1 Comparing international and US resident second language learners' performances in five domains of writing

Kristen di Gennaro
Pace University, New York City, US

Motivation for the research

Higher education programs in the US typically require students to complete courses in writing or composition, that is, courses in which students focus on writing for university-level coursework. To accommodate the large numbers of students in these classes whose first language (L1) is not English, several institutions offer composition courses specifically for second language (L2) writers. Depending on the institution's placement procedures, L2 students enrolling in such courses may be international students who have recently arrived in the US with student visas after having completed secondary education in their home countries, or they may be long-term residents of the US who have completed secondary (or even primary) education in US schools. This group is often referred to as Generation 1.5. Most institutions of higher education enroll students in both categories. Recent statistics on international student enrollment indicate a dramatic rise in the number of international L2 students in US colleges and universities, for example the 'Open Doors Data' from the Institute of International Education (2014). Statistics on school-age students reporting a primary language other than English indicate a rise in US resident L2 learners as well (National Clearinghouse for English Language Acquisition 2008). Based on these demographics, L2 composition courses appear to serve an increasingly heterogeneous student body (di Gennaro 2012). Whether such differences warrant the creation of separate courses for different types of L2 learners is an ongoing debate within the community of L2 composition scholars (di Gennaro 2012, 2013, Doolan 2013, 2014, Matsuda 2008).

Literature review

International L2 learners, that is, students who have arrived in the US after having completed secondary education in their home countries, are likely

to have acquired much of their knowledge of English in classroom environments. Traditionally, such instruction has been described as favoring written English and formal grammar instruction over spoken English and conversational fluency. International learners' completion of secondary school in their home countries assumes advanced literacy in their L1s which, some scholars suggest, may lead them to transfer grammatical rules and organizational preferences from their L1s into their written academic English (see Reid 2006, Thonus 2003). Conversely, US residents who are L2 learners most often attend US post-secondary institutions after having completed secondary or even primary education in the US, thereby, having acquired English primarily through immersion in English-speaking environments, including academic classroom experiences. These experiences result in conversational fluency, as well as familiarity with process approaches to teaching writing. Nevertheless, they may lack an awareness of the differences between informal and academic registers (di Gennaro 2008, 2009). Given differences in prior instruction and exposure to English, L2 writing scholars (see Ferris 2009, Reid 2006, Roberge, Siegal and Harklau 2009) claim that international and resident L2 learners will have noticeably different strengths and weaknesses in their writing abilities, and thus require different types of instruction to improve their writing (see Mikesell 2007).

Until very recently, empirical support for such claims was scant and limited to small-scale studies. For example, in a detailed case study, Leki (1999) noted that the US resident L2 student who was the focus of her qualitative study excelled at informal communication in English, yet did not do well in courses where the focus was on using grammar skills in writing. Frodesen and Starna (1999), who examined multiple writing samples and conducted interviews with two students over the course of several years, suggested that their participants' errors reflected their different backgrounds, and recommended different courses of action for each student. Specifically, they recommended that the long-term US resident L2 learner in their study could benefit more from a mainstream (i.e. L1) composition course than a course directed at L2 learners. The other student, who was a recently arrived L2 learner, would likely prefer a composition course created specifically for L2 learners. Bosher (1998), whose data included interviews, stimulated recalls and text analyses of three students, found that the international L2 participant in her study attended more to content and organization in writing than to other aspects of writing proficiency; conversely, the representative resident L2 participant attended more to surface-level language issues and generating text, and did not appear to focus much on content, discourse, or the overall purpose of the text. A second resident L2 learner demonstrated writing processes more similar to the international participant than to the other resident learner. While these studies are limited in their sample sizes, their findings highlight that L2 students placed in similar composition courses may have

very different strengths and weaknesses in completing academic writing tasks.

Moving beyond small-scale case studies, Bosher and Rowekamp (1998) examined a series of factors to determine which of them best predicted resident L2 writers' success compared to those of international learners in post-secondary education. The study included 56 participants divided into two groups based on whether they had completed secondary school in the US or in their home countries. Participants who had completed secondary school in their home countries scored significantly higher on the objective section of the Michigan English Language Assessment Battery, while the US secondary school graduates scored significantly higher on the listening section. No significant differences appeared across the two groups in terms of their composition scores. In a similar study, Muchinsky and Tangren (1999) found that their 13 resident L2 participants excelled on the Michigan Test of Aural Comprehension, while their nine international L2 learners scored equally well on this test and the Michigan Test of English Language Proficiency, and significantly better than the resident group for the latter test. In terms of participants' writing scores, the international L2 group's scores were higher than those of the resident L2 group. Such studies indicate that differences exist across the two groups in terms of their academic strengths and weaknesses; however, the differences with regard to their writing are inconsistent.

