

# UC Berkeley

## The CATESOL Journal

### **Title**

A Language-Related Comparison of Generation 1.5 and L1 Student Writing

### **Permalink**

<https://escholarship.org/uc/item/5f4742qk>

### **Journal**

The CATESOL Journal, 22(1)

### **ISSN**

1535-0517

### **Author**

Doolan, Stephen M.

### **Publication Date**

2011

### **DOI**

10.5070/B5.36218

### **Copyright Information**

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



## A Language-Related Comparison of Generation 1.5 and L1 Student Writing

“Generation 1.5” is a term being used to describe a type of second language (L2) long-term U.S. resident who may demonstrate persistent language-related challenges (Roberge, Siegel, & Harklau, 2009). Among the difficulties commonly noted with Generation 1.5 students are problems in controlling the academic register expected in university writing tasks. Because of the growth of this population in U.S. schools, tertiary instructors need a better sense of whether patterns of language and literacy challenges are present in the writing of Generation 1.5 students. The goal of this exploratory study was to analyze linguistic/textual features of students’ writing. Specifically, this study was designed to determine whether 25 language-related measures of Generation 1.5 student writing would distinguish their texts from those of L1 classmates of similar writing proficiency. Results indicate similar patterns of textual features between groups. Implications are discussed in relation to the prevalent claims of Generation 1.5 writing.

### Introduction

Careful records of demographic information are not often kept by institutions at the tertiary level (Ignash, 2000); however, demographic information from K-12 education suggests that U.S.-educated language-minority students are dramatically affecting trends in college enrollment. Fix and Passel (2003) state that nationally, one in five students in grades K-12 is a child of immigrants. A more recent report by Batalova, Fix, and Murray (2007) states that enrollment in Limited English Proficiency programs in K-12 schools increased by 57% in the 10 years between 1996 and 2006. According to the Western Interstate Commission for Higher Education (2008), high school graduates in 2014 are projected to be 20% Hispanic and 7% Asian with many of these students being children of immigrants. The impact of this growing population is already being felt in the classroom, with future trends likely to increase this impact. Primary and secondary schools across the nation serve as a pipeline for U.S.-educated language-minority students in postsecondary education, and all indications suggest that this population will represent increasing percentages of college and university students in the years to come.

As tertiary education looks to accommodate these shifting demograph-

ics, it must acknowledge that academic writing takes on increased importance in higher education, serving as a form of gatekeeping by which students with underdeveloped writing and/or underdeveloped English language proficiency are disproportionately and negatively affected. Developmental writing classes at many community colleges are seeing an increasing number of long-term U.S. residents who speak a language other than English at home. Attrition rates are high within this population, perhaps most notably in the community college setting (Craig, personal communication, 2008; Lerman, personal communication, 2008), where the writing of these students with substantial U.S. educational experience but non-English first languages (L1) is claimed to involve a set of language-related difficulties (Roberge, Siegal, & Harklau, 2009).

The quality of the writing instruction that these students receive is affected by their instructors' understanding of student needs. At the tertiary level, while composition instructors are often entrusted with the responsibility of teaching these U.S.-educated language-minority students in developmental courses, instructors' training to teach this group of students is limited by available research on their writing. For example, in my master's in Teaching Writing (MATW), the curriculum included only one course solely devoted to second language (L2) writing issues. Furthermore, in MATW programs, differentiating among types of second language writers is likely to be a very small part of any ESL writing course offered, with the majority of instructional material equating international students or recently arrived immigrants with L2 students. Yet, even for the scholar who extensively searches for research on U.S.-educated language-minority students, virtually nothing can be found beyond qualitative or theoretical discussions of this student population. While interest in this population has grown in the last decade (Harklau, Losey, & Siegal, 1999; Matsuda, 2008; Roberge et al., 2009), surprisingly little empirical research has been published detailing the language-related challenges faced by this subset of language-minority students.

Because the impetus of this study is contingent on group similarities and differences, it is important for readers to have a clear understanding of the subset of language-minority students that is guiding this article. The specific type of second language of English student under investigation in this article is here referred to as a Generation 1.5 student (Harklau et al., 1999; Roberge et al., 2009). These students are commonly viewed as having circumstantial and learner characteristics distinct from first language (L1) of English or L2 international or recently arrived immigrant students (Ferris, 2009; Reid, 1997). For the current study, a Generation 1.5 student is one who (a) has been in the U.S. educational system for more than 3 years, (b) regularly speaks a language other than English at home, and (c) is less than 25 years old. Because defining Generation 1.5 has been inconsistent and sometimes contentious across previous literature (Benesch, 2008; Matsuda & Matsuda, 2009), the rationale for this definition is provided below in more detail.

The evidence is strong that Generation 1.5 students represent a distinct group of students in the U.S. school system up to university levels, with important sociocultural distinctions influencing their literacy performances (see

Leki, 1999; Valdez, 2001). However, it is not clear that these sociocultural differences also lead to qualitatively distinct writing production, as is sometimes claimed or inferred (Ferris, 2009; Frodesen, 2009). To examine this question in more depth, more extensive empirical evidence of Generation 1.5 writing performance is needed. To address this gap, the current study explores possible language-related features particular to their written work. Such an exploration can provide writing instructors with a better sense of the linguistic strengths and difficulties that these students may face. The primary goal of the current study, then, is to use textual analysis to determine whether or not important language-related (e.g., *lexical diversity, cohesion, complexity*, etc.) differences are discernable between a relatively large number of developmental Generation 1.5 and L1 student texts in postsecondary contexts.

While this study in isolation is not generalizable, it is the hope of this researcher that the current study, along with additional future studies, will form a body of work to demonstrate whether Generation 1.5 writing is distinct from the writing of their classmates. If group differences are present in their writing, then catering instruction to the strengths and weaknesses of the groups may be pedagogically expedient. However, if no patterns of difference can be found between the writing of Generation 1.5 and L1 texts, instructors may more appropriately focus on *developmental* issues that affect Generation 1.5 and L1 writing alike.

### Literature Review

Before shifting to empirical Generation 1.5 research on writing, it is worth acknowledging the breadth of contributions to Generation 1.5 research that have addressed issues related to writing. Publications that have advanced our understanding of the Generation 1.5 student experience have focused on literacy (Allison, 2009), motivation (Rodby, 1999), strategic knowledge (Crosby, 2009), identity (Schwartz, 2006), and academic coping strategies (Leki, 1999). Valdez (2001) longitudinally tracked the academic challenges faced by four recently arrived Latino immigrant students 12-13 years old. Scholars such as Roberge (2002) have established the sociological framework for Generation 1.5 studies and grappled with defining this population against the objections to terminology by other scholars (Benesch, 2008; Matsuda & Matsuda, 2009). Generation 1.5 scholars have debated the placement of these students (Matsuda, 2008; Valdez, 1992); the curriculum (Snow & Kamhi-Stein, 1997; Wurr, 2004); teacher perceptions (Goen, Porter, Swanson, & Vandommelen, 2002); and pedagogical approaches (Ferris, 2009). All of these publications have advanced our understanding of Generation 1.5 students in the writing classroom. Yet, without research that examines the actual writing of large numbers of Generation 1.5 students, we cannot say with any degree of confidence that Generation 1.5 students create writing that is different from that of their classmates. While it is much clearer that sociocultural factors allow us to consider these students a group (though this remains a contentious classification to some), there is almost no convincing empirical evidence that Generation 1.5 students form a distinct group of writers—this despite the fact that scholastic attention

to Generation 1.5 students is mostly being conducted within the domain of L2 writing research in higher education. The issues may also be quite different for Generation 1.5 students in middle school and high school literacy research (cf. Valdez, 2001).

