma_{20} + \gamma_{21}IFEP_i + \gamma_{22}RFEP_i + \gamma_{23}LEP_i \tag{4}$$

$$\pi_{3i} = \gamma_{30} + \gamma_{31}IFEP_i + \gamma_{32}RFEP_i + \gamma_{33}LEP_i \tag{5}$$

where $\varepsilon_{ij} \sim N(0, \sigma_1{}^2)$, and $\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right).$

The coefficient γ_{00} represents the average score for EO students at the first wave (the

first measurement point); γ_{10} represents the average initial slope for EO students; γ_{20} represents

the average true acceleration for EO students; and γ_{30} represents the average summer setback (or

gain) for EO students. The random effect ε_{ij} is a Level 1 residual for student i at time j and is

assumed to be drawn from a normal distribution with mean of 0 and variance $\sigma_1{}^2$. Random

effects ζ_{0i} and ζ_{1i} represent Level 2 residuals for the intercept and slope, respectively. They are

both hypothesized to be drawn from a multivariate normal distribution with a mean of zero,

unknown variances σ_0^2 and σ_1^2, and unknown covariance σ_{01}. For each research question, the

parameter estimates that are specific to each language proficiency group (i.e., IFEPs: γ_{03}, γ_{11},

γ_{21}, and γ_{31}; RFEPs: γ_{04}, γ_{12}, γ_{22}, and γ_{32}; LEPs: γ_{05}, γ_{13}, γ_{23}, and, γ_{33}) were compared to the

reference group parameters (i.e., EO: γ_{00}, γ_{10}, γ_{20}, and γ_{30}). Parameters γ_{03}, γ_{04}, and γ_{05} refer to baseline differences in academic vocabulary test scores across language groups with EO as the reference group. Specifically, γ_{03} indicates the difference between EO and IFEP students, γ_{04} the difference between EO and RFEP students, and γ_{05} the difference between EO and LEP students. Similarly, the parameters γ_{11}, γ_{12}, and γ_{13} refer to differences in rate of growth across the language groups; the parameters γ_{21}, γ_{22}, and γ_{23} indicate the differences in the acceleration of growth across the language groups, and the parameters γ_{31}, γ_{32}, and γ_{33} note the differences in the summer setback across language groups. Students' ACV scores from four waves were used as dependent variables to answer our first research question, and the scores from the first wave (ACV_W1) were used as a covariate in the analytical model to address our second research question. The same hypothesized model with general vocabulary (GV) as an outcome variable and academic vocabulary (ACV_W1) as a covariate was used to answer our second research question. The inclusion of the quadratic term improved the model fit in both outcomes (general vocabulary: Δ-2LL = 23.26, $df = 1$, $p < .001$; academic vocabulary: Δ-2LL = 55.15, $df = 1$, $p < .001$); its negative value indicates that learning rates decrease as students get older.

Results

Preliminary Descriptive Analyses

Table 2.3 displays means and standard deviations of academic and general vocabulary for students of different English proficiency status and grade levels across the four waves. Unsurprisingly, students in older grades performed better than students in lower grades on each measure across language proficiency designations. Comparing performance in the fall of Year 1 and spring of Year 1 reveals that students improved during the school year on both measures in

every grade and language category. Trends across summer were much less consistent: some

groups improved during the summer while the average scores of other groups dropped. The

correlation between the two measures was .83 in wave 1.

Table 2.3
Means and Standard Deviations of Academic Vocabulary and General Vocabulary Scores by
Students' Language Designations and Grade Levels

		General Vocabulary				Academic Vocabulary			
		Fall Year 1	Spring Year 1	Fall Year 2	Spring Year 2	Fall Year 1	Spring Year 1	Fall Year 2	Spring Year 2
6th Grade Cohort	EO	518.08	537.33	529.34	540.97	-0.15	0.29	0.27	0.37
		(39.56)	(40.90)	(39.89)	(44.21)	(0.93)	(0.93)	(0.95)	(1.05)
	IFEP	523.03	543.44	536.71	546.13	0.10	0.43	0.60	0.70
		(29.46)	(29.47)	(28.00)	(29.87)	(0.67)	(0.65)	(0.54)	(0.64)
	RFEP	516.92	534.87	529.95	540.25	-0.10	0.29	0.40	0.61
		(28.09)	(30.35)	(25.24)	(28.32)	(0.61)	(0.66)	(0.62)	(0.67)
	LEP	480.87	493.57	494.51	500.38	-0.95	-0.60	-0.52	-0.42
		(23.82)	(25.61)	(28.17)	(29.97)	(0.64)	(0.59)	(0.72)	(0.79)
7th Grade Cohort	EO	525.12	538.91	541.00	550.48	0.07	0.31	0.40	0.58
		(38.18)	(37.79)	(37.09)	(39.74)	(0.95)	(1.04)	(0.92)	(1.00)
	IFEP	537.55	544.37	552.92	560.98	0.35	0.64	0.72	0.91
		(29.23)	(29.86)	(35.30)	(36.85)	(0.78)	(0.78)	(0.75)	(0.78)
	RFEP	526.97	537.06	541.22	552.09	0.20	0.50	0.56	0.79
		(23.30)	(23.82)	(23.44)	(27.02)	(0.62)	(0.62)	(0.63)	(0.63)
	LEP	485.64	495.59	500.58	509.29	-0.86	-0.72	-0.54	-0.30
		(27.73)	(29.46)	(27.95)	(30.18)	(0.63)	(0.73)	(0.72)	(0.80)
8th Grade Cohort	EO	534.91	547.36	-	-	0.37	0.60	-	-
		(40.18)	(40.20)	-	-	(0.94)	(1.00)	-	-
	IFEP	545.16	558.74	-	-	0.60	1.01	-	-
		(31.63)	(31.49)	-	-	(0.80)	(0.76)	-	-
	RFEP	538.62	550.32	-	-	0.54	0.83	-	-
		(25.62)	(26.90)	-	-	(0.64)	(0.69)	-	-
	LEP	490.67	502.33	-	-	-0.78	-0.53	-	-
		(29.37)	(31.00)	-	-	(0.68)	(0.75)	-	-

Note: Standard deviations in parentheses. EO = English-only, IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient. Academic Vocabulary scores are scaled scores from two different forms. General Vocabulary scores are the extended scaled scores from Gates-MacGinitie Reading Test. Grade cohorts indicate student's grade levels in the beginning of the study. The descriptives for academic vocabulary reflect the means and standard deviations of the test scores of students who had both general vocabulary and reading comprehension scores at the first wave and other demographic data. Similarly the descriptive statistics for general vocabulary reflect the means and standard deviations of the test scores of students who had both academic vocabulary and reading comprehension scores at the first wave and other demographic data.

General vocabulary. IFEP students outperformed all other students across all waves in general vocabulary. EO and RFEP students had similar average scores across four waves. Average LEP performance on the general vocabulary knowledge measure was lower than their peers. The rank ordering of average performance across language designations was consistent in all grade level cohorts. All groups grew in their general vocabulary across two years: however, some groups (e.g., EO, IFEP, and RFEP students in the sixth grade cohort) experienced a drop in their general vocabulary scores after their summer break. Although students in the seventh grade cohort did not show such pronounced setback, their vocabulary growth during the summer was smaller than it was during the academic year.

Academic vocabulary. Across all grade levels, IFEP students outperformed their peers, including EOs, in academic vocabulary across all waves. Given the rigorous redesignation processes used in California, RFEP students had the second highest scores followed by the EO and LEP students. All groups in our sample improved in average academic vocabulary knowledge from wave 1 to wave 4. Among them, IFEPs and RFEPs in the sixth and seventh grade cohort showed the most noticeable improvements. Across groups, summer academic vocabulary growth was slow compared to the pace during the school year.

Research Question 1. What are the general vocabulary trajectories of students in each language designation?

Model 1 in Table 2.4 shows students' growth in general vocabulary over two years.

Table 2.4

Multilevel Models for Change Predicting General Vocabulary and Academic Vocabulary Scores

		General Vocabulary	Academic Vocabulary
		Model A	Model B
Fixed Effects	Intercept	377.32***	-10.08***
		(7.40)	(0.15)
	IFEP	-0.63	-0.01
		(1.30)	(0.03)
	RFEP	-5.36***	-0.06*
		(0.97)	(0.02)
	LEP	-9.86***	-0.21***
		(1.15)	(0.03)
	TIME	1.92***	0.03***
		(0.10)	(0.002)
	TIMEXIFEP	-0.05	0.01**
		(0.16)	(0.004)
	TIMEXRFEP	0.05	0.02***
		(0.11)	(0.003)
	TIMEXLEP	-0.26*	0.01*
		(0.13)	(0.003)
	TIME2	-0.02***	-0.001***
		(.004)	(0.0001)
	SUMMER	-9.42***	-0.01
		(1.16)	(0.03)
	SUMMEXIFEP	2.38	-0.07
		(2.25)	(0.05)
	SUMMERXRFEP	1.70	-0.09*
		(1.50)	(0.04)
	SUMMERXLEP	5.79**	0.02
		(1.74)	(0.04)
	GV		0.01***
			(0.0003)
	RC	0.29***	0.01***
		(0.01)	(0.0003)
	AV	20.86***	
		(0.58)	
	GRADE_7	-2.57**	-0.10*
		(0.74)	(0.02)
	GRADE_8	-1.73*	0.02
		(0.82)	(0.02)
	WHITE	4.00**	0.07*
		(1.30)	(0.03)
	BLACK	-4.07**	-0.10**
		(1.29)	(0.03)
	HISPANIC	-2.32**	-0.01
		(0.78)	(0.02)
	OTHER	0.09	-0.004
		(1.13)	(0.03)
	SES	-2.49**	-0.04*
		(0.84)	(0.03)
	SPED	-3.47**	-0.17***
		(1.06)	(0.02)
	GATE	3.38***	0.16***
		(0.74)	(0.02)
Level 1 Variance Component	Residual	194.54***	0.11***
		(4.36)	(0.002)
Level 2 Variance Component	Intercept	187.61***	0.09***
		(9.39)	(0.004)
	TIME	0.53***	0.0003***
		(.06)	(0.00003)
	Covariance	-2.63***	-0.0004
		(.61)	(0.0003)
	N (Students)	3,037	3,141
	N (Observations)	9,162	9,066

There is no difference between the baseline scores of IFEP and EO students (IFEP β = -.63, p = .63). However, the coefficients for RFEP and LEP students were both negative: in controlled models, RFEP (β = -5.36, $p < .001$) and LEP (β = -9.86, $p < .001$) students had lower baseline scores than EOs on average. EO students' general vocabulary scores improved by an average of about 2 points per month (TIME β = 1.92, $p < .001$). Average IFEP and RFEP learning rates did not differ from those of EOs controlling for covariates (i.e., there were no TIME by language status interactions for these two groups). However, the interaction of TIME and LEP was negative and statistically significant (β = -.26, p = .04), meaning LEP students' average rate of growth was lower than that of EO students. The negative and significant TIME^2 coefficient (β = -.02, $p < .001$) indicates that students' rate of growth decreased over time on average. Language group by TIME^2 interaction terms were tested but did not improve the model. EO students did experience change in their general vocabulary learning trajectories during the summer (SUMMER; β = -9.42, $p < .001$), as did IFEP and RFEP students. LEP students also experienced reduced summer learning, but the setback was not as strong (SUMMER X LFEP β = 5.70, $p < .001$). These results are plotted in Figure 2.1.

We saw in the descriptive table that IFEP had the highest average baseline general vocabulary scores. However, the parameter estimate associated with baseline IFEP scores was not significant in our growth model. The apparent difference between the averages in Table 2.4 and plots in Figure 2.1 is most likely explained by the fact that we have good baseline controls in

our model, and IFEP students scored higher than their English-only peers on the reading

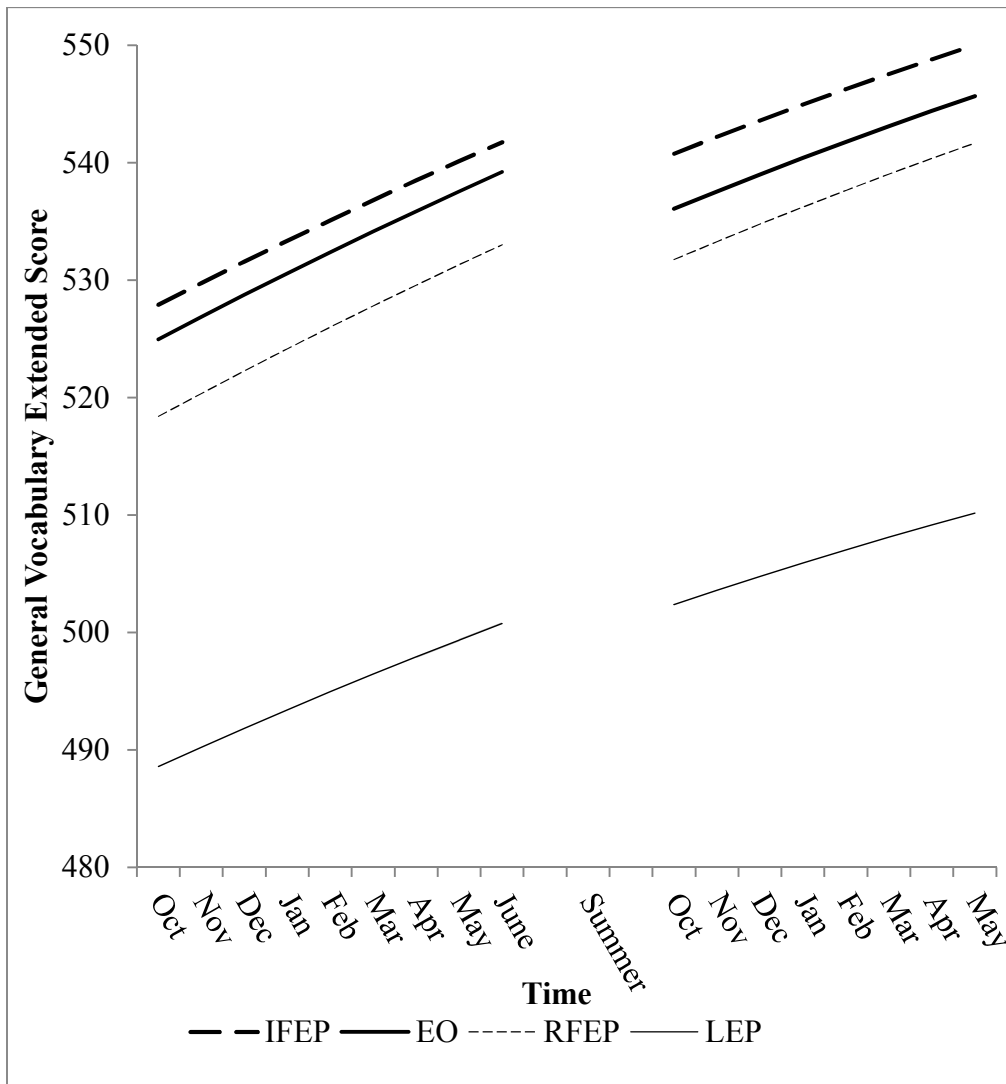comprehension and academic vocabulary covariates.

Figure 2.1. Predicted general vocabulary growth trajectories of sixth-grade students of each language designation.

Our prototypical plots demonstrate the predicted growth trajectories of sixth-grade

students of each language designation based on the mean scores on the covariates for that group

(Figure 2.1)[3]. They demonstrated that IFEP students (heavy dashed line) began the study with the highest general academic scores[4], followed by EO (heavy solid line), RFEP students (light dashed line), and LEP students (light solid line). Although we plotted the parameterized differences in summer setback from Table 2.4, these differences were not significant. In essence, the growth of the highest performing groups was the same across the course of the study although there were significant baseline differences. LEP students had flatter rates of growth during the school year and a less pronounced summer setback than students in the other groups.

Research Question 2. What are the academic vocabulary trajectories of students in each language designation?

Model 2 in Table 2.4 presents results for EO and LM students' growth in academic vocabulary over two years. The coefficient associated with IFEP was not significant ($\beta = -.01$, $p = .74$), which indicates that when other variables are controlled, there is no baseline difference between EO and IFEP students. The coefficients associated with RFEP and LEP baseline scores were negative. On average, RFEP ($\beta = -.06$, $p = .01$) and LEP ($\beta = -.21$, $p < .001$) students had lower scores than their EO peers at baseline controlling for demographic and other assessments.

On average, EO students showed growth in academic vocabulary ($\beta = .03$, $p < .001$; .03 points per month). The interaction terms of TIME and language designation (TIME X IFEP, TIME X RFEP, and TIME X LEP) were all positive ($\beta = .01$, $p = .002$, $\beta = .02$, $p < .001$, and $\beta =$

---

[3] This plot is based on the average covariates and parameter estimates of white sixth-grade students who is not eligible for free and reduced lunch, not on an individualized education plan, and not in the gifted and talented program.

[4] The results of this multilevel model for change indicate that on average sixth grade EO students scored 524.96 at baseline. This number was calculated by summing the following variables: -377.32 (constant) + 20.86 (coefficient for academic vocabulary) X .004 (mean score of academic vocabulary for EO students at the first wave) + .29 (coefficient for reading comprehension) X 517.22 (mean score of reading comprehension for EO students at the first wave).

.01, $p = .04$, respectively), indicating IFEPs, RFEPs, and LEPs had steeper growth than EOs. The negative and significant TIME^2 coefficient ($\beta = -.001$, $p < .001$) indicates that students' rates of growth decreased over time on average. Language group by TIME^2 interaction terms were not significant and so were dropped. EO students did not experience summer setback (SUMMER; $\beta = -.01$, $p = .86$) in academic vocabulary. The interaction terms of SUMMER and two language groups (SUMMER X IFEP and SUMMER X LEP) were not significant ($\beta = -.07$, $p = .22$ and $\beta = .02$, $p = .57$, respectively), indicating neither IFEPs nor LEPs experienced a change in their learning trajectories during the summer. However, the coefficient for SUMMER X RFEP ($\beta = -.09$, $p = .01$) was negative and significant; RFEPs' academic vocabulary learning rates decreased during the summer. These results are illustrated in Figure 2.2.

Figure 2.2 shows prototypical growth trajectories for sixth grade EO, IFEP, RFEP, and LEP students who do not have individualized education plans. At the first wave of data collection, trends were similar to general vocabulary trends: IFEP students started the study with the highest average scores, followed by EO, RFEP, and LEP students[5]. We note that in our descriptive table, EO students' baseline scores were lower than RFEP students' scores. This is largely explained by the fact that the descriptives do not control for differences associated with being on an individualized education plan; a high percentage of EO students had individualized education plans, producing more divergent predicted values for the EO prototypical plots than

---

[5] The results of this multilevel model for change indicate that on average EO students scored .03 at baseline. This number was computed by summing the following variables: -10.08 (constant) + .01 (coefficient for general vocabulary) X 523.22 (mean score of general vocabulary for EO students at the first wave) + .01 (coefficient for reading comprehension) X 515.84 (mean score of reading comprehension for EO students at the first wave). When IFEP students' mean scores of general vocabulary and reading comprehension were taken into account, the constant for IFEP students was .07. The baseline scores for RFEP and LEP students were calculated to be -.06 and -.88, respectively.

for the other groups. The most obvious trend in Figure 2.2 is the improvement of IFEP, RFEP, and LEP students relative to EO students across the four waves of data. It is also worth noting that RFEP students lost ground to EO students during the summer, although they more than made up for that during the school year[6].
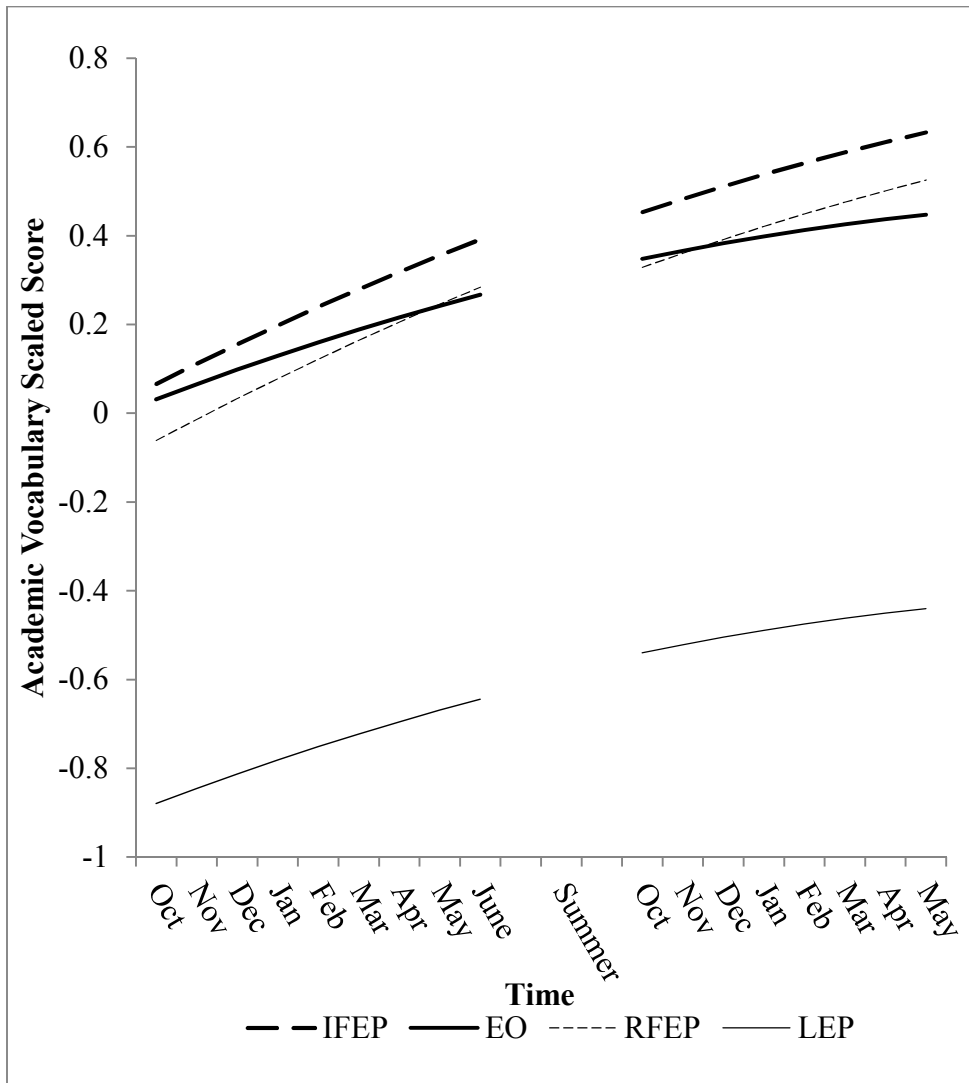

Figure 2.2. Predicted academic vocabulary growth trajectories of sixth-grade students of each language Designation

[6] We tested our analytical models with school dummy variables, but inclusion of these variables did not change our results. We also conducted our analysis with special education by time interaction term. However, this interaction was not significant.

Discussion

The goal of this study was to model the vocabulary growth trajectories of LM and EO students using longitudinal methods. Even if we believe that vocabulary knowledge is unidimensional across kinds of words[7], it is still possible that learning trajectories will differ as a function of student language status and word usage. We probed this hypothesis by modeling growth on two distinct vocabulary measures. While our general vocabulary measure included a wide range of English words, the words that were tested in the academic vocabulary assessment tended to be mid-frequency words that are often encountered in texts across content areas. Thus, not only did we investigate individual differences (i.e., different language designations) among adolescent students in their vocabulary knowledge growth, but we also examined differences in the types of words that students learned.