Perhaps the first large-scale, systematic quantitative study comparing international and resident L2 learners' writing is Levi's (2004) unpublished doctoral dissertation. Based on analyses of 140 participants' writing in search of statistically significant differences in errors in writing, Levi found that international L2 and resident L2 participants produced similar numbers of lexico-grammatical and rhetorical errors. When errors were divided into subcategories, however, differences between the two groups emerged. Similarly, Mikesell (2007) compared grammatical error patterns across international L2 and resident L2 students' writing samples with a specific focus on past participle errors. Mikesell found that both groups produced the same percentage of errors, but they differed in terms of error types. When linguistic context was taken into account, the international L2 learners' errors stemmed mainly from producing the correct form but using it in an inappropriate context, while the resident L2 learners' errors were related primarily to producing an incorrect form. Continuing this line of research focusing on errors, Doolan (2013, 2014) also found statistically significant differences between international and resident L2 groups in terms of error patterns in their writing. Interestingly, while results from these quantitative studies suggest that international and resident L2 learners differ with regard to their writing, the researchers propose different solutions: Levi (2004) recommends creating writing courses for resident L2 learners separate from those for international learners, Mikesell (2007) proposes different types of grammar instruction for

each, and Doolan (2013, 2014) recommends treating resident L2 learners as native English speakers and not as L2 learners at all.

While the studies by Levi (2004), Mikesell (2007), and Doolan (2013, 2014) provide much-needed empirical evidence regarding differences in the writing ability of international and resident L2 writers, they are inadequate. By focusing on learners' errors, such studies are limited in that they reflect a deficit perspective of each group's writing ability, highlighting learners' shortcomings rather than their potential strengths. Perhaps more importantly from an assessment perspective, by focusing almost exclusively on learners' grammatical performance in writing, these studies reflect an impoverished construct definition of writing ability. A more holistic view of learners' writing would permit a focus on each group's strengths and weaknesses. Moreover, it is possible that the two groups differ in aspects of their writing other than in grammatical (in)accuracies.

Adopting a comprehensive construct definition of writing ability, di Gennaro (2009) examined writing placement samples from 97 students (54 international L2 and 43 resident L2) who were scored on five different components of writing ability (grammatical, cohesive, rhetorical, sociolinguistic, and content control) along with essay length. Results showed the two groups differed only with regard to rhetorical control and essay length. While this study reflects an improvement over previous studies in terms of construct representativeness, the definitions of the two learner groups could have been more rigorous. Specifically, international and resident L2 learners were distinguished only in terms of location of high school completion, without consideration of participants' length of residence in the US, which allowed some long-term resident participants who had completed high school overseas to qualify as international L2 participants. More recently, di Gennaro (2013) adopted a more precise distinction between the two groups, including both high school location and length of residence, along with a fine-tuned definition of writing ability. The latter study, which included 134 participants (67 in each group), found that the international L2 learners scored slightly higher than the resident L2 learners in overall writing ability, and a bias analysis revealed that the two groups differed statistically only with regard to grammatical control. When the two groups were analyzed separately, results showed they had opposing strengths and weaknesses in grammatical control and sociopragmatic control. Grammatical control resulted in being the easiest among the five components for the international L2 group, and sociopragmatic control was the easiest for the resident L2 learners. Grammatical control was the second-most difficult component for the resident L2 learners, as was sociopragmatic control for the international L2 learners. Based on these findings, di Gennaro (2013) agrees with Levi (2004) and Mikesell (2007) in concluding that both international and resident L2 learners demonstrate a need for L2 writing instruction at the post-secondary level. Rather than segregate resident

L2 learners from international L2 students as these scholars propose, di Gennaro (2013) agrees with Matsuda (2008) in suggesting that programs can provide instruction relevant to both types of L2 learners in the same courses.