The term “Generation 1.5,” taken literally, presents the idea of being between generations. First-generation immigrants most often move to their new country of residence as adults, and second-generation immigrants are born in the new country of residence. “Generation 1.5” often refers to students who arrive in the US during their formative childhood years. The term “1.5 Generation” was first introduced by the sociologists Rumbaut and Ima (1988). Yet, as research on Generation 1.5 has evolved in the last 20 years, this criterion of arriving in the English-speaking country of residence during childhood has not remained consistent in the literature. According to Schwartz (2006), “Generation 1.5” has become overused, now referring not just to children born in another country but also to those born in ethnic enclaves of the US, technically making them second-generation U.S. citizens. Roberge (2002), on the other hand, embraces the wider range of students this term covers, including (a) “in-migrants” from U.S. territories where other languages are dominant, (b) children who move to the US to live with relatives and attend U.S. schools, (c) transnational families that shift multiple times between the US and their country of origin, (d) U.S.-born children in ethnic enclaves, and (e) immigrant speakers of “Other Englishes” (Nero, 1997). The degree to which all of the above factors contribute to an overarching linguistic profile is unknown; however, for the purpose of the current large-scale research, criteria were established that could include many of the groups listed above.

This project, indeed this line of research, attempts to impose group status onto students with a range of demographic characteristics and language experiences. Most scholars recognize the language and social factors that Generation 1.5 students bring to the classroom as a continuum, and thus they hesitate to construct a working definition. Others have shifted terminology while retaining many demographic characteristics commonly associated with the Generation 1.5 student.<sup>1</sup> While the definition of Generation 1.5 in the current article is broad enough to encompass most of the Generation 1.5 and Generation 1.5–like groups in higher education, readers should be aware that the variability in definition is an ongoing challenge for Generation 1.5 research.

The remainder of this review will focus exclusively on the growing number of claims and observations—some empirical, some anecdotal—specific to Generation 1.5 writing in tertiary education.

### ***Quantitative Research on Generation 1.5 Writing***

One of the most fundamental assumptions in the Generation 1.5 discussion is that Generation 1.5 writers are less proficient or less successful than their classmates. Pathney, Thomas-Spiegel, and Dillon (2009) surveyed information on large numbers of “like Generation 1.5 students” ( $N = 43,964$ ) and then compared them to all their remaining classmates who entered participating writing programs in ESL or precollegiate writing courses ( $N = 238,032$ ). Data were

taken from composition courses in nine community colleges and two universities in California. The researchers investigated how successfully Generation 1.5 students were performing in the contexts of their writing classes. Their results indicated that Generation 1.5 students in community college writing classes had lower GPAs than their classmates but were more likely to progress into subsequent writing classes and pass through the writing program than their classmates. Because this study does not examine students' actual written work, the claim made in this large-scale study—that the lower GPAs of Generation 1.5 writers “suggests that written academic communication presents a challenge for Generation 1.5 students . . .” (p. 143)—needs to be treated cautiously. Writing proficiency may be responsible for these differences in GPA, but many other factors may also affect GPA.

While two other quantitative studies have been conducted on Generation 1.5 writing errors (Foin & Lange, 2007; Mikesell, 2007), only one other quantitative study has investigated language use comparing Generation 1.5 and L1 writing. Connerty (2009) sampled Generation 1.5 ( $N = 24$ ), ESL ( $N = 25$ ), and L1 ( $N = 83$ ) students enrolled in first-year composition writing courses. She used statistical keyword analysis to identify unusual frequencies of lexical items and she analyzed grammatical features using tagged texts. In this sample, with this specific set of students, Connerty found that these Generation 1.5 students used significantly higher frequencies of possessive pronouns, determiners, and prepositional phrases and significantly lower frequencies of adverbs than L1 students. Connerty argues that the linguistic features unique to this sample of Generation 1.5 student writing suggest greater “self-representation” and more of a spoken, conversational style of writing than the writing produced by their L1 classmates. However, Biber (1988) has identified greater use of prepositions with written text types and adverbs with a dimension more generally associated with spoken text types. Connerty's claims are the first of their kind for Generation 1.5 research and would benefit from additional research under more controlled writing conditions and analysis in which the holistic quality of the writing produced was known. Furthermore, because this research is the first of its kind, we cannot know whether these patterns are generalizable.

Apart from these four studies, there does not appear to be other quantitative research on the writing abilities of Generation 1.5 students. Instead, a number of case study reports have been published that describe the writing difficulties encountered by individual Generation 1.5 students. These studies provide important insight for further investigation of Generation 1.5 writing.

### ***Qualitative Research on Generation 1.5 Writing***

A limited number of case studies have used linguistic terminology to discuss the writing of Generation 1.5 students. Two case studies by Frodesen and Starna (1999) and Frodesen (2009) tracked Generation 1.5 students longitudinally to identify strengths, weaknesses, and writing development for three Generation 1.5 students. Some strengths that were reported included a familiarity with academic writing conventions and a willingness to attempt many “complex structures.” Another strength was Generation 1.5 students' use of

references and citation conventions, and an attempt to integrate source material in ways that more closely resemble L1 writing. Writing development was observed in organization, paragraph development, and coherence. The authors report that the students had problems with coherence and illogical reasoning. In addition, a host of language-learning sentence-level challenges were noted, including comparison structures, unidiomatic language (seemingly referring to word-choice difficulties involving appropriate collocations), complement clause structures as well as relative clause structures, verb tense, word forms, appropriacy of function words, and trouble with modal use. Using a format of reporting these features with a few examples, Frodesen and Starna describe these students' lack of subordination and lack of elaboration. Specific reference to the linguistic and discourse-level challenges faced by these students suggests possible hypotheses regarding patterns of use among larger samples of Generation 1.5 students.

While both Frodesen (2009) and Frodesen and Starna (1999) are largely focused on error-related issues, the presence of linguistic and rhetorical terminology provides insight into potential markers of Generation 1.5 students at different proficiency levels. Of course, the three students described in these studies cannot be generalized to all Generation 1.5 students for a number of reasons. Differing definitions of Generation 1.5, the small sample size, the L1 backgrounds, and the specific institution are just a few of the key factors that affect generalizability. Yet these researchers are to be credited for making claims about these specific students. Whether or not these claims are generalizable is an empirical and falsifiable question.

Three remaining studies have made claims about the vocabulary abilities of Generation 1.5 students. Commenting on a specific group of students with demographic features common to Generation 1.5 students, Rodby (1999) claims that the students' finished products contained a lack of lexical variety, resulting in less precision and clarity than would be desired, though few examples are provided. Similarly, Allison (2009) concludes anecdotally that Generation 1.5 students enter college with fewer than 10,000 words while native speakers generally enter with between 10,000 and 100,000. Finally, Ferris (2009) uses a select passage of Generation 1.5 writing as evidence of Generation 1.5 tendencies to perhaps "play it safe" with vocabulary choices. According to Ferris, Generation 1.5 students may have larger vocabularies than more traditional L2 populations, but they may rely more on basic vocabulary because of their "ear" learner experiences with language development. These selective claims from case studies about Generation 1.5 lexical ability set the course for a productive line of research to test these hypotheses in a controlled quantitative analysis.