We analyzed the academic and general vocabulary growth trajectories of EO, IFEP, RFEP, and LEP students across middle grades. Students in the study attended schools in a large urban school district in California, a state with rigorous, multiple criteria for redesignating LEP students. At baseline, the rank order of student performance was the same on both vocabulary measures. IFEP students performed slightly higher than EO students, who outperformed RFEP and LEP students. However, there were interesting differences in the trajectories of students across the study, which we interpret with reference to both contextual and linguistic factors.

IFEP students: High achievers

The results from the multilevel models for change indicate that the IFEPs in our sample outperformed their EO and LM peers in academic and general vocabulary measures. Their rate

---

[7] Our confirmatory factor analysis with items from two measures indicated that the one-factor model (AIC = 661804, BIC = 663751, RMSEA = 0.03) fit the data better than the two-factor model (AIC = 664192, BIC = 666112, RMSEA = 0.032).

of growth and the amount of summer setback in general vocabulary were similar to those of their EO peers, meaning they were the best performing group on measures of general vocabulary at all time points. In fact, IFEP students showed a steeper rate of academic vocabulary growth than their EO peers. Because the magnitude of summer setback was similar to that of their EO peers, the gap between EO and IFEP students in academic vocabulary was wider at the end of the study than it was in the first wave.

We cannot fully explain the reason why IFEP students in our sample were outperforming their peers, however, it was not something that was unexpected. Previous research has shown that bilingual students can have a *bilingual advantage* in various domains (e.g., Adesope, Lavin, Thompson, & Ungerleider, 2010; Bialystok, 2002, 2011). Cummins' (1979) Threshold Hypothesis might explain why the IFEPs are high-achievers. This hypothesis suggests that when a bilingual individual attains a certain threshold level of proficiency and competency in his/her second language (English in this study context), this individual can enjoy cognitive advantages associated with bilingualism. As IFEP students are, by definition, those who were already fluent English speakers by the time they entered school, they may be fully enjoying their cognitive advantage associated with knowing a second language. This could also explain why a high percentage of IFEP students in our sample received gifted and talented education (Table 2.1). On the other hand, kindergarten bilingual proficiency could be a marker of general verbal intelligence, high SES, and high parental investment in education; indeed it is probably a marker of all three. Thus, while these results are consistent with the bilingual advantage hypothesis, they in no way prove it.

EO students: In need of explicit instruction in academic vocabulary

On average, EO students showed growth in both general and academic vocabulary. In

addition, they experienced a summer setback in their general vocabulary. However, this pattern was not found with their academic vocabulary knowledge. Figures 1 and 2 indicate that EO students' scores in both general and academic vocabulary were lower than their IFEP peers. While their general vocabulary scores remained higher than their RFEP peers, the academic vocabulary scores were predicted to be caught up by the RFEP peers in the second year.

It is interesting to note that EO, IFEP, and RFEP students showed similar growth in their general vocabulary knowledge, but IFEP and RFEP had steeper academic vocabulary growth trajectories compared to the EO students. The technical manual for the Gates-MacGinitie Reading Test (our general vocabulary measure) indicates that the words in this test were "words of general usefulness, not obscure or specialized words" (MacGinitie et al., 2000, p. 9) Words used in the test include both high-frequency words and very low-frequency words. One index of word frequency is the number of occurrences that a word has in the *Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995). We found that the average normed frequency of the general vocabulary target words was $M = 43.81$, $SD = 8.01$[8] (range: $22.1 - 70.3$). In contrast, the average word frequency of target words on the academic vocabulary assessment was $M = 48.42$, $SD = 5.78$ (range: $20.8 - 58.2$). The Gates-MacGinitie included words that students are likely to encounter and learn independently, words that are targeted for instruction in school, and also words that are extremely rare and likely only to be encountered from wide independent reading. In that sense, it is not surprising that growth on this measure was stable across groups: the test is intended to sample a wide range of words, and the broader the domain, the more

---

8 We used standard frequency index (SFI) which represents "a logarithmic transformation of weighted frequency of a word that was predicted to appear per million words" (Zeno et al., 1995, p. 12).

difficult it is for individuals or groups to experiences dramatic sustained improvements (Paris, 2005).

It *is* surprising that the IFEP students and RFEP students improved on our academic measure more than the EO students given how stable the growth across groups on the general vocabulary measure was. We were careful that these results were driven by difference in performance, not language status. We used propensity matching to find comparable EO and LM students based on their baseline reading, general vocabulary, and academic vocabulary scores, as well as their SES and GATE status. Our analyses predicting later vocabulary levels are consistent with those presented here: EO students did not make as large gains in their academic vocabulary as their matched LM students.

Given the same average improvements on a measure sampling across word types, how did IFEP and RFEP students manage to pull ahead in academic vocabulary? We might expect a global bilingual advantage to result in improvement across both measures equally. However, one of the central lines of research on the bilingual advantage is related to executive control, especially inhibitory control processes. If these data were evidence of a bilingual advantage, it might be related to students' well-honed ability to determine what are the important and strategic words to focus on and which ones to ignore. Of course, it is also possible that these IFEP and RFEP students have been instructed to be more strategic in word learning by teachers who emphasize the importance of strategy instruction for second language learners (Oxford, 1994). In either case, LM students may develop greater word awareness in English compared to EO students. When taught academic vocabulary in school, LM students may apply these skills to the new set of challenging words. In contrast, EO students may build a great deal of their general vocabulary implicitly through rich exposure. This could mean that they lack the skills or

strategies to learn academic vocabulary as efficiently as their LM peers. In either case, low performing EO students would be good candidates for explicit instruction to word learning strategies and rich exposure to general academic words.

RFEP students: Converging growth trajectory?

In this study, we did not find group differences in rates of general vocabulary learning. EO, IFEP, and RFEP students all learned words at a similar rate, and the estimated setback during the summer was also similar. However, despite the familiar rank ordering of students of groups at baseline, RFEP students had steeper learning trajectories of academic words than EO students. By the end of the study, RFEP students were outperforming the EO students on the academic vocabulary measure. The post-hoc t-test indicated that the difference in test scores between RFEP and EO students at the end of the study was statistically significant, $t(2181) = 2.16$, $p = .03$. Again, we offer a couple of complementary interpretations. The first is that redesignated students have reached a proficiency threshold in their second language, and that they benefit from a bilingual advantage. Ardasheva, Tretter, and Kinny (2012) used very similar method in their study with redesignated students and found that former LEP students (i.e., RFEPs) outperform their EO and LEP peers in both reading and mathematics. However, such bilingual advantage is not found in all studies, and although the criteria for such bilingual advantage is unclear, there is an easily understood sampling problem in comparing students who achieve high proficiency in two languages with a general sample of monolinguals. Nonetheless, these data, including our own, might be interpreted as supporting a weak version of the bilingual advantage hypothesis since it was found in academic vocabulary but not in general vocabulary. Another interpretation is that instruction for these students was more explicit or strategic at some point. Redesignated students may be exposed to more English academic language than LEP

students, and might have a more strategic word learning capacity than EO students, especially for high-leverage words.

LEP students: More support

Unsurprisingly, the LEP students were the lowest performing group in our study. However, our results indicated that LEP students improved more slowly compared to their EO peers in general vocabulary knowledge during the school year, but continued to learn during the summer. The general vocabulary measure targets some high-frequency words which LEP students might learn independently during the summer, even though they were not able to learn some of the low-frequency words as rapidly as their more proficient peers during the school year. In addition, LEP students showed steeper growth and no pronounced summer setback in their academic vocabulary compared to the EO students. These results suggest that LEP students may get some exposure to the general academic words in the summer. Unlike the test of general vocabulary, the academic vocabulary test included no words so infrequent that they would not be encountered in independent summer reading, in an enrichment class, or at a library book club.

On one hand, it is encouraging to see that LEP students are learning both the general vocabulary and demanding, academic vocabulary in their middle school years. On the other hand, despite their growth and relative small amount of summer setback, they were indeed the lowest performing students in our sample. LEP students need explicit instruction and attention in promoting their general and academic vocabulary knowledge.

Other factors influencing students' vocabulary growth trajectories

The students in our sample were from diverse linguistic and ethnic backgrounds as indicated in Tables 2.1 and 2.2. One thing to consider in interpreting the results is that the ethnic groups were not equally distributed across language groups. For instance, while 77% of the

RFEP students were Asians, approximately 30% of the EO students were Asians. Furthermore, while approximately one quarter of the EO student sample were African-American, there were not many African-American students in other language groups (0%-2%). Students' racial and ethnic backgrounds are associated with average educational achievement levels in the United States: White and Asian students tend to outperform African Americans and Hispanics (e.g., Farkas & Beron 2004; Fryer & Levitt, 2006; Kao & Thompson, 2003). We tried to control for students' ethnicity by including students' reported ethnicity in all of our analyses. However, statistical controls do not work well when the sample distribution is uneven. In addition, we need to consider demographic factors when interpreting studies in American urban schools which have larger number of LEP students but also larger number of African-American students, students from low-income homes, and students who struggle on standardized measure of reading (Dogan et al., 2011).

Table 2.5 shows means and standard deviations of general and academic vocabulary test scores of White and African-American EO students in our study. Unsurprisingly, there was a considerable gap between these two groups of students across two years. Because our main aim was to focus on how students' language status (i.e., whether EO or LM) is associated with their vocabulary growth, investigating the gap between White and African-American EO students was out of scope for this paper. However, we acknowledge that African-American students who enter school could also have some difficulty acquiring the language of schooling as their fellow LM students as their home language (i.e., African American Vernacular English) also differs from academic English (e.g., Charity, Scarborough, & Griffin, 2004; Terry, Connor, Petscher, & Conlin, 2012). Our EO sample includes a heterogeneous group of students, and treating them as one reference group is likely to have influenced our findings.

Table 2.5

Means and Standard Deviations of Vocabulary Test Scores of White and African-American English-Only Students

| | | White | | African-American | |
		M	SD	M	SD
General Vocabulary	Fall Year 1	553.46	37.63	504.49	34.23
	Spring Year 1	571.68	35.44	518.72	36.31
	Fall Year 2	564.82	31.89	513.43	33.08
	Spring Year 2	579.74	38.75	524.35	38.12
Academic Vocabulary	Fall Year 1	0.75	0.83	-0.44	0.89
	Spring Year 1	1.08	0.80	-0.16	0.91
	Fall Year 2	0.95	0.77	-0.22	0.85
	Spring Year 2	1.09	0.84	-0.18	1.05

Note. Students from all grade-level cohorts were included in this table.

The sample for this study comes from a large urban school district that serves a relatively diverse group of students: students from socioeconomically disadvantaged homes, students who are still acquiring English, and students with different ethnicities. This is a distinctive sample from one state in the United States, and the results that we found from this sample may not generalize to other school districts that serve different group of students. Thus, while these findings provide insight about how students with different language designations may perform over time, we need to be cautious in generalizing these findings across contexts and language groups.

Limitations and Future Directions

In interpreting the findings of this study, it is important to consider the demographics of our sample. The distribution of first languages in our sample did not match that in the larger United States population; 67% of the LM students in our sample were Asian and 20% Hispanic, whereas approximately 75% of the LM students in the United States were from Spanish-speaking homes (National Clearinghouse for English Language Acquisition, 2011). While this

may not be a serious limitation of the study, it is important to note that our sample was not representative of LM students nationwide. Additionally, language designation of the participants was solely based on the school district report. We were not involved in the identification and classification of LM students and had to trust the district's decision about students' language proficiency designation. However, the initial performance on the vocabulary measures seemed to be consistent with what we would expect from students in each language designation.

In regards to assessing students' vocabulary knowledge, we only used one measure for each of the outcomes (i.e., general and academic vocabulary). The two assessments that we used were all synonym tasks which asked the students to find the closest synonym of the target word that was embedded in a sentence (or a phrase) among four to five choices. Using multiple measures that employ different formats of items (e.g., cloze items) and tap into different dimensions of word knowledge may have produced more precise estimates of students' vocabulary knowledge. However, our ongoing research into vocabulary assessments using multiple measures suggests that the correlations between student performance across item types is very high, especially if latent scores are used. Despite these limitations, this paper improves our knowledge of students' general and academic vocabulary knowledge across middle grades and suggests differential mechanisms for word learning in the relationship between these measures and reading comprehension.

References

Ardasheva, Y., Tretter, T. R., & Kinny, M. (2012). English language learners and academic achievement: Revisiting the threshold hypothesis. *Language Learning, 62*(3), 769-812. doi:10.1111/j.1467-9922.2011.00652.x

August, D., Carlo, M., Dressler, C., & Snow, C. E. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50-57. doi:10.1111/j.1540-5826.2005.00120.x

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

Bailey, A. L., & Kelly, K. R. (2013). Home language survey practices in the initial identification of English learners in the United States. *Educational Policy, 27*(5), 770-804. doi:10.1177/0895904811432137

Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy, 54*(1), 4-12. doi:10.1598/JAAL.54.1.1

Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural equation models* (pp. 132-162). Newbury Park, CA: Sage.

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

Chall, J. S., & Jacobs, V. A. (2003). Poor children's fourth-grade slump. *American Educator, 27*(1), 14-17.

Charity, A. H., Scarborough, H. S., & Griffin, D. M. (2004). Familiarity with school English in

African American children and its relation to early reading achievement. *Child Development, 75*(5), 1340-1356.

Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly, 23*(3), 509-531. doi:10.2307/3586923

Common Core State Standards Initiative. (2010). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

Corson, D. (1997). The learning and use of academic English words. *Language Learning, 47*(4), 671-718. doi: 10.1111/0023-8333.00025

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

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

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education, 2nd Edition, Volume 2: Literacy* (pp. 71-83). New York, NY: Springer Science + Business Media LLC.

Dogan, E., de Mello, V. B., Lewis, S., Simon, C., Uzzell, R., Horwitz, A., & Casserly, M. (2011). *Addendum to pieces of the puzzle: Recent performance trends of urban districts-- A closer look at 2009 NAEP results*. Washington, DC: Council of the Great City Schools.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research, 33*(3), 464-497.

doi:10.1016/j.ssresearch.2003.08.001

Fryer, R. G., & Levitt, S. D. (2006). The Black-White test score gap through third grade.

American Law and Economics Review, 8(2), 249-281. doi:10.1093/aler/ahl003

Goldenberg, C. (2010). Reading instruction for English language learners. In M. Kamil, P. D.

Pearson, E. Moje & P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol. IV, pp.

684-710). Newark, DE: International Reading Association.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2),

127-160. doi:10.1007/BF00401799

Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language

children and its role in predicting word recognition ability. *Applied Psycholinguistics,*

30(1), 153. doi:10.1017/S0142716408090073

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement

and attainment. *Annual Review of Sociology, 29*, 417-442.

doi:10.1146/annurev.soc.29.010202.100019

Kim, J., & Herman, J. L. (2009). A three-state study of English learner progress. *Educational*

Assessment, 14(3-4), 212-231. doi:10.1080/10627190903422831

Lawrence, J. F. (2012). English vocabulary trajectories of students whose parents speak a

language other than English: Steep trajectories and sharp summer setback. *Reading and*

Writing, 25(5), 1113-1141. doi: 10.1007/s11145-011-9305-z

Lawrence, J. F., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. E. (2012). Language

proficiency, home-language status, and English vocabulary development: A longitudinal

follow-up of the Word Generation program. *Bilingualism: Language and Cognition,*

15(3), 437-451. doi: 10.1017/S1366728911000393

Leach, J., Scarborough, H., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95*(2), 211-224. doi:10.1037/0022-0663.95.2.211

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly, 45*(2), 196-228. doi:10.1598/RRQ.45.2.3

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing, 19*(1), 99-131. doi:10.1007/s11145-005-4713-6

Linquanti, R., & Cook, H. G. (2013). Toward a "Common definition of English learner": A brief defining policy and technical issues and opportunities for State Assessment Consortia. Washington, DC: Council of Chief State School Officers.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading test technical report: Forms S and T.* Chicago, IL: The Riverside Publishing Company.

Mancilla-Martinez, J., & Lesaux, N. K. (2011a). Early home language use and later vocabulary development. *Journal of Educational Psychology, 103*(3), 535-546. doi:10.1037/a0023655

Mancilla-Martinez, J., & Lesaux, N. K. (2011b). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*(5), 1544-1560. doi:10.1111/j.1467-8624.2011.01633.x

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

National Center for Education Statistics. (2012). *The nation's report card: Vocabulary results from 2009 and 2011 NAEP reading assessments (NCES 2013-452).* Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

National Clearinghouse for English Language Acquisition. (2011). *What langauges do English learners speak? NCELA Fact Sheet.* Washington, DC: Author.

Olsen, L. (2010). Reparable harm: Fulfilling the unkept promise of educational opportunity for California's long term English learners. Long Beach, CA: Californians Together.

Oxford, R. L. (1994). *Language Learning Strategies: An Update.* Washington DC: Eric Clearinghouse on Languages and Linguistics.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184-202. doi:10.1598/RRQ.40.2.3

Parrish, T., Perez, M., Merickel, A., & Linquanti, R. (2006). Effects of the implementation of Proposition 227 on the education of English learners, K-12: Findings from a five year evaluation (Final Report). Palo Alto and San Francisco, CA: American Institutes for Research and WestEd.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246-256. doi:10.1037/0022-0663.97.2.246

Ragan, A., & Lesaux, N. K. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives, 14*(20), 1-32. doi: 10.14507/epaa.v14n20.2006

Ryan, C. (2013). Language use in the United States: 2011. Washington, DC: US Census Bureau.

Rydland, V., Aukrust, V., & Fulland, H. (2012). How word decoding, vocabulary and prior topic knowledge predict reading comprehension. A study of language-minority students in Norwegian fifth grade classrooms. *Reading and Writing, 25*(2), 465-482. doi:10.1007/s11145-010-9279-2

Shin, H. B., & Ortman, J. (2011). *Language projections: 2010 to 2020.* Paper presented at the Federal Forecasters Conference. Washington, DC.

Short, D., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners: A report to Carnegie Corporation of New York.* Washington, DC: Alliance for Excellent Education.

Singer, J., & Willet, J. (2003). *Applied longitudinal data analysis: Modeling change and even occurrence*. New York, NY: Oxford University Press.

Slama, R. B. (2014). Investigating whether and when English learners are reclassified into mainstream classrooms in the United States: A discrete-time survival analysis. *American Educational Research Journal, 51*(2), 220-252. doi: 10.3102/0002831214528277

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). New York, NY: Cambridge University Press.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934. doi: 10.1037//0012-1649.38.6.934

Terry, N. P., Connor, C. M., Petscher, Y., & Conlin, C. R. (2012). Dialect variation and reading: Is change in nonmainstream American English use related to reading achievement in first

and second grades? *Journal of Speech, Language & Hearing Research, 55*(1), 55-69, doi:10.1044/1092-4388(2011/09-0257)

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

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross☐disciplinary academic☐language proficiency and its association with reading comprehension. *Reading Research Quarterly*. doi:10.1002/rrq.104

Williams, K. T. (2001). *Group reading assessment and diagnostic evaluation technical manual*. Circle Pines, MN: American Guidance Service.

Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery--Revised*. Itasca, IL: Riverside.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

CHAPTER 3

Study 2. Vocabulary and reading performances of

redesignated fluent English proficient students

Currently, there are approximately 11 million language minority (LM) students in the United States (Aud et al., 2011). Considering the large numbers of LM students, it is not surprising that there is a growing body of literature addressing how best to meet their academic needs. There have been research studies about their literacy development (August & Shanahan, 2006), academic achievement (Collier, 1989; Kieffer, 2008), assessment issues (Abedi, 2002; Bedore et al., 2012; Kieffer, Lesaux, Rivera, & Francis, 2009; Solórzano, 2008), pedagogical knowledge (Bunch, 2013; Goldenberg, 2008), instructional environment (Harklau, 1994), and high-quality instructional practices to support these students (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006). Despite the improvements in our understanding of how LM students learn and what we can do to teach them better, more research is still needed to understand LM students' learning across different phases of English skill development.

Traditionally, LM students have been treated as two homogeneous groups in the research literature—those classified as limited in English proficiency and those classified as fully English proficient either at school entry or after some exposure to English instruction. Only recently have scholars begun to consider the degree of variability within each of those groups in their English proficiency levels and the implications for their educational trajectories. In particular, one might expect to find heterogeneity within groups of students recently redesignated as fully English proficient, especially since the criteria for reclassification are somewhat ambiguous in some states and extremely variable across states (Ragan & Lesaux, 2006). Given that there are large numbers of LM students with varying levels of English proficiency, our limited understanding of

how these students perform in school settings is concerning. Further, there is a pressing need to understand how the time at which LM students have been redesignated contributes to their growth in reading skills across adolescence.

Classification of redesignated fluent English proficient students

For our purposes, LM students are any school-aged students in the United States who hear or speak a language other than English at home (August & Shanahan, 2006). Nationally this is a large group that varies on many dimensions, including home language, socioeconomic status (SES), and English proficiency. There is no national guideline for identifying and/or classifying LM students. Thus, there is a great variability in the assessments and classification criteria used across states (Abedi, 2008; Bailey & Kelly, 2013; Kim & Herman, 2009; Ragan & Lesaux, 2006). Most schools use a variation of the following process, although the specific assessments, proficiency criteria, survey instruments, and language use criteria that are used differ widely across states and districts. Students enrolling in a school for the first time take a home-language survey. Students who report hearing or speaking a language other than English at home are identified as LM students. These students are then given an English-proficiency screening assessment. If they meet some minimal criteria set by their state or school district, they are classified as initially fluent English proficient (IFEP) students. If they do not meet the criteria, they are classified as limited English proficient (LEP) students. According to federal mandates, LEPs receive additional support for English language development. LEPs are assessed annually until they meet a minimum proficiency criterion, whereupon they are *redesignated* as fluent English proficient (RFEP) students. RFEP students no longer receive English language development services in their schools. For instance, California law requires schools to use the following criteria to reclassify LEP students: 1) Results on an English-proficiency assessment, 2)

teacher evaluation, 3) parental opinion, and 4) comparison to English-proficient students in basic skills that demonstrate LEPs' ability to participate in academic curriculum (English Language Proficiency Assessment of 1999, 2014). Many California districts use the California English Language Development Test (CELDT) to satisfy the first requirement and California Standards Test (CST) for the fourth requirement[9].

Timing of redesignation is an important issue that has not received much attention until recently. Different costs and benefits are associated with reclassification depending on when LEPs get dismissed from English language development services. On the one hand, although more lenient criteria leading to earlier dismissal from English language development services may restrict RFEP students from receiving additional language services that they may need, it may provide LM students with greater access to the mainstream classroom and higher quality instruction. On the other hand, more rigorous criteria lead to later dismissal from English language support services, yielding a different set of potential costs and benefits. Providing additional English services may be more expensive for school districts and LEPs may have limited access to the general, mainstream curriculum. At the same time, later dismissal could give LEP students more time and opportunity to develop academic English skills through additional support (Robinson, 2011). School districts and policymakers need to walk the line to find the right balance of costs and benefits. As of now, it is difficult to draw definitive conclusions about whether early or late dismissal is better both for individual LEP students and school districts.