Research questions

The current study builds upon di Gennaro (2013) in that it continues the search for empirical evidence confirming (or not) that differences exist in the writing ability of international and resident L2 participants. Drawing upon the same dataset as in di Gennaro (2013), the current study subjected the data to additional analyses, providing another opportunity for differences (or similarities) to emerge. Only by analyzing both groups together for each individual component can such expectations be confirmed (or not). Thus, for the current study, five whole-group analyses were conducted: one for each individual component of writing ability. Examining results from whole-group analyses for each component can reveal how each group performed with respect to the other for each individual component, rather than how each group performed with respect to itself across all five components (as in di Gennaro 2013). The research questions addressed in the current study were:

1. How does the writing performance of international L2 writers compare to that of resident L2 writers in five separate components of writing ability, namely grammatical, cohesive, rhetorical, sociopragmatic, and content control?
2. What implications do the findings have for writing program administrators in post-secondary writing contexts?

Data collection procedures

Participants

Studies comparing international and resident L2 learners typically differentiate the two groups based on participants' educational background (see Bosher and Rowekamp 1998, di Gennaro 2009, Doolan 2013, Levi 2004, Muchinsky and Tangren 1999) or length of residence (see Bitchener and Knoch 2008, Connerty 2009). To strengthen the distinction between the two groups, the current study used both criteria for classifying participants: international L2 participants had completed high school in their home countries and lived in the US for a maximum of three years; resident L2 participants had completed high school and resided in the US for a minimum of three years. These criteria guaranteed that no participant could qualify for both groups. Participants who met one criterion but not the other were excluded from the study.

A total of 134 learners were included in the study: 67 international and 67 resident L2 learners. Participants represented 29 different L1s, with the most prominent being Chinese (61), Spanish (11), Korean (10), Russian (7), and Arabic (7). The median length of residence was less than one year for the international group and six years for the resident group. The median age was 19 years for both groups, as all participants were first-year students at the same post-secondary institution in the US.

Three experienced instructors of post-secondary writing courses for L2 students served as raters. All raters had graduate degrees in Teaching English to Speakers of Other Languages (TESOL) or Applied Linguistics, and had rated placement exams for students entering post-secondary writing courses for several years.

Instruments

All participants responded to the same prompt, instructing them to write an argument essay for or against the point of view that anyone who wants to attend college should be accepted. Five rubrics were developed to score participants' responses, one for each component of the writing construct (see the Appendix). Grammatical control referred to a writer's adherence to lexical and morphosyntactic rules at the sentence level. Moving beyond the sentence, cohesive control referred to the writer's ability to overtly connect ideas within and across clauses and sentences. Rhetorical control differed from cohesive control in that it referred to the writer's ability to organize ideas and supporting information at the discourse level rather than at the sentence level. Organizational cues considered part of rhetorical control might not be overt, as they are in cohesive control. Sociopragmatic control encompassed features categorized as sociolinguistic or pragmatic awareness; that is, it was related to the writer's choice of register, stance and tone within the context of the writing task. Finally, content control was defined as the extent to which a writer elaborated on the topic by providing supporting evidence of the type expected in post-secondary writing contexts.

Participants produced writing samples in class, as part of first-day procedures in their composition courses. They had 45 minutes to read and respond to the writing prompt. Participants were asked if they would be willing to share their writing samples as part of the current study. Participants who agreed completed a demographic information form and signed a consent form.

Three raters were trained to use the five rubrics designed for the current study to evaluate participants' writing and assign scores from 0–5 for each participant and in each of the five components separately. Raters did not have access to information about test takers' backgrounds during the rating process. Following procedures for a fully crossed rating design, each rater read and evaluated all 134 essays in all five components. The sum of the

five component scores from each rater produced three composite ratings per participant. These summed ratings were then averaged to yield one score for each participant. Pearson product-moment correlations for raters ranged from 0.679 to 0.818. While moderate, all correlations were statistically significant at the 0.01 level. Since correlations only refer to agreement of examinee rankings and not agreement in the actual scores, Cronbach's coefficient alpha was also calculated as an additional reliability estimate. The resulting alpha value of 0.898 for the overall group of participants indicated very high internal consistency reliability for the ratings.

Data analysis

The analysis of writing ability is problematic, as raters' judgments are needed to evaluate participants' writing, and yet raters are not part of the writing construct. To account for such construct-irrelevant factors, many researchers of writing ability adopt a many-facet Rasch measurement (MFRM) as a statistical tool because MFRM calculates participants' ability levels after taking into account external factors, such as rater severity; therefore it produces a more accurate depiction of educational performance than inferential statistics do. Another advantage of MFRM is that it transforms participants' scores from ordinal scales to equal-interval scales, a process that inferential statistical procedures cannot do (Bond and Fox 2007). MFRM is also considered sample-independent, allowing findings to be generalizable to a larger population (Sudweeks, Reeve and Bradshaw 2005). For these reasons, MFRM was used to analyze and compare the two groups' writing performances for the current study.