In the studies reviewed, appropriate control of cohesion, organization, lexical diversity and sophistication, clausal structures, and modal verb use are presented as below grade level for many Generation 1.5 students. Yet Frodesen (2009) also claimed that there is a complexity of thought in the syntax of Generation 1.5 writing. As mentioned repeatedly throughout the qualitative section of this literature review, the above studies represent an important first step in hypothesis development. Drawing on the above studies, larger-scale research

that investigates the linguistic markers that distinguish Generation 1.5 writing has the potential to confirm or raise doubts about the generalizability of these claims.

### ***Gap in the Research***

The studies cited above represent groundbreaking efforts in understanding Generation 1.5 writing. Yet despite this growing body of literature, virtually none of the Generation 1.5 studies collected large sets of writing samples under reasonably controlled conditions to determine if or how Generation 1.5 writing is unique in comparison to the writing of their L1 classmates. As such, the current study will investigate the following three research questions:

- Do Generation 1.5 students produce in-class writing of lower holistic quality than their L1 classmates in a matched instructional environment?
- Do Generation 1.5 students exhibit a more limited range of vocabulary in in-class writing than their L1 classmates?
- Do select language markers distinguish Generation 1.5 from L1 texts?

## **Method**

### ***Participants***

Writing samples were gathered from students enrolled in five developmental writing courses at a community college in California. The writing samples used for the current study included those of 41 Generation 1.5 students and 20 L1 students.

While qualitative research on Generation 1.5 often collects and reports a large amount of demographic information on individuals, the definitions used in previous quantitative research have been summarized in Table 1. Based on these previous studies, I define Generation 1.5 students as those who (a) have been in the U.S. educational system for more than 3 years, (b) regularly speak a language other than English at home, and (c) are less than 25 years old.

Furthermore, L1 students who completed the biographical information survey (Appendix A) answered “do/did not regularly speak a language other than English at home” and answered the remaining questions in a manner consistent with an L1 student profile. The majority of Generation 1.5 students had been in the U.S. educational system for more than 10 years (63%), and 92% reported speaking and listening skills in English as “Good,” “Very Good,” or “Excellent.” This is consistent with the widely held belief that Generation 1.5 students have well-developed L2 listening and speaking abilities.

### ***Materials***

Materials for this in-class writing study consisted of (a) an informed-consent document, (b) a biographical information survey (Appendix A), and (c) a writing prompt. The survey consisted of 14 questions mainly eliciting information about the language use and educational background of participants. Those participating in this study responded to the following writing prompt:



Looking back to high school and knowing what you know now, how could you have been better prepared for college? In other words, what might you have done differently regarding your behavior during those high school years—e.g., study habits, attitude, etc.? And what might your high school have done differently? For example, your high school could have made changes in coursework, teaching methods, counseling, etc.

Write an essay where you discuss the changes you would make. Be sure to have an introduction, 3-4 points supported by clear and detailed examples, and a conclusion that draws the essay to a close.

This prompt was created by writing instructors at the participating institution and served as the final examination as well as the writing sample for the current study.

### ***Procedures***

Approximately 90 minutes were used to obtain signatures on the informed-consent document, complete the biographical information survey, and complete the in-class writing assignment. For each of these developmental writing classes, the teacher of record administered the study materials to participants. In some classes, the informed-consent document and survey were completed the week before the writing assignment. Whether all three materials were presented at once or the consent and survey were administered before the final examination, ample time was allowed for in-class writing and none of the essays appears “unfinished” or hindered by time constraints. A memo was also created for administrators with instructions and tips for administering the study to help control writing conditions.

### ***Analysis***

Once writing samples were collected, names were changed to numbers before sending these writing samples to the researcher, allowing for anonymous data collection. Handwritten essays were then typed by the researcher and saved as electronic files.

Essays were scored using a 6-point holistic rubric (see Appendix B) to assess writing quality by two raters. In an effort to increase reliability of scoring, a training session was conducted with the two raters. The raters were doctoral students in Applied Linguistics with at least 1 year of experience in teaching first-year composition and at least 5 years of experience in L2 language instruction. After the training session, the remaining essays were scored by two raters. For discrepancies greater than 1 point (e.g., a score of 3 from one rater and a score of 5 from the other), a third rater was used to resolve the difference. If the third rater’s score was identical to one of the two scores from the other raters, then the two identical scores were used. If the third rater’s score fell between the scores of the other two raters, then the other two scores were averaged with the averaged score and the score from the third rater serving as the two scores used. The two separate scores were added, resulting in possible scores between 0 and 12.

**Table 1**  
**Definitions of Generation 1.5 Students**  
**From Previous Quantitative Studies**

<i>Author</i>	<i>Language criterion</i>	<i>Education criterion</i>	<i>Age criterion</i>	<i>Other</i>
Patthey, Thomas-Spiegel, & Dillon (2009)	has a primary language other than English	graduated from a U.S. high school	is under 22 years old	none
Foin & Lange (2007)	is enrolled in ESL programs	received all or most of his or her education in the U.S. educational system	none	none
Mikesell (2007)	speaks a language other than English at home	scored below 520 on the SAT verbal section;  demonstrated language errors on his or her university's ESL placement exam	no criterion established though the age range was listed as 18-30	has lived in the US for 7 or more years
Connerty (2009)	placed into an L2 composition class	graduated from a U.S. high school or 10 years minimum in the US with no formal education in US	none, but average age was 21.23	living in the US for at least 4 years

To determine the interrater reliability (or the consistency between raters), we conducted a Pearson's correlation statistical procedure. Results from this test indicated moderate agreement with a Pearson's coefficient of .63. Another indicator that reasonable reliability existed between raters was an agreement within 1 (on the 6 point scale) of 96%. For essays in which the two raters disagreed by more than 1, a third rater adjudicated these discrepant scores.

For every variable used in this study, we ran statistical procedures to determine whether the differences between Generation 1.5 and L1 text were significantly different from what we might expect to find merely due to chance. To

run certain statistical analyses, the data must meet certain assumptions. After checking to ensure that the data fit those assumptions, a statistical procedure called independent-sample *t* tests was then conducted to determine whether the holistic scores of Generation 1.5 texts were significantly different from the holistic scores of the L1 texts.

**Vocabulary Measures.** *The Compleat Lexical Tutor* (<http://www.lexutor.ca/vp/eng>; Cobb, 2003) is a free online computational analysis tool for researchers and educators. This site allows teachers and researchers to submit texts for analysis, providing counts for the categories listed below. For each text, the first 200 words were run through *The Compleat Lexical Tutor* to obtain the following lexical counts:

1. 1,000 most frequent English words (1K)—Taken from the General Service List based on a 5 million-word written corpus (West, 1953);
2. 1,000-2,000 most frequent English words (2K)—Taken from the General Service List based on a 5 million-word written corpus (West, 1953);
3. Academic Word List (AWL)—A list of the 570 most frequently occurring words in academic language that are not included in the first 2,000 words;
4. Off List Words (OLW)—All remaining words that are not found on one of the three above lists;
5. Type/Token Ratio (Different Types/Total Words);
6. Lexical Density (Content Words/Total).

For lists of the first three categories, go to [http://www.lexutor.ca/freq/lists\\_download](http://www.lexutor.ca/freq/lists_download).