Redesignated fluent English proficient students' academic performance

---

[9] With the implementation of the Smarter Balanced Assessments (www.smarterbalanced.org), the CSTs are no longer administered and used for the redesignation process.

Bold conclusions about the English proficiency of redesignated students are difficult to draw. In some studies, RFEPs perform as well as or even better than norming groups of English-Only (EO) students. For example, in a cross-sectional study Hwang et al. (2015) showed that adolescent RFEP students' reading comprehension and academic vocabulary scores were comparable to those of EO students and improved with more time since redesignation. Similarly, Ardasheva and colleagues (2012) found that former LEPs (i.e., RFEPs) outperformed LEPs and EO students in both reading and mathematics. On the other hand, Slama (2014) found that a large proportion of the redesignated students experienced academic difficulty, with approximately one quarter of the sample retained at least once after reclassification. The difference between these two sets of findings might be explained by the stringency of the reclassification criteria in the states where the studies were conducted—California and Kentucky vs. Massachusetts.

California has a complex, conjunctive set of criteria for reclassification, whereas Massachusetts relies on only one test, focused on English language proficiency, with no requirement to meet general academic achievement milestones. Umansky and Reardon (2014), studying nine cohorts of Latino students in one California district, found that it took eight years for 50% of the LEPs to be redesignated, and that approximately 25% were never reclassified. In contrast, Slama's (2014) discrete-time survival analysis of over 5,000 LEP students in Massachusetts found 50% redesignation within three years after school entry. Rubio (2014) conducted an analysis very similar to Slama's on data from the New York City Public Schools, and estimated 4 years to 50% reclassification, perhaps because there is access to bilingual education in New York, which has been shown to slow progress toward reclassification but

generate better long-term outcomes (Umansky & Reardon, 2014). Both Rubio and Slama found faster pathways to redesignation for non-Latinos.

Kim and Herman (2009) compared the academic performance of four groups of students -- LEPs, recently redesignated former-LEPs, RFEPs who had been reclassified more than two academic years earlier, and EOs -- across three states. They found achievement differences in LEP and non-LEP students' reading, math, and science assessments across different grades in all three states. However, the magnitude and direction of differences were inconsistent. In one state, RFEP students outperformed current LEPs, but underperformed compared to EO students. In another state, RFEP students outperformed even the EO students. These results again suggest that the stringency of reclassification criteria influence how quickly students exit the LEP status, and therefore the magnitude and direction of achievement gaps among LEP, former LEP, recently reclassified LEP, and non-LEP students. In states like California, then, where the reclassification process requires several kinds of evidence that students are performing at a high level, it is not surprising that RFEPs may outscore EOs. Nonetheless, it is likely that RFEPs in California, and elsewhere, show considerable variability in their post-redesignation learning trajectories (Hwang et al., 2015; Kieffer, 2008).

Vocabulary knowledge and reading comprehension in middle school

All students need to have good amount of vocabulary knowledge in order to comprehend texts in secondary schools (e.g., Hoover & Gough, 1990; Scarborough, 2001). Vocabulary knowledge is a fundamental skill to have as post-primary students *read to learn* (Chall & Jacobs, 2003) new concepts and ideas across multiple content areas. Weak vocabulary knowledge leads to unskilled reading, which can be a critical obstacle in all students' learning in secondary schools as texts become more complicated. Furthermore, for students to perform well in school

settings, they need to master the specific register of schooling, the academic language (Scarcella, 2003; Schleppegrell, 2004). Academic vocabulary is a critical component of academic language and academic vocabulary can be classified into general and discipline-specific academic words (Beck, McKeown, & Kucan, 2002; Nagy & Townsend, 2012). General academic words are high-leverage words that appear across multiple subjects whereas discipline-specific words are closely tied to specific content areas (Beck et al., 2002; Corson, 1997; Nagy & Townsend, 2012). However, teachers may not recognize the need to teach general academic words due to their abstract and polysemous natures, and because they rarely represent key concepts central to the content areas taught (Snow, 2010).

Although LM students, especially those with limited English proficiency, develop basic reading skills (e.g., decoding, word recognition) at rates relatively comparable to EO students, they tend to lag behind on measures of English higher-order linguistic skills, such as vocabulary and syntactic knowledge, compared to the EOs (August, Carlo, Dressler, & Snow, 2005; August & Shanahan, 2006; Lesaux, Lipka, & Siegel, 2006; Mancilla-Martinez & Lesaux, 2011). However, proficient LM students, particularly those designated as IFEP, show performance that is comparable to or even better than their EO peers in these reading-related domains (Hwang et al., 2015; Kieffer, 2011; Lawrence, 2012). It seems likely that the timing of RFEP students' redesignation would relate to their growth in reading-related outcomes in middle school years, but no data on that question are yet available.

Adolescent LM students' growth trajectories in vocabulary and reading

There is a small yet growing literature on adolescent language minority students' growth trajectories in literacy-related outcomes (Hwang, Lawrence, & Snow, under review; Kieffer, 2008, 2011; Lawrence, 2012; Lawrence, Capotosto, Branum-Martin, White, & Snow, 2012;

Mancilla-Martinez, Kieffer, Biancarosa, Christodoulou, & Snow, 2011). One line of results from these studies suggest that EO and LM students' baseline scores in reading predict their later reading scores (Kieffer, 2008, 2011; Mancilla-Martinez et al., 2011; Nakamoto, Lindsey, & Manis, 2007). In other words, students who are below average in earlier grades would continue to be below average in later grades, and those who were above average continue to be above average during their school years. The rate of growth for students at different ability levels does not differ throughout their school year, and their growth trajectories are curvilinear where the rate of growth slowly decreases as students get older. Such findings suggest that it is difficult for students who enter school with below average reading abilities to catch up to their average or above average peers.

Another line of findings shows heterogeneity in the trajectories shown by EO and LM students in vocabulary and reading growth based on their language status. For example, Lawrence (2012) found that proficient adolescent LM students show steeper vocabulary growth during the academic year compared to their EO peers. However, these LM students also experienced more pronounced summer setback (Alexander, Entwisle, & Olson, 2001, 2007) than their EO peers. Alternately, Hwang et al. (under review) showed that while EO and proficient LM students experienced summer setback in their general vocabulary knowledge, LEPs' magnitude of summer loss was smaller than their counterparts. However, LEPs also showed slower growth during the academic year. A somewhat different pattern was found with students' academic vocabulary knowledge. While EO, IFEP, and LEPs did not experience change in their learning trajectories during summer with academic vocabulary, RFEP learning rates were flatter than their EO learning rates during the summer. However, RFEP students' academic vocabulary learning trajectories during the school year were steeper than their EO peers'.

Because these studies were conducted in different contexts with heterogeneous samples using different measures, it is difficult to draw definite conclusions about adolescent LM students' growth trajectories in reading-related outcomes. Furthermore, no study to our knowledge has looked at within-group variation in RFEP students' growth during the academic year and the magnitude of summer setback or gain based on their years since redesignation. There are a very small number of studies on differences within the RFEP student category and this is a serious limitation. Differences in policies, definitions, and criteria across states make it difficult to be reconcile the findings suggesting RFEP students continue to struggle (e.g., Slama, 2014) with data suggesting that they are outperforming their peers (e.g., Ardasheva et al., 2012; Hwang et al., 2015). RFEP students are a large and growing group in the United States, but poorly understood. The current study builds on the previous cross-sectional analysis on RFEP students from California (Hwang et al., 2015) by examining their vocabulary and reading comprehension growth across middle grades for two years. The following research questions were addressed in this study:

1. What are the *general vocabulary* growth trajectories of middle school RFEP students by years since redesignation?

2. What are the *academic vocabulary* growth trajectories of middle school RFEP students by years since redesignation?

3. What are the *reading comprehension* growth trajectories of middle school RFEP students by years since redesignation?

Methods

Sample

Students from six middle schools in a large urban district in California contributed to the data collection of the current study. They were students from the control group in a larger study, which was a randomized efficacy trial of a vocabulary intervention. Initially, there were 1,294 RFEP students who contributed at least one wave of data collection. Of these, there were 59 who did not complete the necessary baseline assessments and 17 who did not have requisite demographic data. Our final analytical sample included 1,226 students (Table 3.1). The majority of the students in this sample (90%) was eligible for free or reduced lunch. Our sample consisted of 76% Asian, 17% Hispanic and 6% of other racial/ethnic backgrounds. In the second year of the study, 453 eighth grade students left the study (graduated middle school).

Table 3.1
Demographic Information of the Participants

	N
Total Students	1,226
Race	
Asian	76%
Hispanic	17%
Other	6%
Eligible for FRL	90%
Special Education	2%
Gifted and Talented Education	58%
Years since Redesignation	
Less than 1 Year	17%
Less than 2 Years	33%
Less than 3 Years	19%
More than 3 Years	31%

Note: FRL = Free or reduced lunch

Procedure

To assess students' academic vocabulary, general vocabulary, and reading comprehension, we administered a researcher-developed academic vocabulary test, and

standardized vocabulary and reading tests. All assessments were administered four times across

two consecutive academic years: once at the beginning of the school year (September/October)

and once at the end of the school year (May).

Measures

Time. TIME is a level-1 variable indicating the time since the start of the study when

students took the assessments. The data were collected the fall and spring of two consecutive

academic years. We coded each wave in months (i.e., wave 1 = 0 month, wave 2 = 7 months,

wave 3 = 12 months, wave 4 = 19 months).

Years since redesignation. The participating school district provided detailed

information about RFEP students' date of redesignation which we used to create a continuous

variable YEARS. This variable indicates how many years RFEP students have been considered

proficient by the district and no longer eligible for additional language support services. Values

on this variable ranged from .18 to 4.93, ($M = 2.29$, $SD = 1.34$). A histogram of YEARS shows

that the distribution is fairly even except for a peak between 1 and 2 years (Figure 3.1).
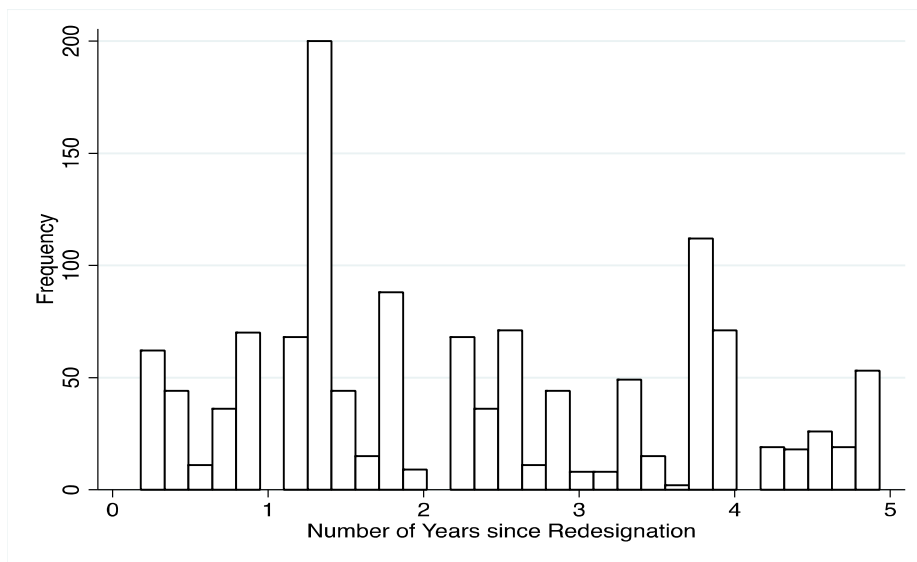
Figure 3.1. Histogram of RFEP students by their years since redesignation

General vocabulary. The Gates-MacGinitie Reading Test (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) is a group-administered assessment that includes reading and vocabulary subtests. As the vocabulary subtest assesses a wide range of general vocabulary knowledge (GEN_VOCAB), the score from this test was used as an indicator of students' general vocabulary knowledge. The level 6 Form T was administered to sixth grade students and the level 7/9 Form T was given to seventh- and eighth-grade students. The vocabulary subtest consists of 45 multiple-choice items that ask students to choose the synonym of a target word. Extended scale scores of the vocabulary subtest were used in the analysis as they allow for estimating growth over time on a single scale (MacGinitie et al., 2000). The internal reliability (Cronbach's α) for our analytical sample was .84 at the first wave. The mean and the standard deviation of the general vocabulary were 528.59 and 26.89, respectively.

Academic vocabulary. Students' academic vocabulary (ACA_VOCAB) was assessed with a 50-item multiple-choice test that was developed by the research team (All test forms can be found in the IRIS digital depository; www.iris-database.org; See Appendix A for the test forms that were used in this study). For each item, target word was embedded in a short sentence where students were asked to choose the closest synonym for a target word among four answer choices. Each target was has been found to be relatively frequent and dispersed across a range of first year college texts (Coxhead, 2000).

The same test form was administered twice each year. However, different forms were used in each of the two years of the study. The two different forms included 20 anchor items that appeared on both test forms, while the remaining 30 items were unique to each test form. To be able to score responses from the two forms on the same metric, we conducted test scaling using the item response theory (IRT) analysis. We ran a unidimensional three-parameter IRT model on

54

the first and third waves of data, constraining the anchor items to have same item parameters (i.e., difficulty and discrimination). This analysis was done with the sample in the larger study (Hwang et al., under review). In this process, we dropped seven anchor items, two unique items from Year 1 test form, and one unique item from Year 2 test form based on the overall fit index and local dependence statistics. The final model fit the data well (RMSEA = .03). Marginal reliability for the first wave with the larger sample was .91 and that of the third wave was .92. Once item parameters were obtained, we used them to score all the waves of our data. The scoring method we used was *expected a posteriori*. These scores were used as an indicator of academic vocabulary in our analysis. The scaled scores had a mean of .25 and standard deviation of .68 at the first wave for our sample.

Reading comprehension. The reading subtest in the Gates-MacGinitie Reading Test (MacGinitie et al., 2000) measures students' reading comprehension skills (READ). The level 6 Form T was administered to sixth grade students and the level 7/9 Form T was given to seventh and eighth grade students. Students were asked to read a passage and answer relevant comprehension questions (there are 48 items). Extended scale scores of reading subtest were used in our analysis. The Cronbach's α for our analytical sample was .86 in the first wave. The mean and standard deviation of our sample on reading comprehension were 538.01 and 29.27, respectively.

Grade-level cohort. To control for different grade levels in the analysis, student-level dummy variables were created for grade six (GRADE_6), seven (GRADE_7), and eight (GRADE_8). There were roughly equal numbers of sixth (27%), seventh (37%), and eighth (37%) grade students in our analytic sample.

Summer. The number of summers (SUMMER) students experience since the start of the study was also included in our analysis (e.g., wave 1 = 0, wave 2 = 0, wave 3 = 1, wave 4 = 1 for sixth and seventh grade cohorts). SUMMER is a time-varying continuous level-1 variable, and its parameter indicates whether or not students experience summer setback.

SES. Eligibility for receiving free or reduced lunch was used as an indicator of students' SES. A student-level dummy variable was created to indicate students who receive free or reduced lunch (FRL = 1) and those who do not (FRL = 0).

Ethnicity. Three student-level dummy variables (ASIAN, HISPANIC, and OTHER) were created to control for students' ethnicity. Asian students were used as the reference group in our analysis because they were by far the largest group (76% of the sample).

Special education status. A student-level dummy variable was created to indicate students who were receiving special education (SPED = 1) and those who were not (SPED = 0).

Gifted and talented education status. A student-level dummy variable was created to indicate students who were receiving gifted and talented education (GATE = 1) and those who were not (GATE = 0).

Data Analysis

To answer our research questions, we conducted series of multilevel models for change (Singer & Willet, 2003). These models allow us to use all waves of data from each student to create a model of vocabulary and reading growth over two years. The hypothesized multilevel model for change for the first research question was

Level-1 (outcomes in four waves across two years):

$$GEN\_\widehat{VOCAB} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}TIME^2{}_{ij} + \pi_{3i}SUMMER_{ij} + \varepsilon_{ij} \qquad (1)$$

Level-2 (student-level):

$$\pi_{0i} = \gamma_{00} + \gamma_{01}YEARSi + \gamma_{02}READ_i + \gamma_{03}ACA\_VOCAB_i + \gamma_{04}GRADE7_i +$$

$$\gamma_{05}GRADE8_i + \gamma_{06}FRL_i + \gamma_{07}HISPANIC_i + \gamma_{08}OTHER_i + \gamma_{09}SPED_i + \gamma_{10}GATE_i + \zeta_{0i}$$

$$(2)$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11}YEARS_i + \zeta_{1i} \tag{3}$$

$$\pi_{2i} = \gamma_{20} + \gamma_{21}YEARS_i \tag{4}$$

$$\pi_{3i} = \gamma_{30} + \gamma_{31}YEARS_i \tag{5}$$

where $\varepsilon_{ij} \sim N(0, \sigma_1{}^2)$, and $\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$.

The coefficient γ_{00} represents the average score for RFEP students who just have been redesignated (i.e., YEARS = 0) at the first wave (the first measurement point); γ_{10} represents the average initial slope for RFEP students with 0 year since redesignation; γ_{20} represents the average true acceleration for these students; and γ_{30} represents the average summer setback (or gain) for these students. The random effect ε_{ij} is a Level 1 residual for student i at time j and is assumed to be drawn from a normal distribution with mean of 0 and variance $\sigma_1{}^2$. Random effects ζ_{0i} and ζ_{1i} represent Level 2 residuals for the intercept and slope, respectively. They are both hypothesized to be drawn from a multivariate normal distribution with a mean of zero, unknown variances σ_0^2 and σ_1^2, and unknown covariance σ_{01}. For each research question, 1 unit change in YEARS variable is associated with γ_{01} change in baseline academic vocabulary test scores, γ_{11} change in their rate of growth, γ_{21} change in the acceleration of growth, and γ_{31} change in the summer learning controlling for all other covariates[10].

---

[10] We tested our analytic models with the continuous variable YEARS and also with the dummy variables for RFEP students based on their years since redesignation. Including series of dummy variables in the model did not significantly improve the model fit (e.g., **Δ**-2LL = 1.74; *df* = 3, *p* = *n.s.* for the model with academic vocabulary as the outcome).

Parameter estimates (γ_{01}, γ_{11}, γ_{21}, and γ_{31}) from our model were used to answer our first research question, controlling for reading comprehension (READ) and academic vocabulary (ACA_VOCAB) at the first wave. To answer our second research question, the same model with academic vocabulary (ACA_VOCAB) as the outcome and general vocabulary (GEN_VOCAB) and reading comprehension (READ) test scores from the first wave as covariates was used. Similarly, a model with reading comprehension (READ) as the outcome and academic vocabulary (ACA_VOCAB) and general vocabulary (GEN_VOCAB) test scores from the first wave as covariates was used to answer our last research question.

We calculated the predicted values for seventh grade RFEP students' general vocabulary, academic vocabulary, and reading comprehension scores based on our final model. We plotted these values as well as the national norms from the Gates-MacGinite Reading Test to compare RFEP students' predicted scores to those of the nationally representative sample.

Results

Preliminary Descriptive Analyses

Table 3.2
Correlations Among General Vocabulary, Academic Vocabulary, Reading Comprehension, and RFEP Students' Years Since Redesignation

	General Vocabulary	Academic Vocabulary	Reading Comprehension	Years since Redesignation
General Vocabulary	1			
Academic Vocabulary	0.79	1		
Reading Comprehension	0.69	0.73	1	
Years since Redesignation	0.47	0.47	0.46	1

Table 3.2 shows correlations among the outcome measures and RFEP students' years since redesignation. The correlation between academic vocabulary and general vocabulary was .79, that between academic vocabulary and reading comprehension was .73, and that between general vocabulary and reading comprehension was .69. As expected, vocabulary and reading measures were highly correlated with one another.

Table 3.3 displays means and standard deviations of students' general vocabulary, academic vocabulary, and reading comprehension test scores by their years since redesignation status across four waves. To clearly illustrate the relationship between RFEP students' years since redesignation and their reading-related outcome scores, we grouped RFEP students into four categories in this table: those who were redesignated less than one year before the first assessment, those who were redesignated one to two years before the first assessment, those who were redesignated two to three years earlier, and those who were redesignated more than three years earlier. Table 3.3 shows that RFEP students redesignated at different intervals from assessment vary in their ability in vocabulary and reading comprehension test scores.

General vocabulary. The first four columns of Table 3.3 show means and standard deviations for general vocabulary. RFEP students who were redesignated earlier performed better than those who were recently redesignated on the general vocabulary measure. This pattern was found in all grade level cohorts. Students gained general vocabulary knowledge over two years on average. There was a noticeable summer loss in general vocabulary scores for sixth grade cohort students (average score at Spring Year 1 = 535.94, average score at Fall Year 2 = 530.35). While the summer setback was not as pronounced for the seventh grade cohort students (average score at Spring Year 1 = 536.72, average score at Fall Year 2 = 539.85), the trajectory for vocabulary growth had flattened during the summer.

Table 3.3

Means and Standard Deviations of General Vocabulary, Academic Vocabulary, and Reading Comprehension Scores for RFEP Students by Their Years Since Redesignation

	Years since Redesignation	General Vocabulary				Academic Vocabulary				Reading Comprehension			
		Fall Year1	Spring Year1	Fall Year2	Spring Year2	Fall Year1	Spring Year1	Fall Year2	Spring Year2	Fall Year1	Spring Year1	Fall Year2	Spring Year2
6th Grade Cohort	Less than 1 Year	504.79 (28.56)	522.94 (30.05)	520.10 (23.83)	528.17 (29.42)	-0.35 (0.57)	0.02 (0.66)	0.14 (0.60)	0.27 (0.66)	510.79 (22.31)	519.59 (24.48)	533.68 (27.65)	536.85 (36.73)
	Less than 2 Years	518.32 (25.65)	534.98 (27.38)	532.74 (24.74)	543.23 (25.89)	-0.04 (0.59)	0.34 (0.60)	0.50 (0.62)	0.72 (0.61)	523.57 (22.91)	537.88 (25.70)	545.72 (26.46)	555.91 (30.79)
	Less than 3 Years	530.67 (24.46)	549.91 (28.69)	538.20 (24.00)	551.39 (24.96)	0.15 (0.59)	0.56 (0.65)	0.60 (0.55)	0.89 (0.60)	525.48 (25.83)	539.88 (30.23)	547.30 (27.65)	557.78 (32.49)
	More than 3 Years	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
7th Grade Cohort	Less than 1 Year	515.86 (23.23)	529.23 (20.91)	532.98 (20.48)	541.66 (27.15)	-0.09 (0.61)	0.18 (0.59)	0.32 (0.57)	0.46 (0.62)	525.68 (29.21)	533.33 (27.93)	541.69 (27.71)	546.14 (31.22)
	Less than 2 Years	522.51 (21.76)	530.97 (22.56)	534.75 (22.21)	546.56 (25.07)	0.14 (0.58)	0.37 (0.61)	0.41 (0.63)	0.69 (0.60)	537.34 (25.44)	545.82 (26.88)	550.51 (29.51)	559.08 (29.29)
	Less than 3 Years	526.03 (23.29)	540.19 (22.31)	539.55 (18.60)	556.68 (27.63)	0.16 (0.60)	0.44 (0.53)	0.69 (0.43)	0.85 (0.52)	538.78 (21.43)	554.65 (25.26)	552.58 (21.00)	555.92 (36.60)
	More than 3 Years	536.85 (21.55)	546.50 (23.59)	552.10 (22.98)	561.12 (26.05)	0.40 (0.62)	0.77 (0.57)	0.81 (0.61)	1.03 (0.61)	547.88 (27.02)	558.29 (29.20)	562.83 (32.35)	571.99 (27.08)
8th Grade Cohort	Less than 1 Year	518.39 (19.61)	529.29 (24.15)	- -	- -	0.11 (0.44)	0.38 (0.52)	- -	- -	527.90 (21.53)	542.30 (20.61)	- -	- -
	Less than 2 Years	525.72 (22.11)	536.68 (21.65)	- -	- -	0.17 (0.57)	0.49 (0.61)	- -	- -	536.29 (19.19)	543.32 (26.73)	- -	- -
	Less than 3 Years	533.59 (27.34)	544.83 (26.25)	- -	- -	0.37 (0.68)	0.62 (0.74)	- -	- -	542.51 (29.81)	553.21 (28.19)	- -	- -
	More than 3 Years	549.35 (22.00)	561.61 (24.39)	- -	- -	0.85 (0.52)	1.13 (0.56)	- -	- -	562.22 (25.32)	570.85 (26.09)	- -	- -

Note: Standard deviations in parentheses. The Academic Vocabulary scores are scaled scores from two different forms. General Vocabulary and Reading Comprehension scores are the extended scaled scores from Gates-MacGinitie Reading Test. Grade cohorts indicate student's grade levels in the beginning of the study.