Five separate MFRM analyses were performed: one for each individual component of the writing construct. Analyses were conducted with the FACETS computer program (Linacre 2009), which converts participants' raw scores into an equal-interval logit scale for each component. The resulting logit scales allowed for comparisons across groups within each component because participants' converted scores have the same frame of reference (Bond and Fox 2007).

Results

For each MFRM analysis, the FACETS program produces a visual summary in the form of a map, illustrating the dispersion of data. Maps produced from the analyses in the current study are presented in Figures 1.1 to 1.5, which include summaries for grammatical, cohesive, rhetorical, sociopragmatic, and content control. The column on the left in each figure is the equal-interval logit scale that is produced after all facets of the measurement procedure have been taken into account. Depending on participants' performance,

the logit scale may have a greater or smaller range. Indeed, the logit scale in Figure 1.1, representing grammatical control, has a greater range (−10 to 10) than the scale in Figure 1.2 representing cohesive control (−8 to 6).

The wider of the two columns in Figures 1.1 to 1.5 displays the 134 participants. Each 'I' or 'R' represents one participant: participants identified with 'I' are international L2 learners; participants identified with 'R' are resident L2 learners. Participants' placement in this column corresponds with each one's logit score, or ability level, for that component. Participants placed higher in the column are described as having greater ability than participants placed lower in the column. For example, in Figure 1.1, the 'I' and 'R' at the top of the participant column indicate that each group had one participant who stood out as having greater ability in the component of grammatical control than the rest of the group. The three 'R's at the bottom of the same column indicate that the three participants with the lowest scores (and, therefore, the least ability) in grammatical control were all resident L2 participants.

Figure 1.1 FACETS summary for grammatical control

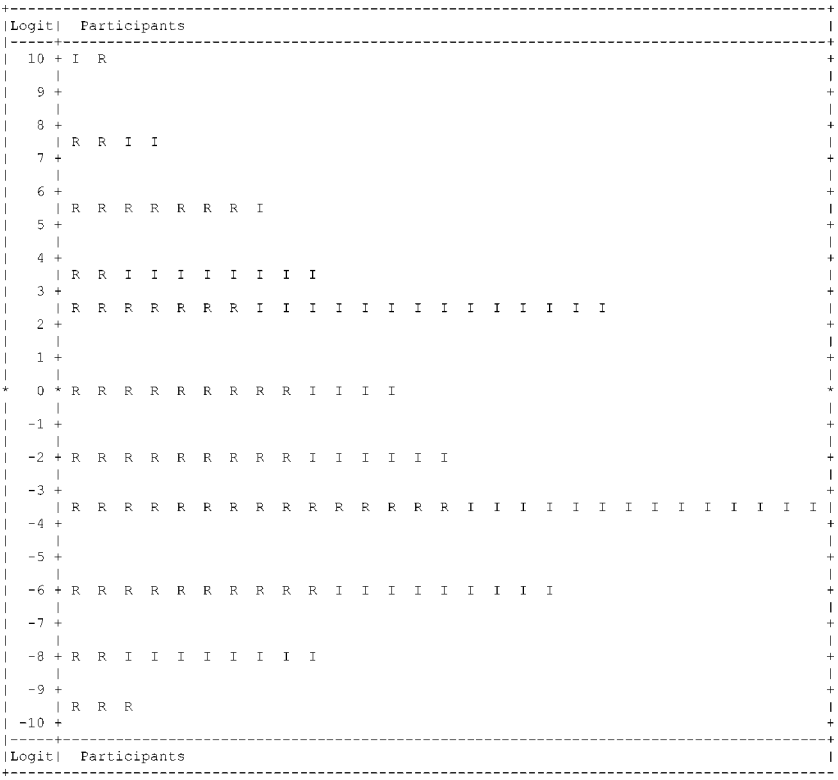
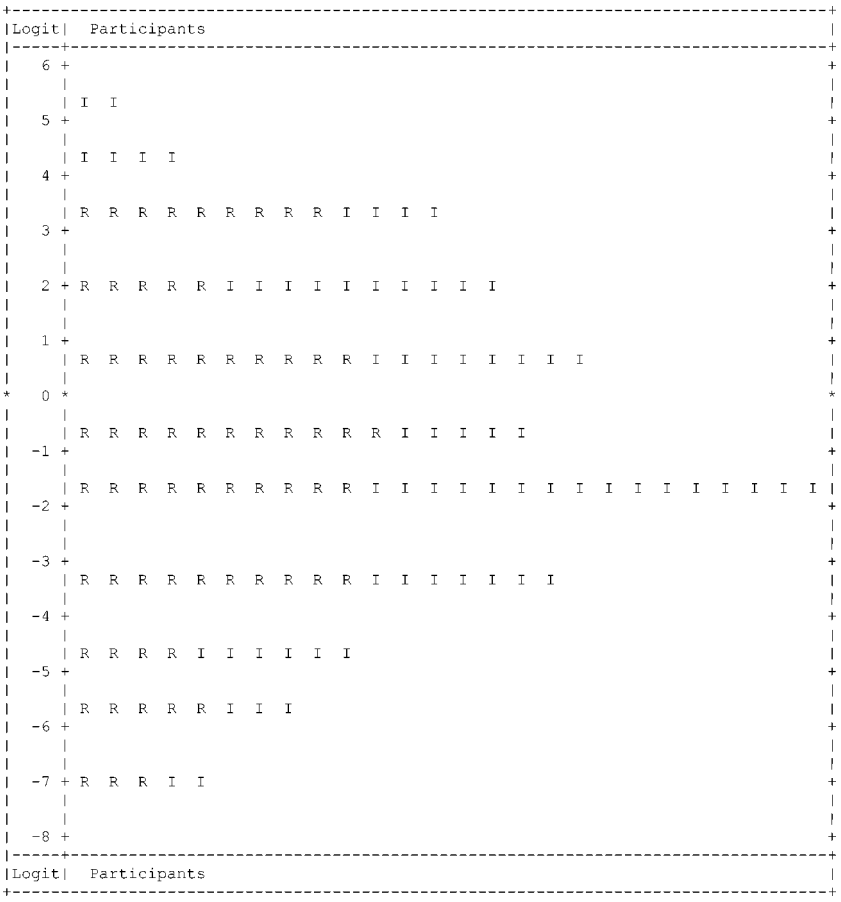