Again a statistical analysis, Mann-Whitney *U*, was conducted to determine whether differences on the scores of these vocabulary measures between Generation 1.5 and L1 text were greater than we would expect due to random chance. Because six comparisons (one for each vocabulary variable) were made, a stricter standard was necessary for what constituted “random chance” ( $\alpha = .01$ ). Mann-Whitney *U* is a nonparametric statistical test and was used instead of the independent-sample *t* test because of nonnormal distribution.

**Coh-Matrix Variables.** Analysis of linguistic features was also conducted using Coh-Matrix (McNamara, Louwse, Cai, & Graesser, 2005) to identify potential markers of Generation 1.5 student writing. Coh-Matrix is a freely available text-analysis program created at the University of Memphis. Using Coh-Matrix, more than 60 computational variables are provided for each text; however, for the purposes of this study, only variables were used from Coh-Matrix that seemed most likely (based on previous research) to identify group differences. There is a growing body of publications using the Coh-Matrix program to demonstrate potentially important textual differences (see Crossley & McNamara, 2009, p. 122). Table 2 provides a list of the variables from Coh-Matrix used for the current study.

**Table 2**  
**Coh-Metrix Variables**

<i>Language variable</i>	<i>Language variable</i>
Cohesion	Lexical sophistication
Incidence of all connectives	Mean hypernym values of nouns
Proportion of content words that overlap between adjacent sentences	Mean hypernym values of verbs
Ratio of pronouns to noun phrases	Average syllables per word
Complexity	Fluency
Mean number of words before the main verb of main clause in sentences	Number of words per essay
Number of conditional expressions, incidence score	Number of paragraphs
Mean of tense and aspect repetition scores	Average sentences per paragraph
Noun Phrase Incidence Score (per 1,000 words)	Number of sentences
Mean number of modifiers per noun phrase	
Average words per sentence	

*Notes.* From McNamara, Louwrese, Cai, & Graesser (2005). For more information visit <http://cohmetrix.memphis.edu/cohmetrixpr/index.html>

Definitions for the above language features are provided to the degree possible; however, the lack of transparency as to how these counts are obtained is a limitation of Coh-Metrix.

Once again the Mann-Whitney *U* statistical test helped determine whether differences on the scores of these Coh-Metrix and corpus measures between Generation 1.5 and L1 text were greater than we would expect due to random chance. Because comparisons (one for each variable) were made, a stricter standard was necessary for what constituted “random chance” ( $\alpha = .002632$ ). As with the vocabulary measures, the Mann-Whitney *U* was a more appropriate statistical test than the independent-sample *t* test because the Coh-Metrix and corpus variables were not normally distributed

**Corpus Variables.** The 61 essays were tagged for linguistic features using a corpus linguistics tagging program designed by Douglas Biber. Counts of corpus variables were normalized per 200 words as detailed in Biber, Conrad, and Reppen (1998). That is, the total number of a specific structural feature was divided by the total number of words in the individual essay and then multiplied by 200. Normalizing the data (e.g., having counts per 200 words) allows for accurate comparisons of specific structural features to be made between differ-

ent texts despite the texts' containing unequal numbers of total words. Counts were normalized to 200 because this number represents a word count achieved by the large majority of participants.

For each linguistic feature that was analyzed, tagger identification was checked by the researcher for reliability. The linguistic features that were identified by the tagger were checked in each occurrence to ensure that the feature sought was the feature found. Definitions of the linguistic features analyzed were largely consistent with Biber (1988), as outlined in Appendix II of his publication. The corpus variables and operationalizations are provided below:

1. Prepositional phrases—Refer to Biber (1988) Appendix II for a complete list of prepositional phrases;
2. Modal verbs—Covering Biber's (1988) three functional classes of modals: (a) possibility, permission, or ability; (b) obligation or necessity; and (c) volition or prediction;
3. Adverbial subordinators—The following lexical items were counted only when beginning finite adverbial clauses: *because, for, since, although, though, whereas, as, while, if, in case, unless, so, such that, after, before, now that, once, till, until, when, whenever, wherever.*

Combining the vocabulary, Coh-Metrix, and corpus variables, 25 variables were investigated in this study, each variable with an empirical rationale providing reason to believe that group differences may emerge (see Appendix C).

## Findings

### *Holistic Scores*

Generation 1.5 and L1 texts received very similar holistic scores. In fact, there was no statistically significant difference between the two groups. Using an alpha level of .05, an independent-sample *t* test was conducted to evaluate whether the average holistic quality score differed significantly as a function of whether student texts came from L1 students or Generation 1.5 students. The test was not significant,  $t(54.65) = -1.48, p > .05$ . An examination of the group means indicated that students in the L1 text group ( $M = 8.30, SD = 1.26$ ) did not receive holistic quality scores that significantly differed from students in the Generation 1.5 group ( $M = 7.71, SD = 1.97$ ).

### *Vocabulary Measures*

The vocabulary measures yielded very similar results between the Generation 1.5 and L1 texts. Statistical analysis revealed that no differences existed between these two groups. The descriptive statistics and results of the Mann-Whitney *U* test for the vocabulary measures are provided in Table 3. Across all six vocabulary measures the means and standard deviations for these two groups remain surprisingly similar. Comparing the group means, we find a slightly larger occurrence of academic words used by L1 writers and a slightly large occurrence of off list word tokens used by L1 writers, yet it should be repeated that no statistically significant differences emerged from these analyses.

**Table 3**  
**Descriptive Statistics and Mann-Whitney U Statistics**  
**for Vocabulary Measures**

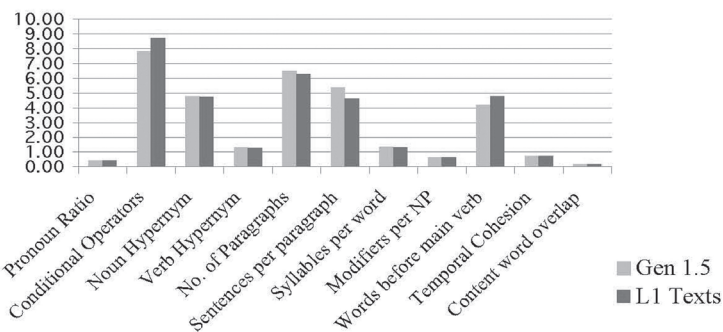
<i>Language variable</i>	<i>Mean and SD</i>		<i>Mean ranks</i>		<i>Stand. test statistic</i>	<i>Sig. (2-tailed)</i>
	Gen 1.5	L1	Gen 1.5	L1		
K1	183.83 (7.45)	182.30 (6.04)	32.11	28.72	-.700	.484
K2	7.76 (3.61)	7.40 (2.23)	31.72	29.52	-.456	.649
AWL	4.17 (2.51)	4.85 (2.96)	29.72	33.62	.813	.416
OLW	6.63 (3.29)	8.00 (3.26)	29.00	35.10	1.266	.206
T/T	.51 (.06)	.53 (.04)	29.83	33.40	.740	.460
L/D	.49 (0.5)	.47 (.03)	33.07	26.75	-1.312	.190

*Note.* K1 = 1,000 most frequent English words, K2 = 1,000-2,000 most frequent English words (2K), AWL = Academic Word List, OLW = Off List Words, T/T = Type/Token Ratio, Lex Den = Lexical Density