Academic vocabulary. The fifth through eighth columns of Table 3.3 describe students'

performance on the academic vocabulary measure by years since redesignation. Across all grade

level cohorts, RFEP students' academic vocabulary scores were correlated with their years since

redesignation. In other words, the earlier redesignated, the higher their baseline test scores were.

Although students in all groups showed growth in their academic vocabulary over time, the score

gap between groups persisted throughout their middle school years. Students did show some

growth in their academic vocabulary knowledge during summer (i.e., in between Spring Year 1

and Fall Year 2); however, from this descriptives table, it was difficult to tell whether such

growth during summer was pronounced enough to change students' overall learning trajectories.

Reading comprehension. As with the two vocabulary measures, RFEP students' reading

comprehension test scores were associated with the years since redesignation, which means that

those who were redesignated early had higher scores compared to those who were redesignated

recently. On average, students showed growth in their reading comprehension over the course of

two years. From this descriptive statistics table, how much students gained reading

comprehension skills seemed to vary according to their years since redesignation. Especially for

students in the seventh grade cohort, those who were redesignated recently showed relatively

large gains over the summer break whereas those who were redesignated earlier showed

relatively small gains or even some in with their reading comprehension skills.

Growth Modeling Results

The descriptive results illustrated earlier do not allow us to leverage the power of

multiple measurement occasions within individuals to improve estimates at each wave, nor

control for any covariates.

Table 3.4
Multilevel Models for Change Predicting General Vocabulary, Academic Vocabulary, and
Reading Comprehension Scores

		General Vocabulary	Academic Vocabulary	Reading Comprehension
		Model A	Model B	Model C
Fixed Effects	Intercept	399.06***	-10.01***	366.15***
		(11.03)	(0.23)	(15.74)
	YEARS	2.07***	0.03**	0.68
		(0.37)	(0.01)	(.47)
	TIME	1.90***	0.05***	1.56***
		(0.09)	(0.002)	(0.11)
	TIME^2	-0.01**	-0.001***	-0.02**
		(0.01)	(0.0001)	(0.01)
	SUMMER	-7.88***	-0.10***	3.24*
		(0.82)	(0.02)	(1.38)
	SUMMER X YEARS			-1.17*
				(0.54)
	GEN_VOCAB		0.01***	0.29***
			(0.001)	(0.03)
	READ	0.24***	0.01***	
		(0.02)	(0.001)	
	ACA_VOCAB	21.24***		15.34***
		(0.89)		(1.23)
	GRADE_7	-4.76***	-0.05*	9.40***
		(1.08)	(0.02)	(1.29)
	GRADE_8	-4.60***	0.07*	12.07***
		(1.28)	(0.03)	(1.52)
	HISPANIC	-3.92***	-0.03	-4.16**
		(1.11)	(0.03)	(1.36)
	OTHER	-1.84	0.04	-0.48
		(1.68)	(0.04)	(2.05)
	FRL	-3.60**	-0.01	1.65
		(1.34)	(0.03)	(1.64)
	SPED	-7.25*	-0.03	-5.64***
		(3.35)	(0.07)	(4.06)
	GATE	0.67***	0.13***	11.42***
		(0.96)	(0.02)	(1.14)
Level 1 Variance Component	Residual	132.47***	0.09***	174.67
		(4.64)	(0.003)	(6.15)
Level 2 Variance Component	Intercept	138.77***	0.06***	197.23***
		(10.57)	(0.01)	(14.39)
	TIME	0.44***	0.0002***	0.79***
		(0.06)	(0.00004)	(0.10)
	Covariance	-1.49*	-0.0004	-0.99
		(0.68)	(0.0004)	(0.94)
	N (Students)	1,190	1,218	1,182
	N (Observations)	3,687	3,658	3,661

We turn now to a longitudinal growth models to answer each research question. Table 3.4 shows the results from the multilevel models for change predicting academic vocabulary, general vocabulary, and reading comprehension across four waves. The inclusion of the quadratic term (TIME^2) improved the model fit in all three outcomes: general vocabulary (Δ-2LL = 7.94; df = 1, $p < .01$), academic vocabulary (Δ-2LL = 18.15; df = 1, $p < .001$), and reading comprehension (Δ-2LL = 8.54; df = 1, $p < .01$). The significant positive linear terms indicate that students tend to improve on these skills across waves. The significant negative quadratic terms indicate that the rates of growth decrease.

General vocabulary (RQ1). Model A in Table 3.4 presents results for RFEP students' growth in general vocabulary over two years. The coefficient associated with YEARS was positive and significant ($\beta = 2.07$, $p < .001$), which indicates that when other variables are controlled, one year interval since redesignation was associated with 2.07 points difference in the baseline score in general vocabulary on average. On average, RFEP students showed growth in general vocabulary ($\beta = 1.90$, $p < .001$; 1.90 points per month) during this two-year period. The interaction term of YEARS and TIME was not statistically significant and was not included in our final model. This means that the rate of growth in academic vocabulary did not differ by RFEP students' years since redesignation. The negative and significant TIME^2 coefficient ($\beta = -.01$, $p = .005$) indicates that RFEP students' rate of growth decreased over time. Students in our sample experienced summer setback in their general vocabulary knowledge on average (SUMMER; $\beta = -7.88$, $p < .001$). These results are demonstrated in Figure 3.2.
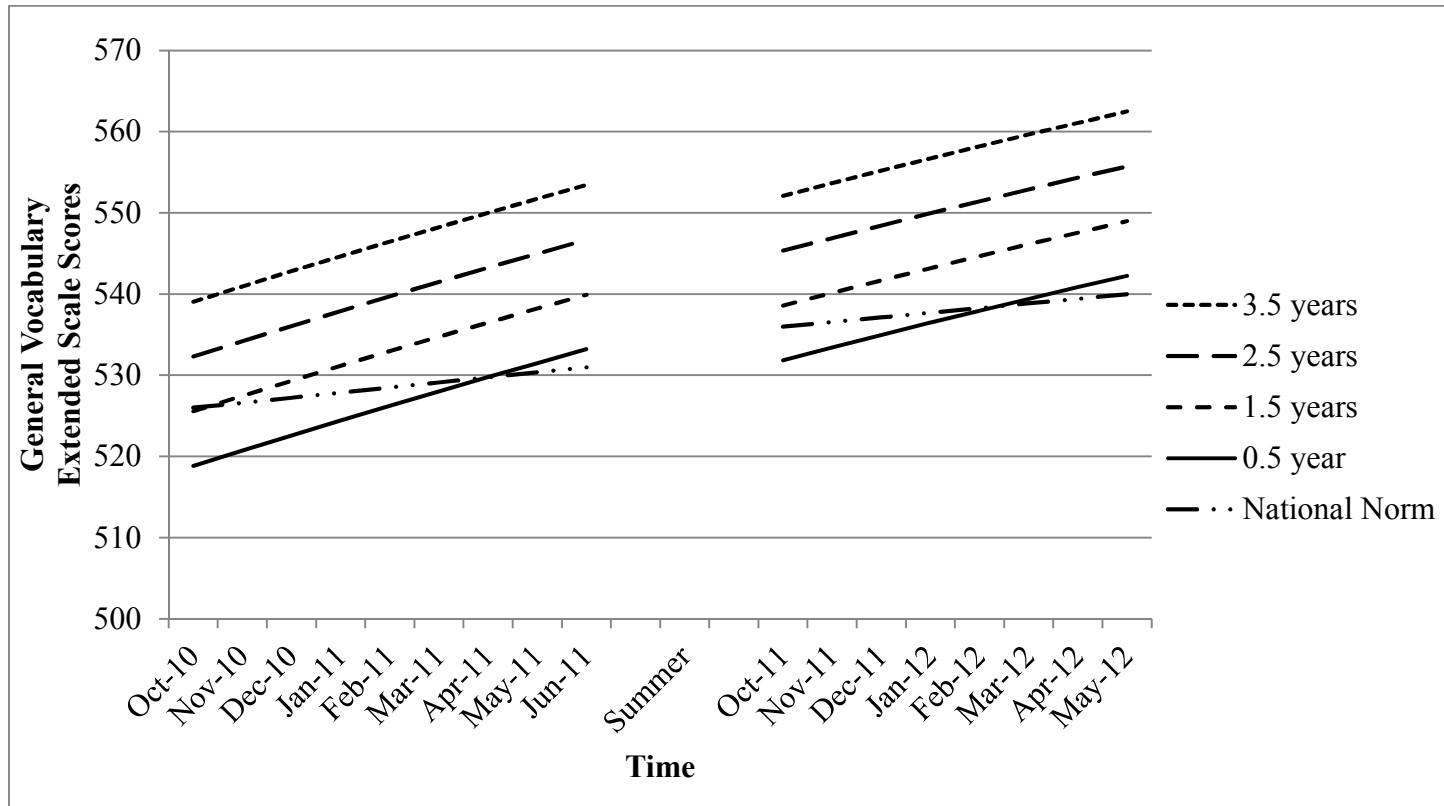
Figure 3.2. Prototypical graph of general vocabulary growth trajectory for RFEP students by their years since redesignation

Figure 3.2 shows prototypical graphs for seventh grade RFEP students' general

vocabulary growth trajectories[11]. On average, students who were redesignated 3.5 years before

the start of the study (squared dot line, average baseline score = 539.08[12]) already demonstrated

significant skill in general vocabulary relative to other RFEP students. Students who were

redesignated 2.5 years before the start of the study (long dashed line) did not perform as well

(average baseline score = 532.33), but they outperformed students who were redesignated 1.5

years before the start of the study (average baseline score = 525.58) or .5 year before the start of

the study (average baseline score = 518.83). As we noted in the final fitted model, there is no

difference across groups in either growth, acceleration, or summer setback: the lines representing

prototypical general vocabulary growth for seventh-grade students are parallel.

Academic vocabulary (RQ2). Model B in Table 3.4 presents results for RFEP students'

growth in academic vocabulary over two years. The coefficient associated with YEARS was

positive and significant (β = .03, p = .003), which indicates that when other variables are

controlled, one year of redesignation was associated with .03 points difference in academic

vocabulary baseline scores. For instance, the predicted difference in the baseline score of RFEP

students who was redesignated at the beginning of the study (YEARS = 0) and one who was

redesignated one year before the start of the study (YEARS = 1) would be .03 when other

---

[11] Although the YEARS variable was a continuous variable, we arbitrarily classified our sample into four groups based on their years since redesignation (i.e., .5, 1.5, 2.5, and 3.5 years) so that our results are easier to interpret.

[12] The average baseline score was calculated by summing the following variables: 399.06 (constant) - 4.76 (coefficient for seventh grade) + 0.24 (coefficient for reading comprehension) X 547.15 (predicted mean score of reading comprehension for seventh grade RFEP students with 3.5 years since redesignation) + 21.24 (coefficient for academic vocabulary) X 0.41 (predicted mean score of academic vocabulary for seventh grade RFEP students with 3.5 years since redesignation). Same process for calculating the constant was done for seventh grade RFEP students with .5, 1.5, and 2.5 years since redesignation.

variables are controlled for. On average, RFEP students showed growth in academic vocabulary (β = .05, p < .001; .05 points per month) during their middle school years. The interaction term of YEARS and TIME was not statistically significant and was not included in our final model. This means that the rate of growth in academic vocabulary did not differ by RFEP students' years since redesignation. The negative and significant TIME^2 coefficient (β = -.001, p < .001) indicates that RFEP students' rate of growth decreased over time. Students in our sample experienced some summer setback on average (SUMMER; β = -.10, p < .001) in their academic vocabulary knowledge. However, the magnitude of change in their learning trajectory over summer did not differ by RFEP students' years since redesignation.

Figure 3.3 shows prototypical graph for seventh grade RFEP students. On average, students with 3.5 years since redesignation before the start of the study (squared dot line, average baseline score = .30) already demonstrated significantly greater skill in academic vocabulary than other RFEP students. Students who were redesignated 2.5 years before the start of the study (long dashed line) did not perform as well (average baseline score = .15), but they outperformed students who were redesignated 1.5 years before the start of the study (average baseline score = .01) or .5 year before the start of the study (average baseline score = -.14). As we noted in the final fitted model, there is no difference across groups in growth, acceleration, or summer setback: the lines representing prototypical academic vocabulary growth for seventh-grade students are parallel.
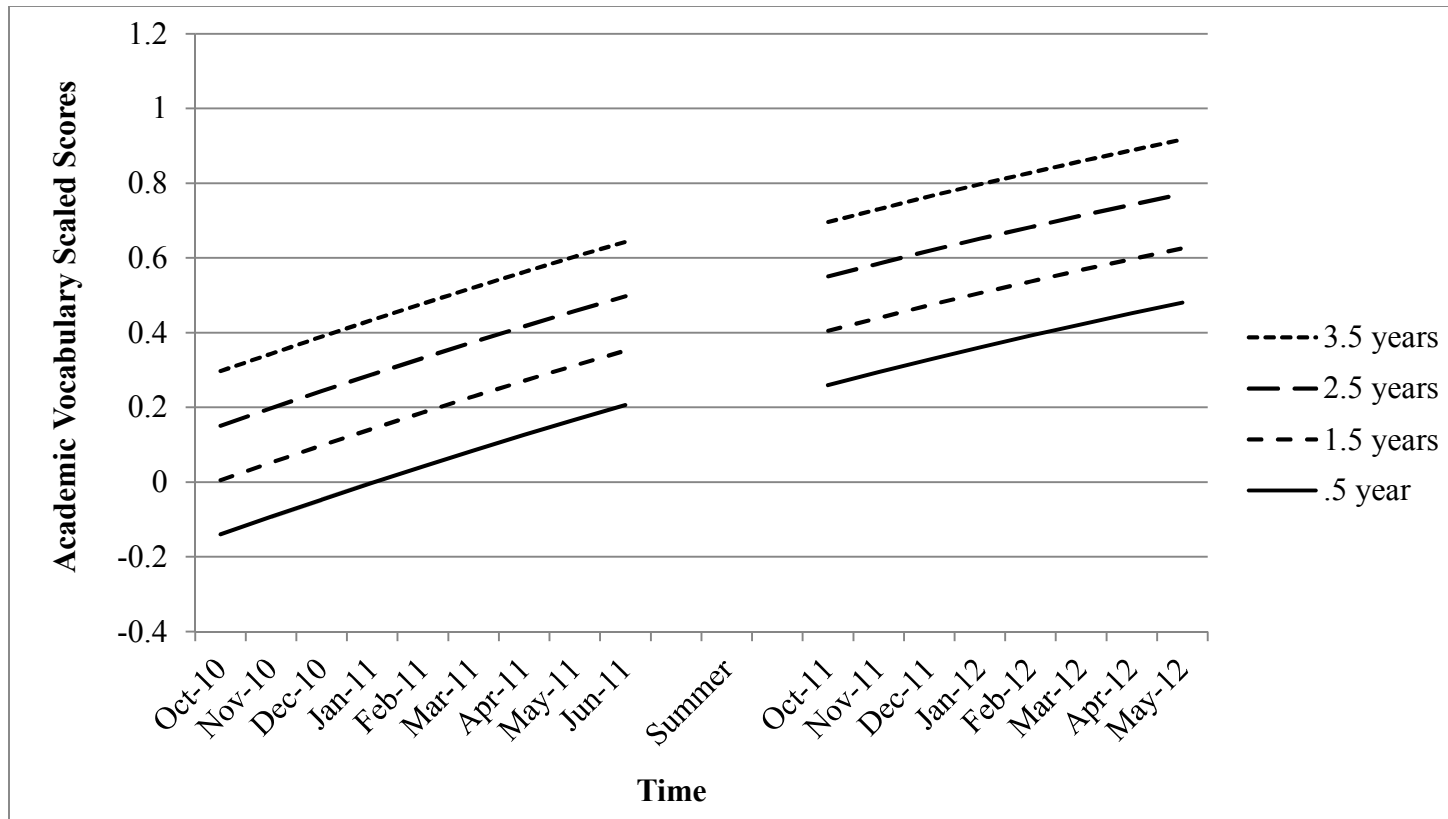
Figure 3.3. Prototypical graph of academic vocabulary growth trajectory for RFEP students by their years since redesignation

Reading comprehension (RQ3). Model C in Table 3.4 presents results for RFEP students' growth in reading comprehension over two years. The coefficient associated with YEARS was not statistically significant ($\beta = .68$, $p = .145$) when other student-level variables are controlled. On average, RFEP students showed growth in reading comprehension ($\beta = 1.56$, $p < .001$; 1.56 points per month) during this two-year period. The interaction term of YEARS and TIME was not statistically significant and was not included in our final model. This means that the rate of growth in academic vocabulary did not differ by RFEP students' years since redesignation. The negative and significant TIME^2 coefficient ($\beta = -.02$, $p = .003$) indicates that RFEP students' rate of growth decreased over time. The coefficient for SUMMER was positive and statistically significant ($\beta = 3.24$, $p < .001$), which means that, on average, students with 0 year since redesignation experienced a gain in their learning trajectories during the summer. However, the coefficient for SUMMER X YEARS was negative and statistically significant ($\beta = -1.17$, $p = .030$). This means that the magnitude of change in their learning trajectories during the summer differed according to RFEP students' years since redesignation. These results are demonstrated in Figure 3.4.

Figure 3.4 shows prototypical graph for seventh grade RFEP students based on our final fitted model. On average, RFEP students who were redesignated 3.5 years before the start of the study (squared dot line, average baseline score = 537.60) demonstrated higher scores in reading comprehension relative to other RFEP students. Students who were redesignated 2.5 years before the start of the study (long dashed line) did not perform as well (average baseline score = 532.75), but they outperformed students who were redesignated 2.5 years before the start of the study (average baseline score = 527.90) or .5 year before the start of the study (average baseline score = 523.05). As we noted in the final fitted model, although there was no difference across groups
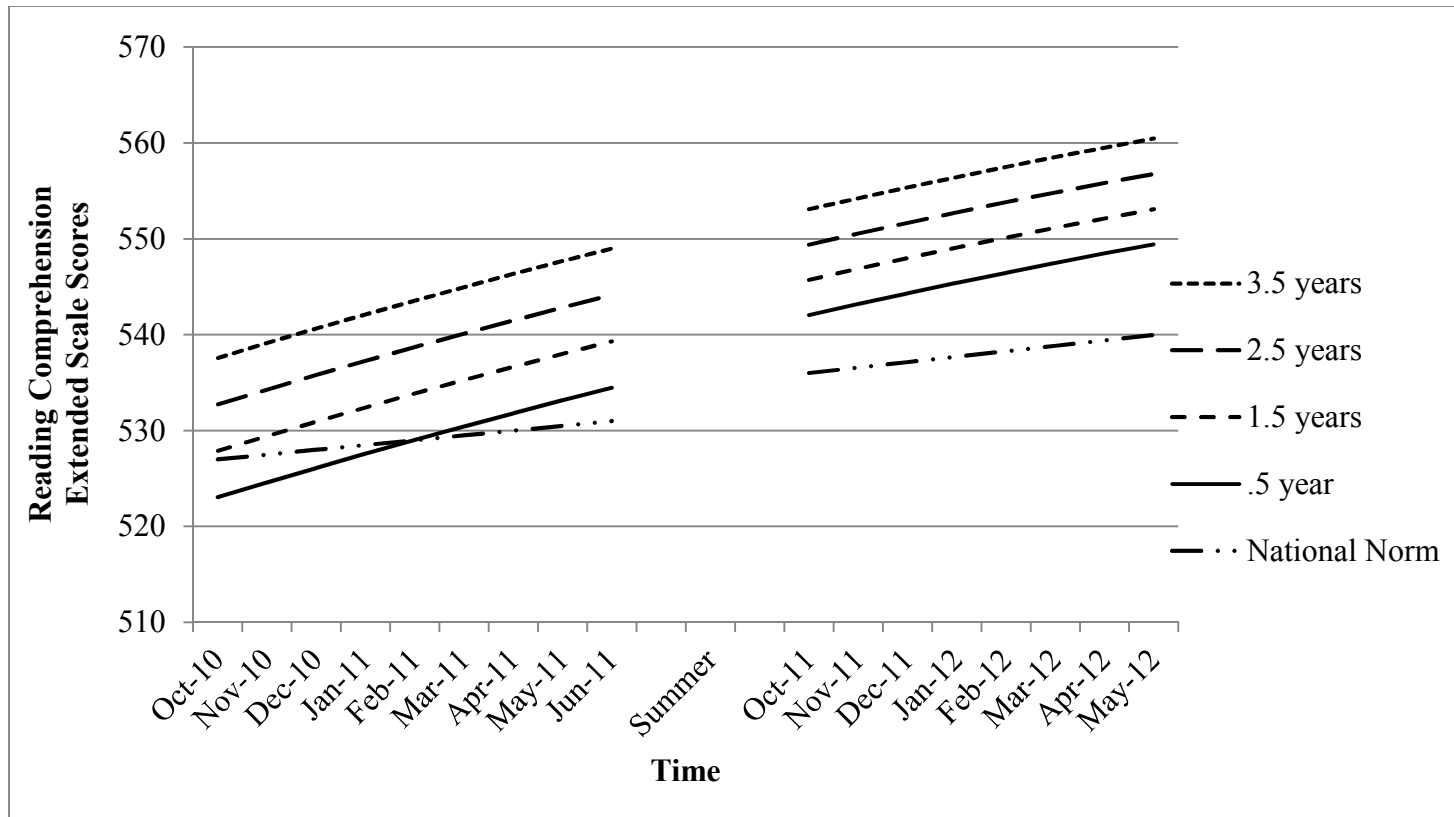
Figure 3.4. Prototypical graph of reading comprehension growth trajectory for RFEP students by their years since redesignation

in either growth or acceleration, there was a differential effect of summer producing a slight narrowing of the gaps between groups during their second academic year.

Discussion

We examined RFEP students' vocabulary and reading growth trajectories to understand the potential variability within the RFEP student population. We discuss some similarities and differences in RFEP students' vocabulary and reading growth trajectories below.