Figure 1.2 FACETS summary for cohesive control



A glance at the maps for each component reveals that representatives from the international L2 group are consistently among the highest scoring participants in each component, a position shared with representatives from the resident group for the components of grammatical, sociopragmatic, and content control. Conversely, the lowest scoring participants on each map are consistently from the resident L2 group, with international participants sharing this position for the components of cohesive and sociopragmatic control. Apart from the extreme scores, participants from both groups achieved a wide range of logit scores, with neither group appearing particularly stronger or weaker than the other. The majority of participants from both groups cluster in the middle of the scale, indicating a normal distribution of scores.

Figure 1.3 FACETS summary for rhetorical control

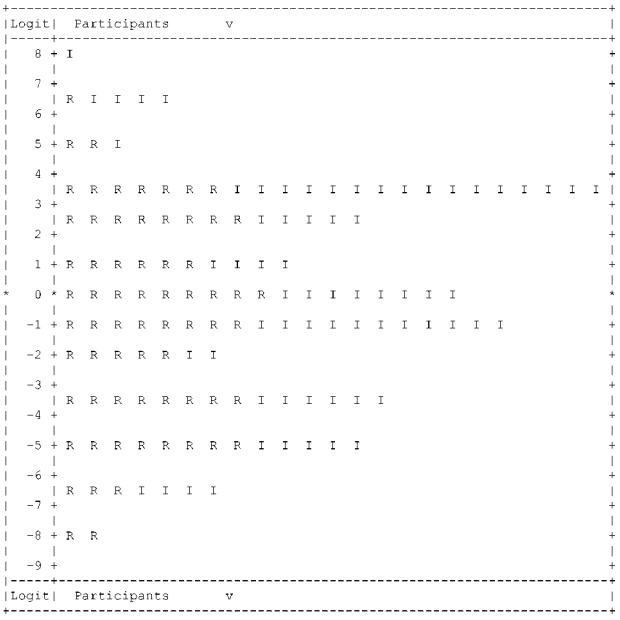


Figure 1.4 FACETS summary for sociopragmatic control