**Coh-Matrix Measures**

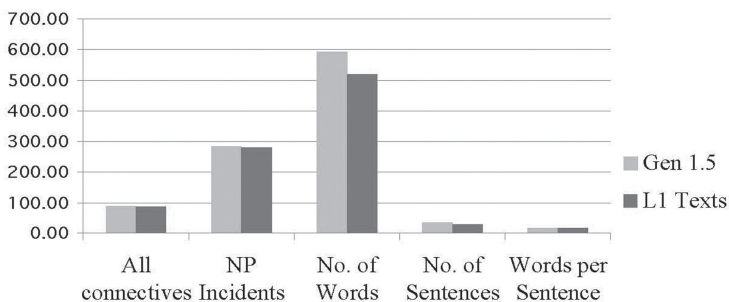
For the 16 Coh-Matrix variables, a comparison between Generation 1.5 and L1 texts once again demonstrates strong similarities. Once again, statistical analysis did not uncover any significant differences in relation to these variables. Results from the Coh-Matrix measures can be found in Figure 1 and Figure 2. The descriptive statistics and results of the Mann-Whitney U test can be found in Appendix D.

**Figure 1**  
**Coh-Matrix Variables Set #1: Comparing Group Means**



When comparing the group means, very slight differences are found in the conditional operators and words before the main verb (both with slightly larger numbers for the L1 texts), and the fluency counts of (a) number of words and (b) sentences per paragraph, both of which were slightly higher in the Generation 1.5 texts. Yet for the large number of variables tested, these counts were very similar, with no discernable patterns of group difference.

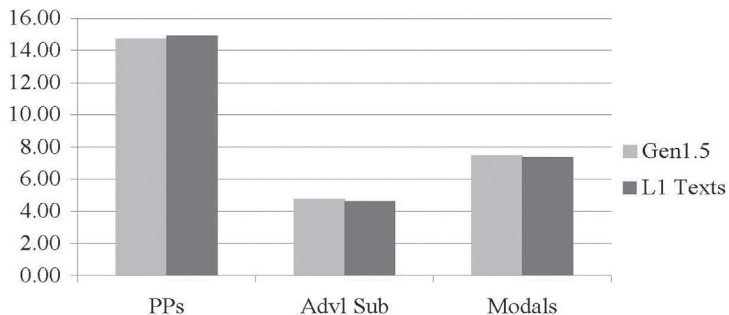
**Figure 2**  
**Coh-Metrix Variables Set #2: Comparing Group Means**



**Corpus Measures**

The three remaining variables analyzed with Biber’s corpus tagger followed the trends present in the first 22 variables, demonstrating no significant differences (see Appendix D). Normalized per 200 words, Figure 3 shows group means of Generation 1.5 and L1 texts for prepositional phrases, adverbial subordinators, and modal verbs all within .20 of the same mean for each variable.

**Figure 3**  
**Corpus Variables: Comparing Group Means (Normed per 200 Words)**



*Note.* PPs = Prepositional phrases, Advl Sub = Adverbial subordinators, Modals = Modal Verbs

**Discussion**

Because the writing obtained from both Generation 1.5 and L1 students in the current study came from developmental writing classes, it should be emphasized that the comparison of these two groups is a comparison between below-grade-level Generation 1.5 writing and below-grade-level L1 writing. In other words, the current study looks to determine whether subsets *within* developmental writing classes at the target institution generate distinct patterns of language-related challenges in their writing. Yet while some Generation 1.5 scholars have noted ways in which Generation 1.5 writing may resemble developmental L1 writing (Ferris, 2009; Reid, 1997), Generation 1.5 research has gained notoriety in L2 writing research circles under the premise (explicit or

implicit) that it is *not* L1 developmental writing, but instead some combination of L1 and L2 writing with features of both.

If no reliable patterns of difference exist between the writing of Generation 1.5 and L1 students at the developmental level, then teachers and researchers might do better to discuss the strengths and weaknesses of *developmental writing* rather than perpetuating these sociological group distinctions despite no tangible language-related patterns of difficulty associated with the writing produced by these groups. The 25 language-related variables used in the current study represent a first step in determining whether or not “Generation 1.5” might serve as a useful distinction as it relates to developmental writers.

### ***Interpreting Vocabulary Measures***

Some researchers (Allison, 2009; Ferris, 2009; Rodby, 1999) have claimed that Generation 1.5 writing is negatively affected by having a less fully developed range of vocabulary than their classmates possess. This assumption is sometimes based on the “ear” learner versus “eye” learner (Reid, 1997) argument, and the belief that because Generation 1.5 students have learned English primarily through speaking and listening, they may rely on “safer” language choices (Ferris, 2009) in their writing. This claim is probably based largely on intuition and anecdotal evidence (Allison, 2009), in which teachers have noticed vocabulary issues within the writing of some Generation 1.5 students. Furthermore, these claims from previous research do not address this issue of whether developmental L1 students would be “ear” learners as well.

Despite these claims, there appears to be no evidence from the current study to suggest that developmental Generation 1.5 students have greater vocabulary deficiencies than their developmental L1 classmates. It should, however, be noted that this group of developmental writers *as a whole* generated very few of the second 1,000 words, academic words, and off word list tokens. *The Compleat Lexical Tutor* website states that a “typical NS” will show scores of approximately “70% from first 1000, 10% from second thousand, 10% academic, and 10% less frequent words.” Following these guidelines, one might expect twice as many words from the second thousand, academic, and off list word categories than we found in the L1 texts. One reasonable interpretation of the vocabulary measures from the current study is that both Generation 1.5 students *and* L1 students are relying much too heavily on the first 1,000 words in the English language and would *both* benefit from instruction looking to increase the diversity of vocabulary choices in their writing.

### ***Interpreting Coh-Metrix and Corpus Measures***

The remaining variables analyzed in the current study involved an exploratory attempt to uncover differences between Generation 1.5 and L1 texts in relation to such constructs as cohesion, lexical sophistication, fluency, and complexity. Nineteen variables were analyzed to determine whether any of these variables that have been used to establish group differences in relation to other independent variables (e.g., proficiency, L1 vs. L2) would differentiate Generation 1.5 and L1 texts. Despite claims of the uniqueness of Generation



1.5 student writing based on anecdotal impressions (Holton, 2002), qualitative research (Frodesen, 2009; Frodesen & Starna, 1999), and teacher impressions of Generation 1.5 challenges (Goen et al., 2002), it appears that Generation 1.5 writers are *using* language features in ways similar to the language features used by their developmental L1 classmates. Contrary to the findings by Connerty (2009) demonstrating quantitative differences between Generation 1.5 texts and L1 texts, the current study did not find significant differences in any of the language-related variables analyzed.

Group means were surprisingly similar, with the only *slight* differences consistent across an entire construct being fluency measures, in which the Generation 1.5 writing was unexpectedly longer (text length, number of paragraphs, number of sentences per paragraph). Of these fluency measures, only text length has a robust tradition in previous research as a viable variable.