Similarities in vocabulary and reading comprehension growth trajectories

Our results demonstrate that RFEP students' general vocabulary, academic vocabulary, and reading comprehension performance are linearly related to the amount of time since they had been redesignated as being proficient in English. That is, although students who are redesignated later in their academic careers may show comparable growth to LM students who have been redesignated in early grades, they did not close the achievement gap in vocabulary and reading comprehension across the middle school grades. Our findings, together with past studies, highlight the need to promote proficiency in English early for LM students (Hwang et al., under review; Kieffer, 2008, 2011; Mancilla-Martinez et al., 2011).

The early-redesignated students outperformed recent-redesignated students in both vocabulary and reading outcomes in the current study. These results are consistent with previous research where students' baseline scores predict their later outcomes in reading-related outcomes (Hwang et al., under review; Kieffer, 2008; Mancilla-Martinez et al., 2011). The rate of growth did not differ among groups of RFEP students with different years since redesignation and the gap among the groups remained throughout their middle school years (especially with their vocabulary knowledge). Even so, these rank orders in students' vocabulary and reading scores

70

suggest that recent-redesignated students still have room to grow in terms of their literacy skills. This indicates that they may benefit from explicit vocabulary or reading instruction in order to expedite their learning rate during their school years and catch up to their early-redesignated peers.

An important consideration when interpreting our findings is that the state of California is known to have stringent criteria for redesignation. Indeed, approximately 60% of secondary LEP students in California do not meet the criteria for redesignation after six years of instruction (Olsen, 2010; Umansky & Reardon, 2014). Further, the assessment benchmarks used to determine eligibility for redesignation, particularly the state English Language Arts tests and the CELDT, become more rigorous with increasing grade levels. Thus, although there was a clear rank order in reading-related outcomes among subgroups of RFEP students, the total sample of RFEP students showed high achievement in vocabulary and reading comprehension, with even the recently redesignated students exceeding the national average eventually.

Differences in vocabulary and reading comprehension growth trajectories

Our findings highlight the importance of considering specific skills when considering academic loss over the summer, rather than treating summer setback as a single construct. For example, we found different growth trajectories over the summer for general vocabulary, academic vocabulary and reading comprehension. For both types of vocabulary, RFEP students learning trajectories slowed during the summer months. They experienced loss of approximately four months-worth of their growth during summer for general vocabulary (i.e., coefficient for TIME = 1.90, coefficient for SUMMER = -7.88) and approximately two months of growth for academic vocabulary (i.e., coefficient for TIME = .05, coefficient for SUMMER = -.10) on average. Our findings were consistent with others reporting that both EO and (English-proficient)

LM students experience summer vocabulary loss in English (Hwang et al., under review; Lawrence, 2012; Lawrence et al., 2012).

However, we were encouraged by the differential summer reading comprehension trajectories. Students who were redesignated most recently showed greater gains over the summer months than they had during the school year. In contrast, although students who had been redesignated two or more years before the start of the study had higher baseline scores, they showed more pronounced summer setback. Although the differential summer learning reduced the gaps between groups of RFEP students, the underlying mechanism is unclear. It is possible that students who had been redesignated early may have read as intensively during the summer, thereby experiencing a steeper setback during summer months. However, this explanation is speculative, as we did not have any data on RFEP students' summer activities to investigate this issue further. In any case, the implication for educators seems to be that summer may be a particularly important leverage point for supporting redesignated language learners.

Concluding remarks

Our findings indicate that RFEP students' years since redesignation is positively correlated with their vocabulary and reading comprehension outcomes. However, we are not suggesting that early redesignation causes students' higher reading performance. In this study, it was not possible to test the causal relationship between redesignation and students' reading-related performance. In other words, we could not test whether early redesignation helped students to benefit from mainstream classroom instruction and develop their language and literacy skills or students who had high linguistic skills were more likely to be redesignated early in their schooling history and were outperforming their peers.

However, we did find that once stringent criteria had been used to redesignate LM students, RFEP students in general perform comparably or even better than students in the national norming sample. Although clear rank order existed among subgroups of RFEP students based on their years since redesignation, the average test scores in reading-related outcomes of the RFEP students in our sample were fairly high. The state of California is known to have rigorous criteria for reclassifying its LEP students. Not only do LEP students need to prove their English proficiency through multiple sources of data (e.g., English proficiency test, content-area standards test, teacher interview, grades in school), they need to meet the cut-off for all the assessments at once. Thus, it is not surprising that RFEPs perform well on literacy-related assessments. Unfortunately, this high bar also explains why the state of California has a high percentage of long-term LEP students, those who do not get redesignated after receiving several years of English instruction ((Olsen, 2010; Umansky & Reardon, 2014). Rigorous redesignation criteria seem to have two seeming contrary consequences: high numbers of high-performing RFEP students and high numbers of long-term LEP students.

It is important to note the demographics of the sample in this study while interpreting our findings. The majority of the students in our sample were Asian, and only 17% were Hispanics as indicated in Table 3.1. Generally, students' racial backgrounds are associated with their language and literacy outcomes in the United States, and Asians have shown to outperform Hispanics and African-Americans on average (e.g., Fryer & Levitt, 2006; Kao & Thompson, 2003; National Center for Education Statistics, 2011). Thus, this was a distinctive sample in an urban school district in California that we need to take caution in applying current findings to other school districts serving different student population.

Limitation and Future Directions

One of the shortcomings of this paper is that we have no data on their after-school activities (e.g., debate club, private tutoring) or summer school activities that could have positively influenced their vocabulary and reading outcomes. Future studies that incorporate adolescent RFEP students' extra-curricular activities would generate better explanations of their growth in reading-related outcomes. Additionally, it is important to note the demographics of the current sample. As noted earlier, the RFEP students in this study are not a nationally representative sample. The participants of this study drawn from a single school district, with most students reporting Chinese as their home language and most receiving gifted and talented education whereas the majority (75%) of the LM students in the United States are from Spanish-speaking homes (National Clearinghouse for English Language Acquisition, 2011). Despite the different student demographics, our findings were convergent with findings involving less advantaged populations (e.g., Kieffer, 2011; Mancilla-Martinez et al., 2011).

As the process for redesignation varies to a great degree across states, we need more research on RFEP students to examine whether the current findings hold in other contexts, such as when criteria for redesignation are more lenient. Although more research is needed and despite the limitations of the current study, these findings fill a gap in the research literature and improve our understanding of adolescent RFEP students' vocabulary and reading development in middle grades. The findings of the current study underscore that adolescent RFEP students with different years since redesignation have parallel growth trajectories in vocabulary and reading comprehension yet have different starting points. Our results highlight that RFEP students who were redesignated with a rigorous criteria are likely to show high performance in reading-related outcomes. However, it may be difficult to close the achievement gap within the RFEP student population despite common response to education and that educators and policymakers need to

invest in earlier supports for LEP students so that they may be redesignated as early as possible and enjoy better reading-related outcomes.

References

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment, 8*(3), 231-257. doi:10.1207/S15326977EA0803_02

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practice, 27*(3), 17-31. doi:10.1111/j.1745-3992.2008.00125.x

Alexander, K., Entwisle, D., & Olson, L. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation & Policy Analysis, 23*(2), 171-191. doi:10.3102/01623737023002171

Alexander, K., Entwisle, D., & Olson, L. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72*(2), 167-180. doi:10.1177/000312240707200202

Ardasheva, Y., Tretter, T. R., & Kinny, M. (2012). English language learners and academic achievement: Revisiting the threshold hypothesis. *Language Learning, 62*(3), 769-812. doi:10.1111/j.1467-9922.2011.00652.x

Aud, S., Hussar, W., Kena, G., Bianco, K., Frohlich, L., Kemp, J., & Tahan, K. (2011). *The Condition of Education 2011 (NCES 2011-033).* Washington, DC: U.S. Department of Education, National Center for Education Statistics, U.S. Government Printing Office.

August, D., Carlo, M., Dressler, C., & Snow, C. E. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50-57. doi:10.1111/j.1540-5826.2005.00120.x

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

Bailey, A. L., & Kelly, K. R. (2013). Home language survey practices in the initial identification

    of English learners in the United States. *Educational Policy, 27*(5), 770-804.

    doi:10.1177/0895904811432137

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary*

    instruction. New York, NY: Guilford Press.

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., . . .

    Gillam, R. B. (2012). The measure matters: Language dominance profiles across

    measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition,*

    15(03), 616-629. doi:10.1017/S1366728912000090

Bunch, G. C. (2013). Pedagogical language knowledge preparing mainstream teachers for

    English learners in the new standards era. *Review of Research in Education, 37*(1), 298-

    341. doi:10.3102/0091732X12461772

Chall, J. S., & Jacobs, V. A. (2003). Poor children's fourth-grade slump. *American Educator,*

    27(1), 14-17.

Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second

    language. *TESOL Quarterly, 23*(3), 509-531. doi:10.2307/3586923

Corson, D. (1997). The learning and use of academic English words. *Language Learning, 47*(4),

    671-718. doi:10.1111/0023-8333.00025

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

    doi:10.2307/3587951

English Language Proficiency Assessment of 1999, California Education Code §§ 313 - 313.5

    (2014).

Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). Practical guidelines for

    the education of English language learners: Research-based recommendations for

    instruction and academic interventions. Portsmouth, NH: RMC Research Corporation,

    Center on Instruction. Available online at http://www.centeroninstruction.org/files/ELL1-

    Interventions.pdf

Goldenberg, C. (2008). Teaching English language learners: What the research does--and does

    not--say. *American Educator, 32*, 7-23, 42-44.

Harklau, L. (1994). ESL versus mainstream classes: Contrasting L2 learning environments.

    TESOL Quarterly, 28(2), 241-272. doi:10.2307/3587433

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2),

    127-160. doi:10.1007/BF00401799

Hwang, J. K., Lawrence, J. F., Mo, E., & Snow, C. E. (2015). Differential effects of a systematic

    vocabulary intervention on adolescent language minority students with varying levels of

    English proficiency. *International Journal of Bilingualism, 19*(3), 314-332.

    doi:10.1177/1367006914521698

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (under review). Differential vocabulary growth

    trajectories among adolescent language minority students: A two-year longitudinal study.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated

    poverty, and the reading growth of language minority learners in the United States.

    Journal of Educational Psychology, 100(4), 851-868. doi:10.1037/0022-0663.100.4.851

Kieffer, M. J. (2011). Converging trajectories reading growth in language minority learners and

    their classmates, kindergarten to grade 8. *American Educational Research Journal, 48*(5),

    1187-1225. doi:10.3102/0002831211419490

Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168-1201. doi:10.3102/0034654309332490

Kim, J., & Herman, J. L. (2009). A three-state study of English learner progress. *Educational Assessment, 14*(3-4), 212-231. doi:10.1080/10627190903422831

Lawrence, J. F. (2012). English vocabulary trajectories of students whose parents speak a language other than English: Steep trajectories and sharp summer setback. *Reading and Writing, 25*(5), 1113-1141. doi:10.1007/s11145-011-9305-z

Lawrence, J. F., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. E. (2012). Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program. *Bilingualism: Language and Cognition, 15*(3), 437-451. doi:10.1017/S1366728911000393

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing, 19*(1), 99-131. doi:10.1007/s11145-005-4713-6

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading test technical report: Forms S and T.* Chicago, IL: The Riverside Publishing Company.

Mancilla-Martinez, J., Kieffer, M. J., Biancarosa, G., Christodoulou, J. A., & Snow, C. E. (2011). Investigating English reading comprehension growth in adolescent language minority learners: Some insights from the simple view. *Reading and Writing, 24*(3), 339-354.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*(5), 1544-1560. doi:10.1111/j.1467-8624.2011.01633.x

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

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing, 20*(7), 691-719. doi:10.1007/s11145-006-9045-7

National Center for Education Statistics. (2011). *The nation's report card: Reading 2011 (NCES 2012-457)*. Retrieved from Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Clearinghouse for English Language Acquisition. (2011). *What langauges do English learners speak? NCELA Fact Sheet*. Washington, DC: Author.

Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational opportunity for California's long term English learners*. Long Beach, CA: Californians Together.

Ragan, A., & Lesaux, N. K. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives, 14*(20), 1-32. doi:10.14507/epaa.v14n20.2006

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267-292. doi:10.3102/0162373711407912

Rubio, E. V. (2014). *English language learners in an accountability landscape: Examining reclassification, long-term ELLs, language programs, and grade retention* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses A&I. (Order No. 3662579)

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97-110). New York: Guilford Press.

Scarcella, R. C. (2003). *Academic English: A conceptual framework.* Berkeley, CA: The University of California Linguistic Minority Research Institute.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective.* Mahwah, NJ: Lawrence Erlbaum.

Singer, J., & Willet, J. (2003). *Applied longitudinal data analysis: Modeling change and even occurrence.* New York, NY: Oxford University Press.

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

Solórzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research, 78*(2), 260-329. doi:10.3102/0034654308317845

Umansky, I. M., & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal, 51*(5), 879-912. doi:10.3102/0002831214545110

CHAPTER 4

Study 3. Investigating the effects of Word Generation on
adolescent language minority students: A longitudinal follow-up study

A substantial number of adolescent students struggle to meet the increasing postsecondary education and workplace literacy demands (Murnane, Sawhill, & Snow, 2012). Many adolescent language minority (LM) learners (i.e., students whose home language is not English) in the United States face dual challenges in school because they need to gain both content-area knowledge and English language skills simultaneously (Short & Fitzsimmons, 2007). Unsurprisingly, LM students with limited English proficiency, fall behind their English-proficient peers in English academic domains such as reading and writing (e.g., National Center for Education Statistics, 2011, 2012). When LM students start school with limited English proficiency, it is difficult to catch up to their English-proficient peers in reading-related outcomes in later grades (Hwang, Lawrence, & Snow, under review; Kieffer, 2011; Lawrence, 2012; Nakamoto, Lindsey, & Manis, 2007; Reese, Garnier, Gallimore, & Goldenberg, 2000).

Academic language interventions have been shown to enhance elementary and secondary LM students' literacy skills (e.g., Lesaux, Kieffer, Kelly, & Harris, 2014; Snow, Lawrence, White, 2009; Townsend & Collins, 2009). Some interventions even support LM students' literacy development more than that of EO students (e.g., Carlo et al., 2004; Hwang, Lawrence, Mo & Snow, 2015; Snow et al., 2009). However, more needs to be known about the long-term treatment effects of such interventions and whether students are able to maintain improvements. Thus, the current study examines the short- and long-term effects of a particular intervention, Word Generation (WG), on adolescent LM students' academic vocabulary knowledge.

Language minority students and academic language

LM students are any school-aged students in the United States who are exposed to a language other than English at home (August & Shanahan, 2006). They have been classified into three groups depending on their English language proficiency (Ragan & Lesaux, 2006): initially fluent English proficient (IFEP), redesignated fluent English proficient (RFEP), and limited English proficient (LEP). IFEP students are those who enter school with sufficient English proficiency (as measured by state- or district-specific assessments). These students do not receive additional English language support in school. LEPs are those who have not met the English proficiency criteria set by the state or school district and receive English language services in school. RFEPs are students who were classified as LEPs when they entered school but eventually met criteria and were reclassified. Because there is no federal guideline used to identify and classify LM students, there is a great variation in how different states and school districts identify them. How subgroups of LM students perform depends largely on how rigorous the classification criteria used to identify them is (Hwang et al., 2015; Kim & Herman, 2009).

Mastery of academic language plays an important role in academic success for all students including those who are LM, especially in post-primary grades. With the advent and implementation of the Common Core State Standards (Common Core State Standards Initiative, 2010) across many states in the United States, more emphasis is being placed on students' academic language skills across disciplines. Although it is difficult to define exactly what academic language is, researchers contend that it is a language that students encounter often in school, and consists of linguistic features that are distinct from social, colloquial language (Bailey, 2007; Scarcella, 2003; Schleppegrell, 2001, 2004; Snow & Uccelli, 2009; Uccelli, Galloway, Barr, Meneses, & Dobbs, 2015).

Academic vocabulary refers to vocabulary used in academic discourse that is relatively low frequency and often abstract and technical (Schleppegrell, 2001; Snow & Uccelli, 2009). Researchers have categorized academic vocabulary into two groups: general and content-specific (Beck, McKeown, & Kucan, 2002; Chung & Nation, 2003; Corson, 1997; Nagy & Townsend, 2012). General academic words are high-leverage words that may appear across multiple content areas. Examples of general academic words include words such as *assess*, *eliminate*, *acquire*, and *obtain*. Unlike general academic words, content-specific academic words are tied to a certain discipline such as *photosynthesis* and *hypotenuse*.

Because LEPs are students who are acquiring English language skills, they still need to develop academic language skills (Hwang et al., 2015; National Center for Education Statistics, 2012a; Slama, 2012; Uccelli et al., 2015). However, results have been mixed regarding the performances of proficient LM students (i.e., IFEPs and RFEPs). On one hand, research consistently shows that IFEP students tend to perform as well as or even outperform their English-only (EO) counterparts (e.g., Hwang et al., under review; Kieffer, 2011). On the other hand, it is difficult to draw conclusions about how RFEPs perform in comparison to their EO peers because states and school districts have different thresholds and criteria for redesignating their LEP students (Kim & Herman, 2009). In places where the redesignation process is rigorous, RFEP students tend to perform comparably or even better than their EO peers (e.g., Ardasheva, Tretter, & Kinny, 2012; Hwang et al., 2015). However, in schools that have lenient guidelines, RFEP students not only perform poorly but may still need academic support post-redesignation (Slama, 2014).

Academic vocabulary interventions

There have been efforts to enhance adolescent students' general and content-specific academic vocabulary knowledge through research-based interventions (August, Branum-Martin, Cardenas-Hagan, & Francis, 2009; Carlo et al., 2004; Lawrence, Capotosto, Branum-Martin, White, & Snow, 2012; Lesaux, Kieffer, Faller, & Kelley, 2010; Lesaux et al., 2014; Proctor et al., 2011; Snow et al., 2009; Townsend & Collins, 2009; Vaughn et al., 2009). In these studies, principles for designing effective vocabulary intervention included the following (Graves, August, & Mancilla-Martinez, 2013): 1) introducing target words in authentic and engaging texts, 2) provide rich definitions of target words in accessible language, 3) giving a variety of examples to show how target words are used, 4) ensuring opportunities for students to use them in writing and orally, 4) providing repeated exposure, and 5) encourage activities that allow students to use the target words in diverse contexts and help students develop word consciousness. There are few interventions that target adolescent students' literacy outcomes, and even fewer that examine potential differential treatment effects for LM students. However, extant evidence suggests that when a vocabulary intervention is well designed with research-based principles, not only do adolescent EO students but also LM students benefit and show significant gains in their literacy outcomes. In the randomized control trial of *Academic Language Instruction for All Students* (ALIAS), Lesaux and colleagues (2014) showed that an academic vocabulary intervention could work well in schools ($N_{\text{School}} = 14$) that serve linguistically diverse students ($N_{\text{Student}} = 2,082$, approximately 70% were LM students). ALIAS is a relatively intense academic vocabulary intervention that targets 70 high-utility words in a 20-weeks long program with 45-minutes daily lessons. All lessons were implemented in the English language arts (ELA) class and each unit consisted of a 9-day lesson cycle. Each cycle of ALIAS included small-group, whole-group, and individual activities that build on prior knowledge, support understanding multiple meanings and

85

uses of words, and teach morphological analysis. These activities are designed to build upon one another to promote students' word learning. Teachers who participated in this study were provided with teacher-friendly materials as well as professional development for appropriate implementation. Lesaux and colleagues (2014) found significant treatment effects on measures that assess students' academic vocabulary knowledge and writing as well as expository text comprehension that included target words. They also found that the treatment effects were generally larger for LM students and students who participated in the intervention with lower vocabulary knowledge than their EO peers. The findings of this study underscore that high-quality vocabulary intervention can positively affect EO and LM students' word learning as well as other literacy skills that are essential for school success. However, there was no follow-up measure to examine whether these treatment effects maintained or faded out in the following years.

Word Generation. WG is a research-based academic language intervention for middle school students. It intends to explicitly teach five general academic words per week that are chosen from the Academic Word List (Coxhead, 2000). WG consists of 15-minute daily activities that are dispersed across discipline (i.e., ELA, mathematics, social studies, and science). Each unit is a week long, and there are 24 units in a yearly WG curriculum. The WG program is designed to expose students to the target words multiple times in diverse contexts and to encourage students to use the newly learned words orally and in writing. A typical week in a WG curriculum might resemble the following layout (for a sample WG unit, see wordgeneration.org): On Monday, students read a short text about a controversial issue (e.g., *Junk food: Should schools sell it? Is the death penalty justified?)* in an ELA classroom. The five target words are embedded in this text. The ELA teacher introduces these words and helps students understand

what those words mean. On Tuesday, the math teacher does a WG math problem. A word problems is related to the weekly topic, and students use the target words in solving it. On Wednesday, students encounter the target words in their science class. Students are given a scenario and problem sets with target words embedded and asked to think scientifically and interpret scientific data. On Thursday, students have a debate around the weekly controversial topic in a social science class. Each student chooses a position, and their teacher moderates the discussion. During this debate, students are encouraged to use the target words. On Friday, students are asked to write a short persuasive essay about the position that they chose. This writing activity concludes the weekly unit and provides an opportunity for students to use the target words in their writing.

There have been both quasi-experimental (Lawrence et al., 2012; Snow et al., 2009) and randomized control efficacy trials (Hwang et al., 2015; Lawrence, Crosson, Paré-Blagoev, & Snow, 2015; Lawrence, Francis, & Snow, under review), of WG. In one quasi-experimental study that evaluated the WG program, Snow et al. (2009) found positive and meaningful treatment effects of WG in enhancing middle school students' literacy skills. In that study, students in the WG schools showed more growth in their academic vocabulary knowledge than those in the comparison schools. Furthermore, LM students showed greater growth than their EO peers in the treatment schools. Researchers also found that students' improvement on the academic vocabulary test predicted their performance on the state standardized test. In another quasi-experimental study, Lawrence and colleagues (2012) examined short- and long-term effects of the WG program on EO and LM students. They found that students in the treatment condition showed more growth in their academic vocabulary knowledge and they maintained this improvement the fall and spring of the subsequent school year. In addition, the researchers found

that proficient LM students in the treatment condition showed even more growth compared to their EO peers while non-proficient LM students (i.e., LEPs) did not benefit as much from their participation in the WG program.

The randomized control trials of WG also have shown small positive effects on middle school students' academic vocabulary knowledge. In a WG evaluation study that was conducted in 28 schools from two urban school districts with 1,554 middle school students, Lawrence and colleagues (2015) found that WG curriculum promoted classroom discussion quality across four content areas and had a positive effect ($ES = .25$) on enhancing students' academic vocabulary knowledge. In addition, they found that classroom discussion partially mediated the WG treatment effects on students' vocabulary outcomes. In other words, the WG curriculum helped improve classroom discussion, which led to increases in students' academic vocabulary knowledge. Results from the second year of the randomized trial demonstrated that treatment effects may differ according to students' level of English proficiency (Hwang et al., 2015). However, there is currently no follow-up that examines the long-term effects of WG on students' academic vocabulary knowledge. The present study would be a replication of a previous quasi-experimental study (Lawrence et al., 2012) in a randomized control setting in another school district with different student population. The present study investigates the long-term effects of WG on students with diverse linguistic backgrounds. The research question that is addressed in this study is the following: Do EO, IFEP, RFEP, and LEP students learn and maintain academic vocabulary knowledge after receiving the WG instruction?

Methods

Sample

The data for this study come from a three-year randomized efficacy trial of WG. Specifically, four waves of data were collected from a large urban school district in California across two school years. In this school district, thirteen middle schools were randomized to either treatment or control conditions. Before randomization was conducted, schools were ranked on a series of school-level variables, such as percent free and reduced lunch, percent LEPs, and prior mean achievement. Then, schools were ranked on the composite score of these variables. In order to maximize comparability of treatment and control schools, each sequential pair of schools formed a dyad within which randomization occurred. Seven schools were assigned to the treatment condition and six schools were assigned to the control condition. In the second year of the study, the same schools remained in the control condition, and one of the seven schools in the treatment condition withdrew from the study.

Table 4.1 presents the demographic information of participating sixth- and seventh-grade students in both WG and control schools who contributed to at least one wave of data collection. There were more students in the WG schools (n = 2,990) than in the control schools (n = 2,062). The participating school district provided information about students' language status. Students in WG and control schools were identified as either EO, IFEP, RFEP, or LEP based on their district record. While EO students were the biggest language group in the WG schools (38%), the RFEP students comprised the largest group in the control schools (38%). IFEPs were the smallest group in both WG and control schools. Asians were the largest racial group in both WG and control schools (48% and 56%, respectively) followed by the Hispanics (24% in the WG schools and 22% in the control schools). In both conditions, the majority of the students were from socioeconomically disadvantaged backgrounds as indicated by their eligibility for free or reduced lunch (84% in the WG schools and 87% in the control schools). Eleven percent of our

sample was receiving special education and approximately 40% of the sample was indicated as

receiving gifted and talented education.

Table 4.1
Demographic Information of Participants in Each Condition

	WG Schools $n = 2,990$	Control Schools $n = 2,062$
Language Status		
EO	38%	30%
IFEP	11%	9%
RFEP	32%	38%
LEP	19%	22%
Race/Ethnicity		
White	12%	7%
Black	8%	9%
Hispanic	24%	22%
Asian	48%	56%
Other	8%	7%
Eligible for FRL	84%	87%
Special Education	11%	11%
Gifted and Talented Education	43%	42%

Note: FRL = Free and reduced lunch, EO = English-only, IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient. This table describes the demographic information of the participants who contributed to at least one wave of data during the present study.

Procedure

Schools randomized to implement WG adopted the cross-content academic literacy

program into their schools. The efficacy trial of WG was conducted for two years in this district,

and the students in the WG schools were taught 120 words (5 words per week) each year. To

assess students' academic vocabulary knowledge, we administered researcher-developed

academic vocabulary tests four times across two years, once at the beginning of the school year

(September/October) and once at the end of the school year (May). In order to assess and control

for students' general language ability, we also administered vocabulary and reading subtests of

the Gates-MacGinitie Reading Test (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) at the first wave of data collection. At the end of this two-year study, all control schools were offered the opportunity to implement the WG program and receive the relevant professional development. As WG is a free program that is available online, the WG schools could also choose to continue to use the program.

Measures

Academic vocabulary. Students' academic vocabulary (ACA_VOCAB) was assessed with a 50-item multiple-choice test that was developed by the Word Generation research team (all test forms can be found in the IRIS digital depository; www.iris-database.org). For each item, a target word was embedded in a short sentence where students were asked to choose the closest synonym for that word among four answer choices. Each target word had been found to be relatively frequent and dispersed across a range of first year college texts (Coxhead, 2000).

Since two different sets of words were taught each year, two forms of the Academic Vocabulary test were used: Form A targeted words taught in Year 1, and form B targeted words taught in Year 2. To assess the (longitudinal) effect of the WG program on students' academic vocabulary knowledge, we also included 11 anchor items that were taught in the first year, and tested in both Year 1 and Year 2. The 11 anchor target words were *acquire, interaction, enforce, relevant, paralyze, incentive, enable, contrast, generate, obtain,* and *recite*. These 11 words were tested in every wave during the two-year period. In order to maximize comparability across scores from four data points, we conducted an item response theory (IRT) analysis. We fit a single-factor 3-parameter model to these 11 items. The unidimensional 3-parameter IRT model fit the data well in all four waves (RMSEA = .00 for all four waves). The marginal reliability ranged from .72 to .77. The item parameters that were obtained from the first wave were used to

produce scaled scores for the subsequent waves. The scoring method we used was expected *a posteriori*. These scaled scores were used as a level-1 outcome variable in our analysis. The mean and standard deviation at the first wave were -.09 and .80, respectively. Table 4.2 displays the number of students who had valid ACA_VOCAB scores at each wave of data collection by their treatment conditions. Although a greater number of students participated in the study, our analytical sample only included students who had valid responses in the 11 anchor words.

Table 4.2
Number of Students who Had Valid Academic Vocabulary Scores at Each Wave of Data Collection in Word Generation and Control Schools

WG Schools	*n*
Fall Year 1	2,904
Spring Year 1	2,434
Fall Year 2	2,362
Spring Year 2	2,081
Control Schools	
Fall Year 1	1,959
Spring Year 1	1,782
Fall Year 2	1,676
Spring Year 2	1,210

General vocabulary. In order to control for students' general vocabulary knowledge (VOCAB) in our analytical models, we used students' extended scale scores from the vocabulary subtest of the Gates-MacGinitie Reading Test (MacGinitie et al., 2000). Gates-MacGinitie Reading Test is a group-administered assessment that includes reading and vocabulary subtests. The level 6 Form T was administered to sixth-grade students and the level 7/9 Form T was given to seventh-grade students. The vocabulary subtest consists of 45 multiple-choice items that ask students to choose the synonym of a target word. The mean and the standard deviation of the general vocabulary assessment in our sample were 516.71 and 37.30, respectively. The Cronbach's alpha for our sample at the fall of Year 1 was .90.

Reading comprehension. In order to control students' overall reading ability (READ), we used students' extended scale scores from the reading subtest of the Gates-MacGinitie Reading Test (MacGinitie et al., 2000). This 48-item subtest measures students' reading comprehension skills; students are asked to read a passage and answer relevant comprehension questions. Extended scale scores of the reading subtest were used in our analysis. The mean and standard deviation of our sample on reading comprehension were 518.40 and 37.76, respectively. The Cronbach's alpha for our sample at the fall of Year 1 was .91.

Time. TIME is a level-1 variable indicating the time since the start of the study when students took the assessments. We coded each wave in months (i.e., wave 1 = 0 month, wave 2 = 7 months, wave 3 = 12 months, wave 4 = 19 months).

Instruction. INSTRUCTION is a time-varying individual (level-1) variable that indicates how many instructional encounters students had with the 11 target words. Students in the treatment condition were taught the target words during Year 1 but not during Year 2 (i.e., wave 1 = 0, wave 2 = 1, wave 3 = 1, wave 4 = 1). Students in the control condition were coded as 0 for all four waves.

Summer. The number of summers (SUMMER) students experienced since the start of the study was also included in our analysis (i.e., wave 1 = 0, wave 2 = 0, wave 3 = 1, wave 4 = 1). SUMMER is a time-varying continuous level-1 variable, and its parameter indicates whether or not students experience summer setback with their academic vocabulary knowledge.

WG School. WG is a time-invariant student-level (level-2) dummy variable indicating whether students attended a WG school (WG = 1) or not (WG = 0) in the duration of this current study.

Language status. Dummy variables for EO, IFEP, RFEP, and LEP students were created to be used as student-level predictors in our analysis. Although some of the LEP students' status changed over time (i.e., 425 LEPs in Year 1 were reclassified as RFEP in Year 2), this variable is a time-invariant initial classification of students' language status at the beginning of this current study. EO students were the reference group in our analysis.

Grade-level cohort. To control for different grade levels in the analysis, student-level dummy variables were created for grade six (GRADE_6) and seven (GRADE_7). The sixth grade students were the reference group in the analysis.

Ethnicity. Three student-level dummy variables (ASIAN, WHITE, BLACK, HISPANIC, and OTHER) were created to indicate students' ethnicity. Asian students were used as the reference group in our analysis because they were by far the largest group (52% of the sample).

Socioeconomic status (SES). Eligibility for receiving free or reduced lunch was used as an indicator of students' SES. A student-level dummy variable was created to indicate students who receive free or reduced lunch (SES = 1) and those who do not (SES = 0).

Special education status. A student-level dummy variable was created to indicate students who were receiving special education (SPED = 1) and those who were not (SPED = 0).

Gifted and talented education status. A student-level dummy variable was created to indicate students who were receiving gifted and talented education (GATE = 1) and those who were not (GATE = 0).

Data Analysis

To answer our research questions, we conducted a series of multilevel models for change (Singer & Willet, 2003). These models allow us to use all waves of data from each student to create a model of vocabulary growth over two years to examine potential treatment effects in the

first year and possible summer setback during the summer break. The data were prepared in a person-period dataset where each student had up to four rows of data. The hypothesized multilevel model for change for the first research question was:

Level-1 (outcomes in four waves across two years):

$$ACA\_\widehat{VOCAB} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}TIME^2_{ij} + \pi_{3i}INSTRUCTION_{ij} +$$

$$\pi_{4i}SUMMER_{ij} + \varepsilon_{ij} \tag{1}$$

Level-2 (student-level):

$$\pi_{0i} = \gamma_{00} + \gamma_{01}IFEP + \gamma_{02}RFEP_i + \gamma_{03}LEP + \gamma_{04}VOCAB_i + \gamma_{05}READ_i + \gamma_{06}WG_i +$$

$$\gamma_{07}GRADE7_i + \gamma_{08}WHITE_i + \gamma_{09}BLACK_i + \gamma_{10}HISPANIC_i + \gamma_{11}OTHER_i + \gamma_{12}SES_i +$$

$$\gamma_{13}SPED_i + \gamma_{14}GATE_i + \zeta_{0i} \tag{2}$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11}IFEP_i + \gamma_{12}RFEP_i + \gamma_{13}LEP_i + \zeta_{1i} \tag{3}$$

$$\pi_{2i} = \gamma_{20} \tag{4}$$

$$\pi_{3i} = \gamma_{30} + \gamma_{31}IFEP_i + \gamma_{32}RFEP_i + \gamma_{33}LEP_i \tag{5}$$

$$\pi_{4i} = \gamma_{40} + \gamma_{41}IFEP_i + \gamma_{42}RFEP_i + \gamma_{43}LEP_i \tag{6}$$

where $\varepsilon_{ij} \sim N(0, \sigma_1{}^2)$, and $\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$.

The coefficient γ_{00} represents the average score for sixth grade, Asian, EO students in the control condition at the first wave (the first measurement point); γ_{10} represents the average initial slope for EO students; γ_{20} represents the average true acceleration for EO students; and γ_{40} represents the average summer setback (or gain) for EO students. The random effect ε_{ij} is a level 1 residual for student i at time j and is assumed to be drawn from a normal distribution with mean of 0 and variance $\sigma_1{}^2$. Random effects ζ_{0i} and ζ_{1i} represent Level 2 residuals for the

intercept and slope, respectively. They are both hypothesized to be drawn from a multivariate normal distribution with a mean of zero, unknown variances σ_0^2 and σ_1^2, and unknown covariance σ_{01}. The coefficient γ_{06} refers to the baseline difference in academic vocabulary test scores between students in the treatment and control conditions. The coefficient γ_{30} represents the difference in academic vocabulary test scores between EO students in the WG and control conditions at the second wave of data collection. Parameter estimates specific to each language proficiency group (i.e., IFEPs: γ_{31}; RFEPs: γ_{32}; LEPs: γ_{33}) were compared to the reference group parameters (i.e., EO: γ_{30}) to answer our research question. Different interaction terms were created and tested in the analytical models to examine whether students in the WG schools had different growth trajectories compared to those in the control schools, and whether the treatment effect faded away as time passed.

Results

Table 4.3 provides means and standard deviations of the 11 target items by students' language status and their treatment status across two years of data collection. The first four columns present the IRT scaled academic vocabulary scores by language group, and the last four columns show average raw scores for each group. We can see that students in both WG and control schools improved in their academic vocabulary knowledge during the two-year period of this study. This table suggests that LEPs experience greater summer setback than their EO, IFEP, and RFEP peers in both conditions. LEPs had the lowest scores across four waves of data and IFEP students had the highest scores in both conditions. At baseline, students in the WG schools had higher scaled scores ($M = -.06$) than students in the control schools ($M = -.15$). However, students in the treatment condition improved more across the two-year period than students in control schools (WG schools: $\Delta = .64$; Control schools: $\Delta = .53$).

Table 4.3
Means and Standard Deviations of Scaled and Raw Scores of Eleven Longitudinal Items by Treatment Condition

		Scaled				Raw			
		Year 1		Year 2		Year 1		Year 2	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
WG Schools	EO	.12 (.88)	.40 (.94)	.53 (.99)	.75 (1.00)	6.16 (2.47)	6.97 (2.69)	7.39 (2.70)	7.84 (2.81)
	IFEP	.22 (.77)	.58 (.80)	.68 (.86)	.82. (.88)	6.39 (2.22)	7.46 (2.22)	7.79 (2.33)	8.13 (2.29)
	RFEP	.02 (.67)	.35 (.70)	.50 (.75)	.69 (.80)	5.79 (2.04)	6.78 (2.11)	7.25 (2.12)	7.76 (2.20)
	LEP	-.69 (.60)	-.48 (.70)	-.45 (.65)	-.26 (.81)	3.94 (1.79)	4.50 (2.03)	4.64 (1.96)	5.12 (2.33)
	All Students	-.06 (.81)	.24 (.88)	.36 (.92)	.58 (.96)	5.64 (2.35)	6.51 (2.53)	6.90 (2.57)	7.41 (2.67)
Control Schools	EO	-.04 (.87)	.21 (.89)	.25 (.94)	.45 (1.01)	5.68 (2.53)	6.46 (2.48)	6.62 (2.64)	6.94 (3.00)
	IFEP	.16 (.72)	.41 (.75)	.55 (.80)	.66 (.84)	6.20 (2.08)	6.97 (2.07)	7.41 (2.17)	7.41 (2.62)
	RFEP	.02 (.66)	.28 (.67)	.38 (.72)	.63 (.73)	5.77 (1.97)	6.63 (1.89)	6.89 (2.01)	7.54 (2.22)
	LEP	-.71 (.58)	-.52 (.64)	-.53 (.65)	-.34 (.76)	3.80 (1.86)	4.33 (1.93)	4.37 (2.04)	4.83 (2.39)
	All Students	-.15 (.78)	.11 (.81)	.16 (.87)	.38 (.92)	5.34 (2.30)	6.13 (2.30)	6.30 (2.47)	6.79 (2.74)

Note: Standard deviations in parentheses. EO = English-only, IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient.

Table 4.4
Multilevel Models for Change Predicting Academic Vocabulary for Students with Different
Language Status

		Model A	Model B	Model C
		Unconditional	With Covariates	With Interactions
Fixed Effects	Intercept	-.08***	-8.32***	-8.31***
		(.011)	(.133)	(.133)
	IFEP		.01	-.01
			(.023)	(.027)
	RFEP		-.01	-.06**
			(.018)	(.020)
	LEP		-.09***	-.06*
			(.021)	(.024)
	TIME	.03***	.03***	.03***
			(.003)	(.003)
	TIME X IFEP			.01
				(.004)
	TIME X RFEP			.01***
				(.003)
	TIME X LEP			-.001
				(.003)
	TIME2		-.0002*	-.0003*
			(.0001)	(.000)
	SUMMER		-.03*	.02
			(.015)	(.025)
	SUMMER X IFEP			-.05
				(.050)
	SUMMER X RFEP			-.06
				(.034)
	SUMMER X LEP			-.10*
				(.041)
	INSTRUCTION		.06***	.07***
			(.017)	(.017)
	WG SCHOOL		.03*	.03
			(.016)	(.016)
	READ		.01***	.01***
			(.0002)	(.0003)
	VOCAB		.01***	.01***
			(.0003)	(.0003)
	GRADE 7		-.03*	-.03*
			(.013)	(.013)
	WHITE		.04	.04
			(.023)	(.023)
	BLACK		-.14***	-.14***
			(.026)	(.026)
	HISPANIC		-.02	-.02
			(.016)	(.016)
	OTHER		-.03	-.03

			(.024)	(.024)
	SES		-.02	-.02
			(.018)	(.018)
	SPED		-.04*	-.04*
			(.021)	(.021)
	GATE		.16***	.16***
			(.015)	(.015)
Level 1 Variance Component	Residual	.21***	.20***	.20***
		(.003)	(.003)	(.003)
Level 2 Variance Component	Intercept	.46***	.08***	.08***
		(01)	(.01)	(.006)
	TIME	.0004***	.0003***	.0003***
		(.000004)	(.00003)	(.00004)
	Covariance	.004***	.0004	.001
		(.001)	(.0004)	(.0004)
	N (Students)	5,052	5,052	5,052
	N (Observations)	16,408	16,408	16,408

Note. *$p < .05$; **$p < .01$; *** $p < .001$; The reference group in the analysis was English-only student who is in 6th grade cohort, Asian, and does not receive free and or reduced lunch, and not receiving either special or gifted and talented education. IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient, SES = socioeconomic status, SPED = special education status, GATE = gifted and talented education status

Table 4.4 displays the results from fitting a series of multilevel models for change predicting ACA_VOCAB across four waves of data. The inclusion of the quadratic term improved the model fit (Δ-2LL = 5.78, df = 1, p < .05); its negative value (β = -.0003, p = .016) indicates that learning rates decrease as students get older. Model A in Table 4.4 shows the results from an unconditional growth model, which was a baseline model for our analysis. Model B displays results from a growth model with predictors and covariates. Including the predictors and covariates improved the model fit (Δ-2LL = 6006.18, df = 17, p < .05), and 84% of the between-person variation was associated with these predictors and covariates. Model C in Table 4.4 shows results from our final fitted model. This model shows that EO students improved in their academic vocabulary on average (β = .03, p < .001). IFEP and LEP students learned new

words at the same rate as EOs when controlling for covariates. However, the TIME X RFEP interaction term was positive and significant ($\beta = .01$, $p < .001$), indicating that RFEP students had more rapid growth than their EO peers on average. In this sample, EO, IFEP, and RFEP students did experience "summer setback" in their academic vocabulary knowledge. Notably, the SUMMER X LEP interaction term was negative and significant ($\beta = -.10$, $p = .019$) indicating that LEPs' academic vocabulary learning rates decreased during the summer. The coefficient for INSTRUCTION was positive and significant ($\beta = .07$, $p < .001$). This means that students who received Word Generation instruction scored significantly higher than those who did not. INSTRUCTION did not interact with language status; all students benefited from instruction similarly. We tested whether students in the WG schools experience differential growth or summer setback by creating WG X TIME and WG X SUMMER interaction terms; neither was statistically significant. We also tested for a TIME X INSTRUCTION interaction to examine whether vocabulary growth varied between groups post instruction (i.e., if there was a treatment fade out). However, this interaction term was not statistically significant.

Figure 4.1 presents prototypical academic vocabulary growth trajectories for students who start the study in sixth grade. The bold lines model prototypical students attending WG schools and thin lines indicate students attending control schools. At the first wave of data collection, IFEP (short-dashed lines with square markers) and EO (straight lines with circle markers) students were the highest performing groups, followed by the RFEP (long-dashed lines with triangle markers) students. LEPs (long-dashed lines with two dots) had the lowest vocabulary knowledge to start with. Looking across the groups, we can confirm that students showed growth in academic vocabulary on average over this two-year period.
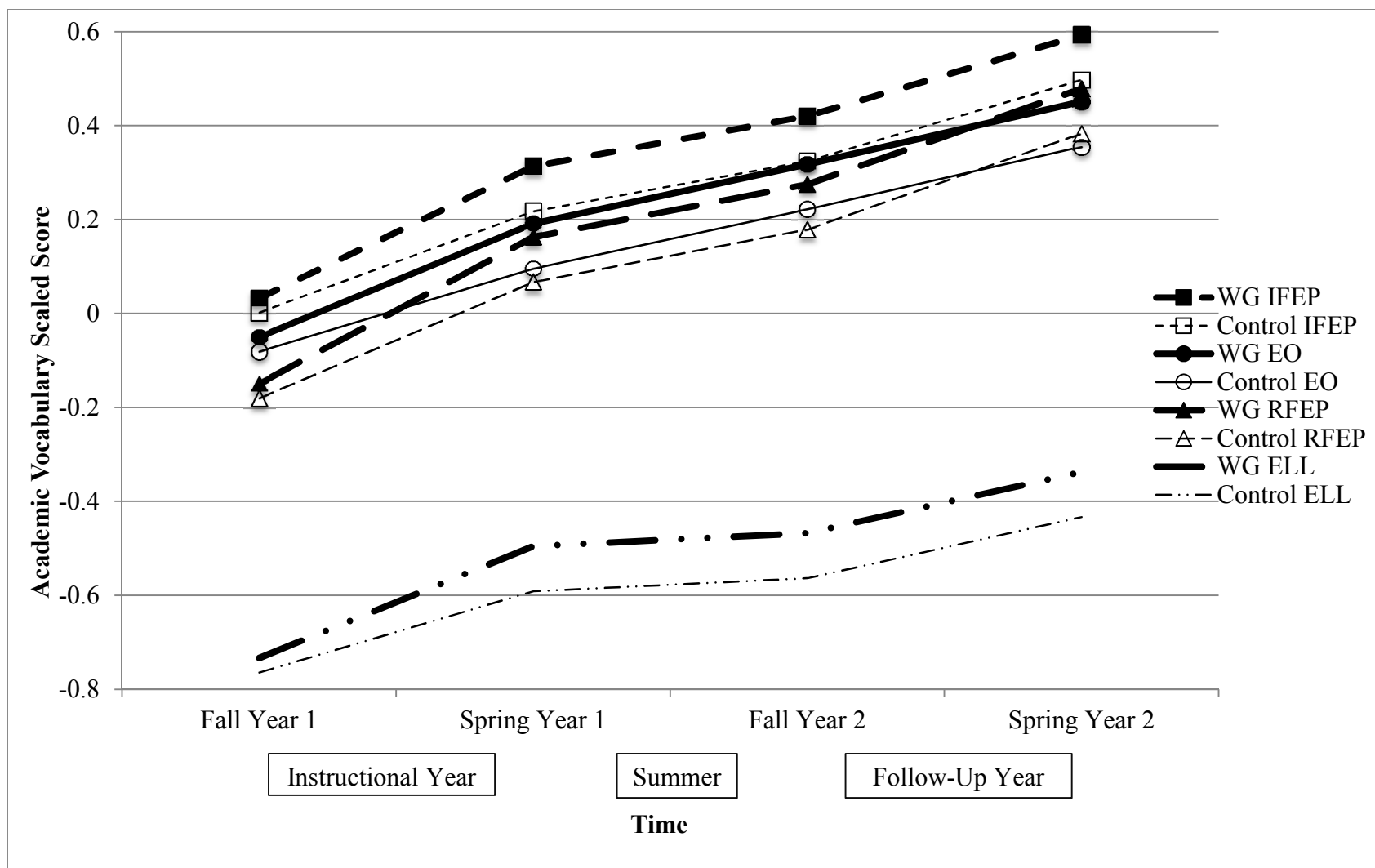
Figure 4.1. Prototypical plot of sixth-grade students in Word Generation and control schools by their language status

Students in the WG schools showed additional growth during the first year (i.e., the instructional year) compared to peers in the control condition. They maintained that relative improvement during the summer and the next school year. Our final analytical model indicated that RFEP students had more expedited growth over time compared to their EO peers. As a result, the RFEPs who joined the study in sixth grade in both conditions were predicted to outperform their EO peers by the end of seventh grade when other variables were controlled for. Moreover, the RFEPs in the WG schools were predicted to catch up to the IFEPs in the control condition by the end of the follow-up year. Due to the negative and significant LEP X SUMMER interaction, LEP students in both treatment conditions showed learning curves that were almost flat during the summer months. This pattern was not found in other language groups. Although LEPs in the treatment condition showed more growth over two years than their peers in the control condition, this was not enough to catch up to their non-LEP peers in both treatment and control conditions.