### ***Interpreting Holistic Similarities***

Obtaining holistic quality scores for the current study served as a control. While one might expect that holistic scores of students placed into the same developmental classes would have similar writing proficiencies, previous research suggests that this may not be the case. Patthey et al.'s (2009) large-scale quantitative research provides convincing evidence that Generation 1.5 students are achieving lower GPAs than their classmates in writing classes at the developmental and transfer levels. Based on the findings of the current study, however, it may not be the case that those lower GPAs found in Patthey et al. (2009) are the result of group-specific lower writing proficiency. The writing produced in the current study by Generation 1.5 students was not significantly different from the writing produced by the L1 students, indicating that, at least for this sample, Generation 1.5 students did not represent the less proficient writers in developmental classes at this community college. Another interpretation from these holistic findings is that the community college participating in the current study is more successfully placing students of similar proficiency levels within the same class than the classes sampled in the Patthey et al. (2009) study.

The second benefit for scoring these essays is to make additional comparisons with a knowledge of how similar the writing quality is between groups. That is, comparing linguistic variables between different groups should take into account differences in writing proficiency. Such differences would help sort out whether differences might be due to writing proficiency differences or differences among demographic groups (e.g., Generation 1.5, L1). Because there is no difference between the holistic quality of the essays in this sample of Generation 1.5 and L1 texts, it is reasonable to assume that the analyses of linguistic variables are made between writers of similar writing proficiency levels.

### **Conclusion**

Implications from the current study are restricted to a comparison between two student groups. The variables investigated in the current study can be associated with vocabulary abilities, lexical sophistication, fluency, cohesion, and complexity. While much more research is needed, findings from the cur-

rent study suggest that Generation 1.5 and L1 developmental writing may be very similar. If patterns of difference do exist between developmental Generation 1.5 and L1 texts, either this sample was atypical, or the language variables chosen for the current study did not uncover the differences.

For professionals who work with Generation 1.5 students at the community college level, findings from the current study have two important implications. First, the similarities between these two samples based on these variables are undeniable. Though developmental L1 writers may also be in need of language-related instruction, there is no evidence from the current study that their Generation 1.5 classmates are in *greater* need of language instruction.

Second, the current study is exploratory and the current article should open a discussion. Would instructors agree that the academic literacy challenges faced by developmental Generation 1.5 and L1 writers are largely the same (allowing for individual variation)? If so, what are the pedagogical advantages to identifying Generation 1.5 students in the writing classroom? If not, what language-related differences exist between the writing of these groups?

### ***Limitations***

The current study used frequency counts as a point of departure for identifying differences in patterns of language use. The assumption for those following this methodology is that differences in frequency counts could then be followed by a functional analysis to investigate how a select variable is being used differently. Because no sizable frequency differences were found, no functional analysis was conducted; however, it should be noted that one can make no claims as to how effectively these groups were using these features based on frequency counts. That is, just because a group of writers is using a linguistic feature does not necessarily mean they are using it well, using it accurately, or using the same feature in different ways.

On a separate note, research suggests that gathering demographic information from Generation 1.5 students through self-reporting may be problematic (Chiang & Schmida, 1999). Obviously, using school records and more extensive interviews would be ideal for gaining background information on participants; however, gathering large amounts of data from many Generation 1.5 students may require self-reporting as a necessary limitation.

Finally, inferential statistics have the potential to use data as a means to make generalizations about a larger population. Unfortunately, because this research took writing from only one community college, in one writing program, in a nonrandom fashion, the results of this study cannot be generalized beyond this specific environment.

### ***Conclusion and Future Research***

Interest in Generation 1.5 writing is both timely and important. Writing programs need to adjust to changing demographics within tertiary writing classes and better address the needs of students in developmental writing classes. Making a commitment to better understand and improve instruction for developmental writers could decrease attrition rates and improve student

learning outcomes. As teachers and researchers, the better sense we have of how Generation 1.5 students are unique, the more capable we will be of providing quality instruction. For example, what *specifically* do we mean when we say that a student is struggling with:

- Competence using basic syntax;
- A written accent that is too “nonnative” sounding; or
- Comfort with English (Goen et al., 2002).

Teacher impressions, intuition, and qualitative work are extremely important to our understanding of Generation 1.5 writing. The next step, however, is to determine whether the patterns we think we see in the classroom, or within the work of one or two students, are, in fact, generalizable to larger numbers of students and whether those patterns are unique to Generation 1.5 writing or more generally are a challenge for developmental writers.

From the current study, an objective observer could determine that, at least based on this sample of developmental writers, and on these 25 variables, there do not appear to be differences between the writing produced by developmental Generation 1.5 and developmental L1 students. Clearly, a great deal more research is needed before such a claim could be generalized to these developmental populations at large. Perhaps the next analysis that should be conducted is an error analysis, in which the amounts and types of errors produced by Generation 1.5 and L1 writers are reliably coded by multiple raters. A similar analysis that would be valuable to this line of research would be a carefully controlled attempt to document “idiomaticity,” a feature that factors prominently in Generation 1.5 commentary (Frodesen, 2009; Frodesen & Starna, 1999; Holten & Mikesell, 2007).

Generation 1.5 research has the potential to transform the discussion of L2 writing beyond the current dichotomy of L1 versus L2 texts. Yet, without empirical, quantitative support demonstrating that their writing involves patterns of use that are unique, the legitimate claim can be made that their writing is no different from the writing being produced by their classmates. This claim is entirely separate from sociocultural claims of group difference, but it should nevertheless be of great interest to Generation 1.5 research that is largely being published within the interest group of L2 writing.

### **Author**

*Stephen Doolan is a PhD candidate in Applied Linguistics studying at Northern Arizona University in Flagstaff. He received an MA in Teaching Writing from Humboldt State University. His dissertation looks to identify language-related and discourse patterns that distinguish Generation 1.5 student writing from that of their L1 and more traditional L2 classmates.*

### **Note**

<sup>1</sup>Such closely related alternative terms for Generation 1.5 include cross-over students (Schwartz, 2006), early-arriving resident students (Ferris, 2009), de-

developmental immigrant students (Crosby, 2010), U.S.-educated multilingual writers (Nakamaru, 2010), and resident nonnative speakers of English (Levi, 2004).

## References

- Allison, H. (2009). High school academic literacy instruction and the transition to college writing. In M. Roberge, M. Siegal, & L. Harklau (Eds.), *Generation 1.5 in college composition* (pp. 75-90). New York, NY: Routledge.
- Batalova, J., Fix, M., & Murray, J. (2007). *Measures of change: The demography and literacy of adolescent English learners*. New York, NY: National Center on Immigrant Integration Policy, Migration Policy Institute.
- Benesch, S. (2008). "Generation 1.5" and its discourses of partiality: A critical analysis. *Journal of Language, Identity, & Education*, 7, 294-311.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge, England: Cambridge University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. New York: Cambridge University Press.
- Chiang, Y. D., & Schmida, M. (1999). Language identity and language ownership: Linguistic conflicts of first-year university writing students. In L. Harklau, M. Kay, & M. Siegal (Eds.), *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL* (pp. 81-98). Mahwah, NJ: Erlbaum.
- Cobb, T. (2003). *Web vocabprofile*. Retrieved June 1, 2009, from <http://www.lex-tutor.ca/vp>, an adaptation of Heatley & Nation's (1994) *Range*.
- Connerty, M. (2009). *Variation in academic writing among Generation 1.5 learners, native English-speaking learners and ESL learners: The discursive self of G1.5 student writers*. Unpublished doctoral thesis, University of Birmingham, Birmingham, England. Retrieved November 29, 2010, from <http://etheses.bham.ac.uk/274>
- Crosby, C. (2009). Academic reading and writing difficulties and strategic knowledge of Generation 1.5 learners. In M. Roberge, M. Siegal, & L. Harklau (Eds.), *Generation 1.5 in college composition* (pp. 105-119). New York: Routledge.
- Crosby, C. (2010, March). *Linking literacies for developmental immigrant students*. Paper presented at the TESOL Conference, Boston, MA.
- Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18, 119-135.
- ETS (2004). *Test of written English guide*. Princeton, NJ: Educational Testing Service.
- Ferris, D. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly*, 28, 414-420.
- Ferris, D. (2009). *Teaching college writing to diverse student populations*. Ann Arbor: University of Michigan Press.