Discussion

This is the first longitudinal study conducted on the WG program in a randomized control setting that examined short- and long-term effects of WG on students with differing levels of English proficiency. By employing individual growth modeling, we were able to simultaneously investigate how subgroups of LM students perform in comparison to their EO peers and how they respond to the WG instruction both immediately after the program implementation and at the follow-up year. In some ways, the findings from this study are quite consistent with previous studies on the WG program. Students with different language designations showed differential growth trajectories over two years and students in the WG schools improved more rapidly during

the instructional year than student in the control schools. Furthermore, in the follow-up year,

students in the WG schools maintained that relative improvement.

Short- and long-term effects of WG

We found both short- and long-term effects of WG on middle school students' academic

vocabulary knowledge. During the first year of the study (i.e., instructional year), students

attending WG schools showed more rapid growth than their peers in the control schools. This

finding is consistent with previous experimental and quasi-experimental evaluations of WG.

However, we did not find the heterogeneous treatment effects of WG across student by language

status in the current study. In the current study, *all* students benefited from program participation

equally; previous studies have demonstrated differential treatment effects favoring LM learners

(Hwang et al., 2015; Lawrence et al., 2012)[13]. These differences might be related to differences

in the sample. The study by Lawrence et al. (2012) was conducted in a school district in

Massachusetts and the current study was conducted in California. How each state identifies and

---

[13] In one of our previous analyses we found that only RFEP students benefited from the program (Hwang et al., 2015). However, in the current study, we estimated a main effect of treatment and we did not find the treatment effect moderated by students' language status. When we tested for statistical models constraining all students to have equal rate of growth and same amount of summer setback, we did find that the treatment effect was moderated by students' language status. However, we believed that EO and LM students would have differential academic vocabulary growth trajectories based on our previous work (c.f., Hwang et al., under review) that we decided on the current final model. The goodness-of-fit statistics also indicated that our current final model better fit the data. Such difference seems to be due to different analytical models that were used in these two studies. In Hwang et al. (2015), multilevel modeling where students were nested in schools (i.e., hierarchical linear modeling) was conducted. This model used academic vocabulary and reading comprehension scores as covariates at the school and student level. This analytical approach was designed to explore treatment by language status interactions and was conservative to determining overall treatment effects. In fact, was say "this study is not powered to detect a treatment effect with data from this small number of schools" (p. 329). In contrast, the current study uses individual growth modeling where individuals are *not* nested within schools and grade levels. We ran models that included school dummy variables to account for students attending different schools. The models with and without school dummy variables gave similar results. This approach allows for the flexible use of data, and we were able to include students who contributed at least on wave of data in the duration of this study. More importantly, we are able to describe average trajectories across four waves of data. Overall, the finding of a main treatment effect is consistent with the majority of the other studies and the efficacy analysis from the second year of the randomized trial that uses data from all of the school sites (Lawrence et al, under review).

categorizes LM learners varies to a great extent (Abedi, 2008; Ragan & Lesaux, 2006). California uses a relative complex set of criteria for identifying and reclassifying LM students whereas Massachusetts relies on one test that is focused on English language proficiency. More longitudinal studies that examines follow-up treatment effects of a particular intervention and LM students' growth trajectories in different states are needed to understand differential treatment effects across students' language proficiency.

In the second year (i.e., follow-up year), we found that students in the WG schools maintained relative improvement in their academic vocabulary knowledge. This finding is consistent with the results from the quasi-experimental study of WG (Lawrence et al., 2012). Our findings suggest that WG may work well in schools that serve students with diverse English proficiency levels with relatively small amount of time (approximately 15 minutes per day). Because WG is a program that focused on promoting students' academic vocabulary knowledge across content areas, it provided opportunities for students to use target words in diverse settings (e.g., in writing, during discussion). These various opportunities to practice using academic vocabulary could have eventually led students in the treatment condition to not only learn the target words in the short-term but also to maintain that knowledge in the following years.

Academic vocabulary trajectories

Because we employed individual growth modeling approach, we were also able to simultaneously examine student's underlying academic vocabulary growth trajectories along with treatment effects. LEPs were the lowest performing group at the beginning of our study, and they remained so through the end of the study. There are differences in the average academic vocabulary knowledge of LEP and non-LEP (i.e., EO, IFEP, RFEP) students (August, Carlo, Dressler, & Snow, 2005; Kieffer, 2008; Mancilla-Martinez & Lesaux, 2011). The finding from

the current study underscores the fact that students' initial English proficiency is associated with their later reading-related outcomes (in this case, academic vocabulary knowledge) and that dramatically changing their learning trajectories may not be feasible with regular business-as-usual instruction (Kieffer, 2011; Mancilla-Martinez, Kieffer, Biancarosa, Christodoulou, & Snow, 2011; Reese et al., 2000). On one hand, LEPs in the treatment schools showed more growth in their academic vocabulary than their peers in the control schools in the first year of this study. Furthermore, they maintained these relative improvements even a year after instruction had ended. These finding indicate that LEPs benefit from explicit vocabulary instruction and can learn quickly during the school year. On the other hand, LEPs' learning trajectory during the summer slowed dramatically compared to their peers. Their loss during the summer was equivalent to approximately three months of school year learning. We did not have data on student extracurricular activities during the summer. However, these finding suggest that summer could be a particularly important period for supporting LEP academic vocabulary and literacy skills. Researchers have found that factors such as decreased access to books and less time spent on voluntary reading may lead to summer loss, and the summer setback may be greater for students from low SES, ethnic minority and/or LM homes (Alexander, Entwisle, & Olson, 2001, 2007; Heyns, 1978; Lawrence, 2012). Research suggests that well-designed summer reading programs may be effective for decreasing this summer loss in reading and providing more reading opportunities for students (Kim, 2006, 2007; Kim & Guryan, 2010; Kim & White, 2008). High-quality reading programs that incorporate explicit vocabulary instruction may be needed to accelerate LEP students' vocabulary learning during the summer.

Other LM students performed comparably or even better than their EO peers in academic vocabulary. Consistent with previous research findings, IFEPs students were among the highest-

performing groups in both conditions (Hwang et al., 2015; Hwang et al., under review; Kieffer, 2011). Their rate of growth and acceleration were similar to those of their EO peers and the effect of summer on their academic vocabulary knowledge was also the same compared to their EO peers. Their average general vocabulary and reading comprehension scores, which were included as covariates in the final model, were also the highest among the language groups. With the data that we have, we cannot fully explain why IFEPs are at an advantage with literacy-related outcomes. However, it highlights that students' English proficiency at school entry can be an important predictor in their later reading-related outcomes and that there needs to be much support in the early grades for students from LM homes to gain sufficient English proficiency as early as possible. It was somewhat encouraging to find that even the high-performing students in our sample benefited from WG intervention. IFEP students in the treatment condition showed more growth over time compared to their peers in the control condition in the first year of the study and they were able to maintain that gap in the follow-up year.

At baseline, the RFEP students performed slightly lower than their EO peers. However, they showed steeper learning trajectories over two-year period that they eventually caught up to their EO peers by the end of the study. Such growth was not sufficient to close the gap between RFEP and IFEP students. There was no significant interaction between student RFEP status and Summer in the current study. In another study RFEP students showed more rapid growth over time, their learning trajectories may slow down during the summer compared to their EO peers (Hwang et al., under review). This may indicate that these former LEP students may also benefit from summer activities that allow them to be exposed to new English words that appear frequently in secondary texts.

Limitation and Future Directions

In interpreting the results of this study, it is important to note that we only used one type of assessment (i.e., choosing a synonym to a target word) to measure students' academic vocabulary knowledge. This test was easy to administer to students in a whole-group setting with less burden on participating teachers and students. However, using multiple forms of assessments that tap on different kinds of vocabulary knowledge could have produced richer and more accurate measure of students' vocabulary knowledge. Furthermore, we only selected students' responses in 11 items in our analysis. While this procedure allowed us to examine their growth and maintenance of academic vocabulary knowledge over time, it considerably reduced the amount of data and variability within students' responses. This was the best and cleanest way for us to examine the longitudinal treatment effects of WG, but future research studies that aim to examine the longitudinal treatment effects should consider including more items to assess students' academic vocabulary knowledge more reliably. Another thing to note is that the current study was conducted in one school district in California. As mentioned earlier, California holds a relatively high bar for classifying its LM students. Thus, the current findings may not generalize to other states that have different standards for identifying and classifying LM students. Despite the limitations, the current study contributes to the research literature by examining the long-term effects of a research-based academic vocabulary intervention on adolescent EO and LM students. There is still limited amount of intervention research conducted on adolescent LM students that examine short- and long-term treatment effects. Future research studies should continue to examine these issues so that educators can provide appropriate instructional support for their students.

References

Abedi, J. (2008). Classification system for English language learners: Issues and

    recommendations. *Educational Measurement: Issues and Practice, 27*(3), 17-31.

    doi:10.1111/j.1745-3992.2008.00125.x

Alexander, K., Entwisle, D., & Olson, L. (2001). Schools, achievement, and inequality: A

    seasonal perspective. *Educational Evaluation & Policy Analysis, 23*(2), 171-191.

    doi:10.3102/01623737023002171

Alexander, K., Entwisle, D., & Olson, L. (2007). Lasting consequences of the summer learning

    gap. *American Sociological Review, 72*(2), 167-180. doi:10.1177/000312240707200202

Ardasheva, Y., Tretter, T. R., & Kinny, M. (2012). English language learners and academic

    achievement: Revisiting the threshold hypothesis. *Language Learning, 62*(3), 769-812.

    doi:10.1111/j.1467-9922.2011.00652.x

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an

    instructional intervention on the science and language learning of middle grade English

    language learners. *Journal of Research on Educational Effectiveness, 2*(4), 345-376.

    doi:10.1080/19345740903217623

August, D., Carlo, M., Dressler, C., & Snow, C. E. (2005). The critical role of vocabulary

    development for English language learners. *Learning Disabilities Research & Practice,*

    20(1), 50-57. doi:10.1111/j.1540-5826.2005.00120.x

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners:*

    Report of the national literacy panel on language minority children and youth. Mahwah,

    NJ: Erlbaum.

Bailey, A. (2007). *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford Press.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., . . . White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188-215. doi:10.1598/RRQ.39.2.3

Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language, 15*(2), 103-116.

Common Core State Standards Initiative. (2010). Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects. Washington, DC: Authors.

Corson, D. (1997). The learning and use of academic English words. *Language Learning, 47*(4), 671-718. doi:10.1111/0023-8333.00025

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

Graves, M. F., August, D., & Mancilla-Martinez, J. (2013). *Teaching vocabulary to English language learners*. New York: Teachers College Press.

Heyns, B. (1978). *Summer learning and the effects of schooling*. New York, NY: Academic Press.

Hwang, J. K., Lawrence, J. F., Mo, E., & Snow, C. E. (2015). Differential effects of a systematic vocabulary intervention on adolescent language minority students with varying levels of

English proficiency. *International Journal of Bilingualism, 19*(3), 314-332. doi:10.1177/1367006914521698

Hwang, J. K., Lawrence, J. F., & Snow, C. E. (under review). Differential vocabulary growth trajectories among adolescent language minority students: A two-year longitudinal study.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100*(4), 851-868. doi:10.1037/0022-0663.100.4.851

Kieffer, M. J. (2011). Converging trajectories reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal, 48*(5), 1187-1225. doi:10.3102/0002831211419490

Kim, J., & Herman, J. L. (2009). A three-state study of English learner progress. *Educational Assessment, 14*(3-4), 212-231. doi:10.1080/10627190903422831

Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis, 28*(4), 335-355. doi:10.3102/01623737028004335

Kim, J. S. (2007). The effects of a voluntary summer reading intervention on reading activities and reading achievement. *Journal of Educational Psychology, 99*(3), 505-515. doi:10.1037/0022-0663.99.3.505

Kim, J. S., & Guryan, J. (2010). The efficacy of a voluntary summer book reading intervention for low-income Latino children from language minority families. *Journal of Educational Psychology, 102*(1), 20.

Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3

    to 5: An experimental study. *Scientific Studies of Reading, 12*(1), 1-23.

    doi:10.1080/10888430701746849

Lawrence, J. F. (2012). English vocabulary trajectories of students whose parents speak a

    language other than English: Steep trajectories and sharp summer setback. *Reading and*

    Writing, 25(5), 1113-1141. doi:10.1007/s11145-011-9305-z

Lawrence, J. F., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. E. (2012). Language

    proficiency, home-language status, and English vocabulary development: A longitudinal

    follow-up of the Word Generation program. *Bilingualism: Language and Cognition,*

    15(3), 437-451. doi:10.1017/S1366728911000393

Lawrence, J. F., Crosson, A., Paré-Blagoev, E. J., & Snow, C. E. (2015). Word Generation

    randomized trial: Discussion mediates the impact of program treatment on academic

    word learning. *American Educational Research Journal, 52*(4), 750-786.

    doi:10.3102/0002831215579485

Lawrence, J. F., Francis, D., & Snow, C. E. (under review). The poor get richer: Effectiveness of

    a school-level intervention for academic language in ameliorating the consequences of

    low initial achievement.

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of

    implementation of an academic vocabulary intervention for linguistically diverse students

    in urban middle schools. *Reading Research Quarterly, 45*(2), 196-228.

    doi:10.1598/RRQ.45.2.3

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic

    vocabulary Instruction for linguistically diverse adolescents evidence from a randomized

field trial. *American Educational Research Journal, 51*(6), 1159-1194. doi:10.3102/0002831214532165

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading test technical report: Forms S and T*. Chicago, IL: The Riverside Publishing Company.

Mancilla-Martinez, J., Kieffer, M. J., Biancarosa, G., Christodoulou, J. A., & Snow, C. E. (2011). Investigating English reading comprehension growth in adolescent language minority learners: Some insights from the simple view. *Reading and Writing, 24*(3), 339-354.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*(5), 1544-1560. doi:10.1111/j.1467-8624.2011.01633.x

Murnane, R., Sawhill, I., & Snow, C. (2012). Literacy challenges for the twenty-first century: Introducing the issue. *The Future of Children, 22*(2), 3-15. doi:10.1353/foc.2012.0013

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

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing, 20*(7), 691-719. doi:10.1007/s11145-006-9045-7

National Center for Education Statistics. (2011). *The nation's report card: Reading 2011 (NCES 2012-457)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2012a). *The nation's report card: Vocabulary results from 2009 and 2011 NAEP reading assessments (NCES 2013-452).* Washington, D.C: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2012b). *The nation's report card: Writing 2011 (NCES 2012-470).* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C. E., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing, 24*(5), 517-544. doi:10.1007/s11145-009-9218-2

Ragan, A., & Lesaux, N. K. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives, 14*(20), 1-32. doi:10.14507/epaa.v14n20.2006

Reese, L., Garnier, H., Gallimore, R., & Goldenberg, C. (2000). Longitudinal analysis of the antecedents of emergent Spanish literacy and middle-school English reading achievement of Spanish-speaking students. *American Educational Research Journal, 37*(3), 633-662.

Scarcella, R. C. (2003). *Academic English: A conceptual framework.* Berkeley, CA: The University of California Linguistic Minority Research Institute.

Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education, 12*(4), 431-459.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective.* Mahwah, NJ: Lawrence Erlbaum.

Short, D., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners: A report to Carnegie Corporation of New York*. Washington, D.C.: Alliance for Excellent Education.

Singer, J., & Willet, J. (2003). *Applied longitudinal data analysis: Modeling change and even occurrence*. New York, NY: Oxford University Press.

Slama, R. B. (2012). A longitudinal analysis of academic English proficiency outcomes for adolescent English language learners in the United States. *Journal of Educational Psychology, 104*(2), 265-285. doi:10.1037/a0025861

Slama, R. B. (2014). Investigating whether and when English learners are reclassified into mainstream classrooms in the United States: A discrete-time survival analysis. *American Educational Research Journal, 51*(2), 220-252. doi:10.3102/0002831214528277

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2*(4), 325-344. doi:10.1080/19345740903167042

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). New York, NY: Cambridge University Press.

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: An intervention study. *Reading and Writing, 22*(9), 993-1019.

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross☐disciplinary academic☐language proficiency and its

association with reading comprehension. *Reading Research Quarterly*.

doi:10.1002/rrq.104

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis,

D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade

English language learners: Findings from two experimental studies. *Journal of Research*

on Educational Effectiveness, 2(4), 297-324.

CHAPTER 5

Summary and Conclusion

There are great numbers of language minority students in the United States (Kena et al., 2014), and it is critical that educators and policymakers are well aware of their English literacy developmental trajectories. My dissertation aims to carefully examine vocabulary growth trajectories of adolescent language minority students with varying levels of English proficiency and investigate how subgroups of language minority students respond to an academic vocabulary intervention during the instructional and follow-up year. To address these goals, I analyzed information on students' language proficiency status and their scores on vocabulary and reading comprehension measures. Below I summarize the results from three studies. Following the summary of findings, I discuss overarching themes that emerged within and across studies. Finally, I close with a discussion of directions for future research.

Summary of Findings

In Study 1 (Differential vocabulary growth trajectories among adolescent language minority students: A two-year longitudinal study), I investigated general vocabulary and academic vocabulary growth trajectories of sixth- to eighth-grade English-only (EO) and language minority students ($N = 3,161$) using an individual growth modeling approach. The language minority student sample in this analysis included initially fluent English proficient (IFEP), redesignated fluent English proficient (RFEP), and limited English proficient (LEP) students from a large urban school district in California. Students were assessed at four time points on a standardized measure of general vocabulary (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) and a researcher-developed measure of academic vocabulary. In regards to general vocabulary, IFEP students outperformed their peers on average in all time points. The average

baseline score of EO students was higher than that of the RFEP students, and LEP students were the lowest-performing group in this sample. EO, IFEP, and RFEP students showed similar rates of growth and amounts of summer setback in general vocabulary. LEPs improved more slowly compared to their EO peers in general vocabulary knowledge during the school year, but continued to learn during the summer. In terms of academic vocabulary growth trajectories, IFEPs were again the highest-performing group in this sample. All subgroups of language minority students had steeper academic vocabulary growth trajectories than their EO peers. In academic vocabulary, only RFEP students experienced loss in their academic vocabulary knowledge during the summer months. Even so, because RFEP students had steeper learning trajectories during the school year, they were predicted to catch up to their EO peers by the end of the second year of this study. The findings of this study suggest subgroups of language minority students experience differential learning trajectories in vocabulary and they may have different mechanisms for word learning in their middle school years.

In Study 2 (Vocabulary and reading performances of redesignated fluent English proficient students), I examined academic vocabulary, general vocabulary, and reading comprehension growth trajectories of sixth- to eighth-grade RFEP students using individual growth modeling analysis. The sample included 1,226 RFEP students from six middle schools in an urban school district in California. RFEP students completed up to four waves of reading-related measures during a two-year period. The results indicate the RFEP students in this study were performing comparably or even better than the students in the national norming sample on average. Additionally, RFEP students' scores on vocabulary and reading assessments were positively correlated with their years since redesignation. Moreover, students on average showed growth over time on all three outcomes of interest, and the rate of growth did not differ by their

years since redesignation. In other words, RFEP students who were redesignated early in their schooling career tended to outperform those who were recently redesignated.

In Study 3 (Investigating the effects of Word Generation on adolescent language minority students: A longitudinal follow-up study), I examined longitudinal treatment effects of an academic language intervention, Word Generation, on 5,052 adolescent EO and language minority students' academic vocabulary knowledge. Thirteen middle schools in an urban district in California were randomized to treatment and control conditions. Using individual growth modeling across four waves of data, I tested if EO and language minority students learned the target vocabulary words during the instructional year and maintained that vocabulary knowledge one year after during the follow-up year. The results indicate that there was a main effect of treatment on students' academic vocabulary knowledge. EO and language minority students in the treatment condition showed more growth in their academic vocabulary knowledge than those in the control condition. In addition, students in the treatment condition were able to maintain this improvement in their academic vocabulary knowledge in the follow-up year.

Themes

Each study in this dissertation answered specific research questions regarding adolescent language minority students' vocabulary growth trajectories. Taken together, there are overarching themes that emerged.

Heterogeneity of language minority students

The results from the three studies confirm that language minority students indeed are a diverse group of students with different levels of vocabulary and reading skills. Results from Study 1 and Study 3 indicate that IFEP students are the highest-performing group on average in regards to their vocabulary knowledge. EO and RFEP students tended to show a relatively

similar performance on vocabulary outcomes. Unsurprisingly, LEP students were the lowest-performing group. With the limited data that I used, it was not possible to further examine why IFEPs were at this advantage at baseline and throughout the study. One possible explanation could be due to how the state of California identifies and classifies its language minority students. All language minority students are required to take the English proficiency test (i.e., California English Language Development Test [CELDT]) when they first enter school. If the test were rigorous and difficult to pass, only those who have excellent mastery of English language skills would then be classified as IFEP. IFEP students who enter school with English proficiency to pass the test are more likely to gain more vocabulary knowledge and reading skills in their school years (i.e., Matthew Effect; Stanovich, 1986). This phenomenon is also shown in Study 2 where early-redesignated students consistently outperformed the recently-redesignated students during their middle school years. These results suggest that students who obtained sufficient English proficiency in order to reclassify in primary grades (e.g., third grade) may have had more opportunities and skills to learn new words and master reading comprehension strategies in- and out-of-school settings compared to students who were redesignated in the later grades (e.g., sixth grade). The results from these studies highlight that it could be difficult to close the achievement gap within the language minority student population despite their general developmental growth and common response to education. Thus, one take-away from these findings is that educators and policymakers need to invest in earlier supports to language minority students so that they may be redesignated as early as possible to better enjoy the reading-related outcomes.

Instructional practices that support language minority students' vocabulary development

Results from Study 3 underscore that vocabulary knowledge is amenable to targeted instruction for both EO and language minority students. Both EO and language minority students

119

in the treatment condition showed growth in their academic vocabulary knowledge and they were able to maintain that growth in the follow-up year. This is consistent with the research literature where researchers found well-designed interventions are beneficial for enhancing both EO and language minority students' general and content-specific vocabulary knowledge (August, Branum-Martin, Cardenas-Hagan, & Francis, 2009; Carlo et al., 2004; Lawrence, Capotosto, Branum-Martin, White, & Snow, 2012; Lesaux, Kieffer, Faller, & Kelley, 2010; Lesaux, Kieffer, Kelley, & Harris, 2014; Proctor et al., 2011; Snow, Lawrence, & White, 2009; Townsend & Collins, 2009; Vaughn et al., 2009). Researchers contend that academic vocabulary knowledge is essential for students' reading comprehension across different subject areas, especially in post-primary grades (Nagy & Townsend, 2012; Scarcella, 2003; Uccelli et al., 2015). However, it is often difficult to teach general academic vocabulary words during regular instruction hours because they are not tied specifically to any academic concepts or ideas. Thus, interventions, such as the Word Generation program, that facilitate word learning across different content areas may be beneficial for both the content-area teachers who may struggle with incorporating vocabulary lessons during class and students who need to gain academic vocabulary knowledge for improved reading comprehension. Given that Word Generation is a program that requires only about 15 minutes per day, the fact that both EO and language minority students were able to learn and maintain academic vocabulary knowledge is encouraging.

Methodological considerations

In conducting research, there can be diverse ways to analyze data based on research questions of interest. The analytical approach that I used in all three studies was multilevel models for change (Singer & Willet, 2003). Because I was interested in examining developmental trajectories of adolescent students' vocabulary and reading outcomes, these

120

multilevel models seemed most appropriate. These models allowed me to use up to four waves of data from each student to create a model of vocabulary and reading growth over the course of two years. Particularly in Study 3, I had competing analytical methods that I could use. If my primary concern was to evaluate Word Generation and examine treatment fadeout, I could have used multilevel regression analysis taking into account that students are nested in schools. However, I was more interested in analyzing the data from a language-learning, developmental point of view. From this perspective, I believed that students in both control and treatment conditions would gain some academic vocabulary knowledge regardless of receiving Word Generation instruction. In order to examine students' natural learning trajectory and change in growth due to Word Generation, longitudinal growth models seemed most appropriate.

Future work

This dissertation provides some detailed information about adolescent language minority students' vocabulary growth trajectories and their short- and long-term response to an academic language intervention. The results from these three studies provide some directions for future research. In closing, I discuss potential areas for future research that would help us extend our knowledge about language minority students' language and literacy development.

Different reclassification criteria of language minority students

For this dissertation work, I used students' data from a school district in California. California is known to have relatively rigorous classification criteria for reclassification as it uses multiple sources of data in the process. This could have been one of the reasons why RFEP students were performing relatively well in comparison to their EO peers. The rigor or leniency of the criteria is very likely to influence how RFEPs perform in language and literacy outcomes and the time they get redesignated.

121

There is not much research conducted on timing and process of redesignation of LEP students. Thus, it is still not certain what the costs and benefits would be for early redesignation resulting from lenient criteria or late redesignation due to rigorous classification process. As discussed in Study 2, early or late redesignation can have impacts on both individual LEP students and schools that serve them. More research that examines the assessments and criteria that are used in reclassification process across states and how these differences affect the students and school districts would be needed.

Subgroups of LEP students

California is known to have a large percentage of long-term LEP students, or LEP students who did not get redesignated after several years of English language development instruction (Olsen, 2010). The LEPs who were included in this dissertation could have been categorized as three different subgroups: the newcomers, developing LEPs, and long-term LEPs. The newcomers group would include students who just moved and joined the school system in the United States. The developing LEPs would be students who have been receiving English language development instruction for a couple of years. The long-term LEPs would be LEPs who have spent about seven years in the American schools but did not gain sufficient English proficiency to be reclassified. However, due to lack of data, I could not make this distinction with the LEPs in my sample. Understanding how students in each LEP group perform in regards to their language and literacy outcomes and designing intervention that are specifically targeted towards their needs would be another fruitful area for research.

Vocabulary assessments

One of the weaknesses in the vocabulary assessments used in the three studies of this dissertation is that they were all synonym tasks: all items asked students to find the synonyms of

the target words. These types of vocabulary assessments have been criticized because if students are expected to leverage word knowledge to comprehend complex texts, they need to know and understand multiple dimensions of vocabulary words such as literal meaning, connotative meaning, semantic relationship with other words, and morphology (Nagy & Scott, 2000; Pearson, Hiebert, & Kamil, 2007). If we were to more comprehensibly assess students' vocabulary knowledge, we would need to take into account multiple dimensions of vocabulary words and think about ways to address them.

Another avenue for research related to vocabulary assessment would be to examine potential factors that affect accurate measurement of vocabulary knowledge among EO and language minority students. For both general and academic vocabulary assessments that were used in this dissertation, the reliability coefficients (Cronbach's alpha) were different for each subgroup of students. They were relatively high for EO students but low for LEPs. These different alpha coefficients indicate that these measures are not equally reliable across groups of students. In addition, concurrent validity as measured by correlations among general vocabulary, academic vocabulary, and reading comprehension varied by students' language designations (see Appendix B for the reliability and validity coefficients for EO and LM students). Better understanding of item features that affect item difficulty for groups of students would provide a foundation for developing assessments that are theoretically and psychometrically sound.

Conclusion

Adolescent language minority students are still an under-studied student population in the research literature. The overarching goal of my dissertation was to better understand their language development in their middle school years. The findings of the three studies within this dissertation underscore that adolescent language minority students are heterogeneous in terms of

their English language abilities and both EO and language minority students can benefit from an academic language intervention.

References

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness, 2*(4), 345-376. doi:10.1080/19345740903217623

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., . . . White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188-215. doi:10.1598/RRQ.39.2.3

Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., . . . Kristapovich, P. (2014). *The condition of education 2014 (NCES 2014-083)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Lawrence, J. F., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. E. (2012). Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program. *Bilingualism: Language and Cognition, 15*(3), 437-451. doi:10.1017/S1366728911000393

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly, 45*(2), 196-228. doi:10.1598/RRQ.45.2.3

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary Instruction for linguistically diverse adolescents evidence from a randomized

field trial. *American Educational Research Journal, 51*(6), 1159-1194. doi:10.3102/0002831214532165

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading test technical report: Forms S and T*. Chicago, IL: The Riverside Publishing Company.

Nagy, W., & Scott, J. A. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269-284). Mahwah, NJ: Erlbaum.

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

Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational opportunity for California's long term English learners*. Long Beach, CA: Californians Together.

Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly, 42*(2), 282-296. doi:10.1598/RRQ.42.2.4

Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C. E., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing, 24*(5), 517-544. doi:10.1007/s11145-009-9218-2

Scarcella, R. C. (2003). *Academic English: A conceptual framework*. Berkeley, CA: The University of California Linguistic Minority Research Institute.

Singer, J., & Willet, J. (2003). *Applied longitudinal data analysis: Modeling change and even occurrence*. New York, NY: Oxford University Press.

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2*(4), 325-344. doi:10.1080/19345740903167042

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407. doi:10.1598/RRQ.21.4.1

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: An intervention study. *Reading and Writing, 22*(9), 993-1019. doi:10.1007/s11145-008-9141-y

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sanchez, E. (2015). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics, 36*(5), 1077-1109. doi:10.1017/S014271641400006X

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: Findings from two experimental studies. *Journal of Research on Educational Effectiveness, 2*(4), 297-324.

Academic vocabulary test forms

Year 1 academic vocabulary test form that was used in Study 1, 2, and 3

Vocabulary Assessment

681

<u>Directions:</u> Read the sentence and choose the word or group of words that mean the same or almost the same as the **<u>underlined word</u>**. Fill in the appropriate bubble/s on this answer sheet (a, b, c, or d).

1. I did not think I could <u>recite</u> the poem.

   - a. rehearse
   - b. perform
   - c. understand
   - d. memorize

2. The essay needed <u>cohesion</u>.

   - a. comments
   - b. disorganization
   - c. detail
   - d. consistency

3. I had to <u>constrain</u> my emotions.

   - a. repress
   - b. communicate
   - c. release
   - d. show

4. The lawyer <u>amended</u> the document.

   - a. damaged
   - b. authored
   - c. modified
   - d. read

5. They <u>assess</u> your skills.

   - a. adjust
   - b. evaluate
   - c. ignore
   - d. discuss

6. Her comment was not <u>relevant</u>.

   - a. real
   - b. kind
   - c. related
   - d. restricted

7. The new artwork in the office created a lot of <u>controversy</u>.

   - a. approval
   - b. conversation
   - c. appreciation
   - d. disagreement

8. The couple found they were <u>compatible</u> with each other.

   - a. angry
   - b. conflicted
   - c. bored
   - d. harmonious

9. They were <u>paralyzed</u> with fear.

   - a. paranoid
   - b. shaken
   - c. immobilized
   - d. active

10. I always <u>obtain</u> permission.

   - a. ignore
   - b. receive
   - c. lose
   - d. obstruct

WG_ASS_G06_XXX_11_X_X

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

21. Her injury <u>restricted</u> her.

- ○ a. strengthened
- ○ b. resonated
- ○ c. limited
- ○ d. damaged

22. The water was <u>contaminated</u>.

- ○ a. contradicted
- ○ b. clean
- ○ c. refreshing
- ○ d. impure

23. The athlete <u>attained</u> his goal.

- ○ a. achieved
- ○ b. discussed
- ○ c. earned
- ○ d. missed

24. He could not <u>retain</u> his confidence.
.

- ○ a. let go of
- ○ b. stay away from
- ○ c. risk losing
- ○ d. hold on to

25. He felt <u>apathy</u>.

- ○ a. appreciation
- ○ b. unconcern
- ○ c. excitement
- ○ d. tenderness

26. The teacher <u>enforced</u> the rule.

- ○ a. neglected
- ○ b. imposed
- ○ c. broke
- ○ d. ensured

27. The criminal received <u>amnesty</u>.

- ○ a. forgiveness
- ○ b. animosity
- ○ c. punishment
- ○ d. presents

28. The team <u>excluded</u> the new player.
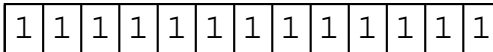
- ○ a. invited in
- ○ b. tested out
- ○ c. left out
- ○ d. expanded with

29. There was an <u>invasion</u>.

- ○ a. evacuation
- ○ b. intrusion
- ○ c. examination
- ○ d. instruction

30. The man <u>documented</u> what happened.

- ○ a. dodged
- ○ b. forgot
- ○ c. recorded
- ○ d. decided
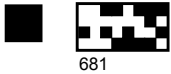
31. The team <u>comprises</u> eight players.

    ○ a.   includes

    ○ b.   co-opts

    ○ c.   nominates

    ○ d.   excludes

32. He <u>maintained</u> a good reputation.

    ○ a.   lost

    ○ b.   manipulated

    ○ c.   received

    ○ d.   kept

33. He <u>perceived</u> her anger.

    ○ a.   inspired

    ○ b.   noticed

    ○ c.   prepared

    ○ d.   ignored

34. He <u>acquired</u> a pet.

    ○ a.   got

    ○ b.   trained

    ○ c.   lost

    ○ d.   adored

35. She showed great <u>aptitude</u>.

    ○ a.   apprehension

    ○ b.   talent

    ○ c.   clumsiness

    ○ d.   bravery

36. The plan had several <u>components</u>.

    ○ a.   parts

    ○ b.   strategies

    ○ c.   combinations

    ○ d.   problems

37. The <u>interaction</u> was friendly.

    ○ a.   invitation

    ○ b.   interruption

    ○ c.   exchange

    ○ d.   parting

38. The pattern on the quilt was <u>complex</u>.

    ○ a.   complicated

    ○ b.   beautiful

    ○ c.   simple

    ○ d.   contextual

39. The benefits of the surgery <u>outweigh</u> the risks.

    ○ a.   are equally important as

    ○ b.   are less important than

    ○ c.   are more important than

    ○ d.   are important with

40. Helping poor people seemed <u>intrinsic</u> to Josie.

    ○ a.   natural

    ○ b.   external

    ○ c.   maternal

    ○ d.   intimate
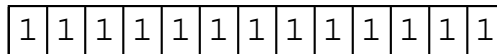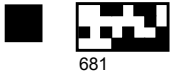
41. She <u>prescribed</u> a different plan.

- ○ a. enjoyed
- ○ b. prepared
- ○ c. prohibited
- ○ d. recommended

42. She <u>assumed</u> the worse.

- ○ a. described
- ○ b. doubted
- ○ c. supposed
- ○ d. assailed

43. This decision is <u>critical</u>.

- ○ a. difficult
- ○ b. important
- ○ c. criminal
- ○ d. trivial

44. The room was not <u>altered</u>.

- ○ a. normal
- ○ b. alarmed
- ○ c. changed
- ○ d. cleared

45. She had a <u>distinct</u> accent.

- ○ a. unique
- ○ b. typical
- ○ c. disappointing
- ○ d. foreign

46. The baby had unusual <u>attributes</u>.

- ○ a. clothes
- ○ b. features
- ○ c. attitudes
- ○ d. emotions

47. We can <u>conserve</u> energy.

- ○ a. save
- ○ b. conduct
- ○ c. use
- ○ d. waste

48. In <u>subsequent</u> weeks, we learned of the hero's actions.

- ○ a. earlier
- ○ b. later
- ○ c. superior
- ○ d. main

49. We came to a <u>conclusion</u>.

- ○ a. beginning
- ○ b. conviction
- ○ c. reception
- ○ d. decision

50. The argument was not <u>logical</u>.

- ○ a. reasonable
- ○ b. angry
- ○ c. longitudinal
- ○ d. ridiculous

Year 2 academic vocabulary test form that was used in Study 1, 2, and 3

Vocabulary Assessment

Draft

<u>Directions:</u> Read the sentence and choose the word or group of words that mean the same or almost the same as the
<u>underlined word</u>. Fill in the appropriate bubble/s on this answer sheet (a, b, c, or d).

1. We are trying to <u>promote</u> a healthy diet.

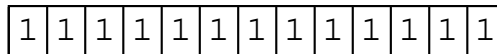   ○   a.   encourage

     ○   b.   promise

     ○   c.   eat

     ○   d.   buy

2. I <u>interpreted</u> the directions differently.

     ○   a.   understood

     ○   b.   interrelated

     ○   c.   was given

     ○   d.   repeated

3. The new artwork in the office created a lot of <u>controversy</u>.

     ○   a.   approval

     ○   b.   conversation

     ○   c.   appreciation

     ○   d.   disagreement

4. The volunteer will <u>contribute</u> to the project.

     ○   a.   agree

     ○   b.   contrast

     ○   c.   give

     ○   d.   run

5. He will <u>analyze</u> the information.

     ○   a.   ignore

     ○   b.   anchor

     ○   c.   remember

     ○   d.   examine

6. We had <u>sufficient</u> food at the party.

     ○   a.   delicious

     ○   b.   too much

     ○   c.   standard

     ○   d.   enough

7. The science <u>exhibit</u> was interesting.

     ○   a.   display

     ○   b.   experiment

     ○   c.   laboratory

     ○   d.   creature

8. A <u>decade</u> ago, the school had another name.

     ○   a.   a long time

     ○   b.   series

     ○   c.   ten years

     ○   d.   decimal

9. Scientists have a <u>hypothesis</u> about how the world began.

     ○   a.   question

     ○   b.   history

     ○   c.   guess

     ○   d.   photosynthesis

10. There is a <u>conflict</u> between the two groups.

     ○   a.   fight

     ○   b.   treaty

     ○   c.   inflict

     ○   d.   concerted

11. He was <u>transferred</u> to another school.

  ○ a. referred to

  ○ b. sent to

  ○ c. treated by

  ○ d. rejected by

12. The police <u>eliminated</u> her from the list of suspects.

  ○ a. removed

  ○ b. added

  ○ c. reduced

  ○ d. elected

13. They were <u>paralyzed</u> with fear.

  ○ a. paranoid

  ○ b. shaken

  ○ c. immobilized

  ○ d. active

14. The student will <u>monitor</u> his own writing.

  ○ a. improve

  ○ b. add to

  ○ c. motivate

  ○ d. check

15. The magician <u>revealed</u> his secret.

  ○ a. discovered

  ○ b. reviewed

  ○ c. showed

  ○ d. hid

16. The dentist <u>extracted</u> her tooth.

  ○ a. cleaned

  ○ b. examined

  ○ c. replaced

  ○ d. removed

17. She <u>indicated</u> that she was hungry.

  ○ a. denied

  ○ b. thought

  ○ c. showed

  ○ d. indeed

18. The new law will <u>substitute</u> for the old law.

  ○ a. replace

  ○ b. change

  ○ c. support

  ○ d. deny

19. In <u>subsequent</u> weeks, we learned of the hero's actions.

  ○ a. earlier

  ○ b. later

  ○ c. superior

  ○ d. main

20. In science class, we are looking at the <u>impact</u> of smoking.

  ○ a. weight

  ○ b. biology

  ○ c. imposition

  ○ d. effect

133

21. The athelete <u>attained</u> his goal.

    ○ a. achieved

    ○ b. discussed

    ○ c. earned

    ○ d. missed

22. This coach <u>prohibits</u> jumping jacks.

    ○ a. prefers

    ○ b. insists on

    ○ c. doesn't allow

    ○ d. stops

23. Her comment was not <u>relevant</u>.

    ○ a. real

    ○ b. kind

    ○ c. related

    ○ d. restricted

24. There are <u>diverse</u> ideas about freedom.

    ○ a. singular

    ○ b. divisive

    ○ c. active

    ○ d. varied

25. The tests were <u>suspended</u>.

    ○ a. allowed

    ○ b. hard for students

    ○ c. suspicious

    ○ d. stopped for awhile

26. The candidate's <u>ethics</u> were a problem in his campaign.

    ○ a. morals

    ○ b. finances

    ○ c. emotions

    ○ d. relatives

27. The couple found they were <u>compatible</u> with each other.

    ○ a. angry

    ○ b. conflicted

    ○ c. bored

    ○ d. harmonious

28. The <u>interaction</u> was friendly.

    ○ a. invitation

    ○ b. interruption

    ○ c. exchange

    ○ d. parting

29. That was a <u>dramatic</u> story.

    ○ a. non-fiction

    ○ b. drastic

    ○ c. exciting

    ○ d. very boring

30. He doesn't <u>react</u> well when he's tired.

    ○ a. study

    ○ b. bounce

    ○ c. receive

    ○ d. respond

134

31. I always <u>obtain</u> permission.

    ○   a.   ignore

    ○   b.   receive

    ○   c.   lose

    ○   d.   obstruct

32. The team <u>comprises</u> eight players.

    ○   a.   includes

    ○   b.   co-opts

    ○   c.   nominates

    ○   d.   excludes

33. The building <u>collapsed</u> after the earthquake.

    ○   a.   exploded

    ○   b.   fell apart

    ○   c.   stayed standing

    ○   d.   collated

34. The new law was <u>enforced</u>.

    ○   a.   neglected

    ○   b.   entrapment

    ○   c.   broken

    ○   d.   executed

35. I did not think I could <u>recite</u> the poem.

    ○   a.   rehearse

    ○   b.   perform

    ○   c.   understand

    ○   d.   memorize

36. There is a <u>cycle</u> to the seasons.

    ○   a.   pattern

    ○   b.   weather

    ○   c.   reason

    ○   d.   cyclone

37. It can be hard to <u>generate</u> ideas.

    ○   a.   destroy

    ○   b.   gather

    ○   c.   produce

    ○   d.   gesture

38. The boy took a new <u>route</u>.

    ○   a.   quest

    ○   b.   bus

    ○   c.   routine

    ○   d.   path

39. He <u>acquired</u> a pet.

    ○   a.   got

    ○   b.   trained

    ○   c.   lost

    ○   d.   adored

40. The <u>bulk</u> of students got a good score.

    ○   a.   majority

    ○   b.   minority

    ○   c.   bounty

    ○   d.   best group

41. Please, <u>release</u> the baby lion.

   ○   a.   cure

   ○   b.   rescue

   ○   c.   let go of

   ○   d.   lock up

42. The governor <u>allocated</u> funds for the new playground.

   ○   a.   took back

   ○   b.   asked for

   ○   c.   tapped into

   ○   d.   set aside

43. That was an <u>enormous</u> breakfast.

   ○   a.   nutritious

   ○   b.   small

   ○   c.   huge

   ○   d.   encased

44. The colors <u>contrast</u>.

   ○   a.   compare

   ○   b.   differ

   ○   c.   blend

   ○   d.   shine

45. <u>Incentives</u> are not always necessary.

   ○   a.   rewards

   ○   b.   punishments

   ○   c.   intentions

   ○   d.   changes

46. The benefits of the surgery <u>outweigh</u> the risks.

   ○   a.   are equally important as

   ○   b.   are less important than

   ○   c.   are more important than

   ○   d.   are more important with

47. We are studying American <u>culture</u> in school.

   ○   a.   way of life

   ○   b.   history

   ○   c.   way of voting

   ○   d.   couture

48. Speaking Korean will <u>enable</u> you to find a better job.

   ○   a.   force

   ○   b.   allow

   ○   c.   forbid

   ○   d.   encourage

49. Scientists <u>project</u> that the earth will be hotter in 1000 years.

   ○   a.   program

   ○   b.   believe

   ○   c.   predict

   ○   d.   deny

50. He <u>acknowledged</u> that he was responsible.

   ○   a.   admitted

   ○   b.   tackled

   ○   c.   bragged

   ○   d.   denied

APPENDIX B

Reliability and validity coefficients by students' language designations

Reliability coefficients of general and academic vocabulary tests by students' language designations

	General Vocabulary	Academic Vocabulary
Total Sample	0.89	0.91
EO	0.91	0.92
IFEP	0.86	0.89
RFEP	0.84	0.86
LEP	0.74	0.78

Note. The coefficients are Cronbach's alpha values obtained from the sample in Study 1 (wave 1). The reported reliability (Kuder-Richardson formula 20) in the technical manual of the Gates-MacGinitie Reading Test ranges from .90 to.92. EO = English-only, IFEP = Initially fluent English proficient, RFEP = Redesignated fluent English proficient, LEP = Limited English proficient.

Concurrent validity coefficients among general vocabulary, academic vocabulary, and reading comprehension for total sample

	General Vocabulary	Academic Vocabulary	Reading Comprehension
General Vocabulary	1		
Academic Vocabulary	0.81	1	
Reading Comprehension	0.76	0.77	1

Note. Concurrent validity coefficients were obtained by estimating the correlations among measures. Students' general vocabulary and reading comprehension scores were obtained from the vocabulary and reading subtests of the Gates-MacGinitie Reading Test, which is a validated and standardized measure that is used widely in literacy research. These correleations were obtained from the sample in Study 1 (wave 1).

Concurrent validity coefficients among general vocabulary, academic vocabulary, and reading comprehension for English-only students

	General Vocabulary	Academic Vocabulary	Reading Comprehension
General Vocabulary	1		
Academic Vocabulary	0.79	1	
Reading Comprehension	0.77	0.78	1

Concurrent validity coefficients among general vocabulary, academic vocabulary, and reading comprehension for initially fluent English proficient students

	General Vocabulary	Academic Vocabulary	Reading Comprehension
General Vocabulary	1		
Academic Vocabulary	0.78	1	
Reading Comprehension	0.76	0.75	1

Concurrent validity coefficients among general vocabulary, academic vocabulary, and reading comprehension for redesignated fluent English proficient students

	General Vocabulary	Academic Vocabulary	Reading Comprehension
General Vocabulary	1		
Academic Vocabulary	0.76	1	
Reading Comprehension	0.68	0.7	1

Concurrent validity coefficients among general vocabulary, academic vocabulary, and reading comprehension for limited English proficient students

	General Vocabulary	Academic Vocabulary	Reading Comprehension
General Vocabulary	1		
Academic Vocabulary	0.56	1	
Reading Comprehension	0.50	0.51	1