- Fix, M., & Passel, J. S. (2003). *U.S. immigration—Trends and implications for schools*. *Immigration Studies Program of the Urban Institute*. Presented at the National Association for Bilingual Education's NCLB [No Child Left Behind], Implementation Institute, New Orleans, LA.
- Foin, A. T., & Lange, E. J. (2007). Generation 1.5 writers' success in correcting errors marked on an out-of-class paper. *The CATESOL Journal*, *19*, 146-163.
- Frodesen, J. (2009). The academic writing development of a Generation 1.5 "latecomer." In M. Roberge, M. Siegal, & L. Harklau (Eds.), *Generation 1.5 in college composition* (pp. 91-104). New York, NY: Routledge.
- Frodesen, J., & Starna, N. (1999). Distinguishing incipient and functional bilingual writers: Assessment and instructional insights gained through second-language writing profiles. In L. Harklau, M. Kay, & M. Siegal (Eds.), *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL* (pp. 61-80). Mahwah, NJ: Erlbaum.
- Goen, S., Porter, P., Swanson, D., & Vandommelen, D. (2002). Working with Generation 1.5 students and their teachers: ESL meets composition. *The CATESOL Journal*, *14*, 121-171.
- Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, *9*, 123-145.
- Harklau, L., Losey, K. M., & Siegal, M. (Eds.). (1999). *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL*. Mahwah, NJ: Erlbaum.
- Heatley, A., & Nation, P. (1994). *Range*. Victoria University of Wellington, NZ. [Computer program, available at <http://www.vuw.ac.nz/lals/>.]
- Holten, C. (2002). Charting new territory: Creating an interdepartmental course for Generation 1.5 writers. *The CATESOL Journal*, *14*, 173-189.
- Holten, C., & Mikesell, L. (2007). Strategies to address the lexicogrammatical development of Generation 1.5 ESL writing. *The CATESOL Journal*, *19*, 35-52.
- Ignash, J. M. (2000). Linguistic diversity, immigration, and today's undergraduate demographics. In J. W. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 3-27) Mahwah, NJ: Erlbaum.
- Jarvis, S., Grant, L., Bikowski, D., & Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing*, *12*, 377-403.
- Leki, I. (1999). "Pretty much I screwed up:" Ill-served needs of a permanent resident. In L. Harklau, M. Kay, & M. Siegal, (Eds.), *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL* (pp. 17-45). Mahwah, NJ: Erlbaum.
- Levi, E. I. (2004). *A study of linguistic and rhetorical features in the writing of non-English language background graduates of U.S. schools*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.
- Matsuda, P. K. (2008). International and U.S. resident ESL writers cannot be taught in the same class. In J. Reid (Ed.), *Writing myths: Applying second*

- language research to classroom teaching (pp. 159–176). Ann Arbor: University of Michigan Press.
- Matsuda, P. K., & Matsuda, A. (2009). The erasure of resident ESL writers. In M. Roberge, M. Siegal, & L. Harklau (Eds.), *Generation 1.5 in college composition* (pp. 50-64). New York: Routledge.
- McNamara, D. S., Louwerse, M. M., Cai, Z., & Graesser, A. (2005). *Coh-Metrix version 1.4*. Retrieved October 20, 2009, from <http://cohmetrix.memphis.edu>
- Mikesell, L. (2007). Differences between Generation 1.5 and English as a second language writers: A corpus-based comparison of past participle use in academic writing. *The CATESOL Journal*, 19, 7-29.
- Nakamaru, S. (2010). Lexical issues in writing center tutorials with international and US-educated multilingual writers. *Journal of Second Language Writing*, 19, 95-113.
- Nero, S. J. (1997). English is my native language ... Or so I believe. *TESOL Quarterly*, 31, 585-593.
- Patthey, G., Thomas-Spiegel, J., & Dillon, P. (2009). Educational pathways of Generation 1.5 students in community college writing courses. In M. Roberge, M. Siegal, & L. Harklau (Eds.), *Generation 1.5 in college composition* (pp. 135-149). New York, NY: Routledge.
- Reid, J. M. (1997). Which non-native speaker? Differences between international students and U.S. resident (language minority) students. *New Directions for Teaching and Learning*, 70, 17-27.
- Roberge, M. M. (2002). California's generation 1.5 immigrants: What experiences, characteristics, and needs do they bring to our English classes? *The CATESOL Journal*, 14, 107-129.
- Roberge, M., Siegal, M., & Harklau, L. (Eds.). (2009). *Generation 1.5 in college composition: Teaching academic writing to U.S.-educated learners of ESL*. New York: Routledge.
- Rodby, J. (1999). Contingent literacy: The social construction of writing for nonnative English-speaking college freshmen. In L. Harklau, M. Kay, & M. Siegal (Eds.), *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL* (pp. 45-60). Mahwah, NJ: Erlbaum.
- Rumbaut, R. G., & Ima, K. (1988). *The adaptation of Southeast Asian refugee youth: A comparative study*. (Final report to the Office of Resettlement). San Diego, CA: San Diego State University. (ERIC Document Reproduction Service No. ED 299372)
- Schwartz, G. G. (2006). *Vietnamese students in mainstream composition: An ethnographic study of academic identities in Generation 1.5 students who cross over*. Unpublished doctoral dissertation, University of Arizona, Tucson.
- Shaw, P., & Liu, E. T. (1998). What develops in the development of second-language writing? *Applied Linguistics*, 19, 225-254.
- Snow, M. A., & Kamhi-Stein, L. D. (1997). Teaching academic literacy skills: A new twist on the adjunct model. *Journal of Intensive English Studies*, 11, 93-208.

Valdez, G. (1992). Bilingual minorities and language issues in writing. *Written Communication*, 9, 85-136.

Valdez, G. (2001). *Learning and not learning English: Latino students in American schools*. New York: Teachers College Press.

Western Interstate Commission for Higher Education. (2008). *Knocking at the college door, 2008*. Boulder, CO: Western Interstate Commission for Higher Education.

West, M. (1953). *A general service list of English words*. London: Longman, Green.

Wurr, A. J. (2004). English studies and generation 1.5: Writing program administration at the crossroads. *The Reading Matrix*, 4, 14-23.

**Appendix A**  
**Biographical Information Survey**

1. How old are you?  
  
19 or Less    20-24    25-29    30-34    35-39    40-49    Over 50
  
2. Were your parents born in the US?  
  
Yes        No
  
3. Were you born in the US?  
  
Yes        No
  
4. If you were born in another country, how long have you been living in the United States?  
  
Less than 5 years    5-10 years    10-15 years    Over 15 years
  
5. With your family (or in the house where you spent most of your childhood), do/did you regularly speak a language other than English?  
  
Yes        No (If no, skip to question # 7)  
  
If yes, what language? \_\_\_\_\_  
  
Hmong    Spanish    Russian    Armenian    Punjabi    Hindi    Other
  
6. In this language other than English, how many years of formal (school) education have you had?  
  
None    1-3 years    3-5 years    5-8 years    More than 8

7. Outside of the home, what percentage of your day do you use this non-English language?

5-15% 15-30% 30-50% 50-75% 75-95%

8. How old were you when you started learning English?

1-3 years old 3-5 years old 5-8 years old 8-16 years old 17 years+

9. How many years of formal (school) education have you received in the United States?

1-3 years 3-5 years 5-10 years More than 10

10. Did you graduate from high school in the US?

11. How would you describe your speaking and listening abilities in the non-English language listed above?

Very limited Weak Good Very Good Excellent

12. How would you describe your reading and writing abilities in the non-English language listed above?

Very limited Weak Good Very Good Excellent

13. How would you describe your speaking and listening abilities in English?

Very limited Weak Good Very Good Excellent

14. How would you describe your reading and writing abilities in English?

Very limited Weak Good Very Good Excellent



## Appendix B Holistic Scoring Guide

- 6** An essay at this level
- effectively addresses the writing task
  - is well organized and well developed
  - uses clearly appropriate details to support a thesis or illustrate ideas
  - displays consistent facility in use of language
  - demonstrates syntactic variety and appropriate word choice though it may have occasional errors
- 5** An essay at this level
- may address some parts of the task more effectively than others
  - is generally well organized and developed
  - uses details to support a thesis or illustrate an idea
  - displays facility in the use of language
  - demonstrates some syntactic variety and range of vocabulary, though it will probably have occasional errors
- 4** An essay at this level
- addresses the writing topic adequately but may slight parts of the task
  - is adequately organized and developed
  - uses some details to support a thesis or illustrate an idea
  - demonstrates adequate but possibly inconsistent facility with syntax and usage
  - may contain some errors that occasionally obscure meaning
- 3** An essay at this level may reveal one or more of the following weaknesses:
- inadequate organization or development
  - inappropriate or insufficient details to support or illustrate generalizations
  - a noticeably inappropriate choice of words or word forms
  - an accumulation of errors in sentence structure and/or usage
- 2** An essay at this level is seriously flawed by *one* or more of the following
- serious disorganization or underdevelopment
  - little or no detail, or irrelevant specifics
  - serious and frequent errors in sentence structure or usage
  - serious problems with focus
- 1** An essay at this level
- may be incoherent
  - may be undeveloped
  - may contain severe and persistent writing errors
- 0** A paper is rated 0 if it contains no response, merely copies the topic, is off-topic, is written in a foreign language, or consists of only keystroke characters.

## Appendix C

### L2 Studies with Linguistic Variables that Distinguish Student Writing

<i>Author</i>	<i>Task</i>	<i>Students</i>	<i>Types of comparisons</i>	<i>Features identifying more advanced L2 writing</i>
Shaw & Liu (1998)	Two 30-minute placement essays (descriptive and persuasive) completed pre- and posttest	EAP undergraduate students	Pre- versus post (2 to 3 months apart) within-group comparison	Reduction of contractions, increase in subordinate clauses, passives, and connectors.
Ferris (1994)	35-minute writing prompt about “the effects of culture shock”	160 ESL students’ placement compositions of four L1 groups (Arabic, Chinese, Japanese, Spanish)	Holistic scores separating the low versus high essays	Higher-scored essays showed greater text length, fewer passive constructions, more nominal structures, more adverbials, more prepositional phrases
Grant & Ginther (2000)	30-minute TWE writing assignment	178 TWE examinees of three L1 groups (Arabic, Chinese, Spanish)	Groups correspond to TWE proficiency scores of 3, 4, and 5.	Higher-scored essays showed greater use of nominalizations, prepositions, and modals (after normalizing for text length)*
Jarvis, Grant, Bikowski, & Ferris (2003)	Combined data taken from Ferris (1994) and Grant and Ginther (2000)	Combined data taken from Ferris (1994) and Grant and Ginther (2000)	Cluster analysis done on high-rated essays. Clusters of linguistic variables formed groups.	<u>Significant in all clusters</u> text length, diversity of vocabulary, more emphatics, downtoners, and adverbials <u>Significant in some clusters but not others</u> More nominalizations, nouns, impersonal pronoun <i>it</i> , prepositions, adverbial subordination, longer mean word length
Crossley & McNamara (2009)	500- to 1,000-word take-home essays, mostly argumentative covering four prompts	(a) 195 L1 Spanish ESL students (ICLE corpus), and (b) 208 university L1 freshmen	Used Coh-Metrix to analyze coreferentiality, word frequency measures, word information measures, hypernymy and polysemy values, spatiality, causality	L1 writers obtained significantly higher values in hypernymy, polysemy, argument overlap, motion verbs, age of acquisition words, latent semantic analysis givenness, word meaningfulness, and incidence of causal verbs. L2 writers used significantly more locational nouns than L1 writers.

\*The author of the current study performed the reanalysis of data from Grant and Ginther (2000).

**Appendix D**  
**Descriptive Statistics and Mann-Whitney *U* Statistics**  
**for Language Variables**

<i>Language variable</i>	<i>Mean &amp; SD</i>		<i>Mean ranks</i>		<i>Stand. test statistic</i>	<i>Sig. (2-tailed)</i>
	Gen 1.5	L1 Texts	Gen 1.5	L1		
All connectives	88.07 (14.84)	87.10 (15.60)	31.98	29.00	-.615	.539
Content word overlap	.18 (.05)	.18 (.03)	31.51	29.95	-.323	.747
Pronoun ratio	.41 (.09)	.46 (.08)	27.27	38.65	2.351	.019
Words before main verb	4.19 (1.10)	4.78 (1.46)	28.74	35.62	1.421	.155
Conditional operators	7.83 (5.97)	8.68 (6.67)	30.21	32.62	.500	.617
Temporal cohesion	.74 (.10)	.72 (.08)	32.02	28.90	-.645	.519
NP incidents	282.78 (18.30)	279.76 (17.66)	31.85	29.25	-.538	.591
Modifiers per NP	.62 (.12)	.64 (.10)	29.91	33.22	.684	.494
Noun hypernym	4.78 (.29)	4.74 (.35)	31.71	29.55	-.446	.656
Verb hypernym	1.32 (.13)	1.30 (.10)	32.06	28.82	-.668	.504
“No. of words”	592.83 (180.26)	519.25 (121.93)	33.52	25.82	-1.590	.112
“No. of paragraphs”	6.51 (.90)	6.30 (.57)	31.82	29.32	-.581	.561
Sentences per paragraph	5.36 (1.82)	4.60 (1.03)	33.59	25.70	-1.629	.103
“No. of sentences”	35.10 (12.75)	29.10 (7.77)	34.11	24.62	-1.961	.05
Words per sentence	17.43 (3.11)	18.21 (3.43)	30.02	33.00	.615	.539
Syllables per word	1.34 (.07)	1.32 (.06)	32.41	28.10	-.891	.373
Prep phrases	14.74 (2.89)	14.93 (3.31)	30.61	31.80	.246	.806
Advl Sub	4.76 (2.10)	4.65 (2.04)	31.12	30.75	-.077	.939
Modals	7.48 (3.15)	7.35 (2.43)	31.00	31.00	.000	1.000