

# Effects of task type practice on the Japanese EFL university student's writing from a dynamic systems perspective: A longitudinal study utilizing multi-level text analysis

**Kyoko BABA & Ryo NITTA\***

Kinjo Gakuin University  
\*Nagoya Gakuin University

## **Abstract**

The longitudinal effects of repeating a timed classroom writing activity on the EFL (English as a Foreign Language) student's L2 (second language) writing development were investigated. Data for 46 students were collected 30 times over one year. The students' compositions were analyzed for fluency and for grammatical and lexical complexity. Text analysis using Coh-Metrix and visual analysis of non-linear individual developmental patterns showed that task repetition has an overall effect on L2 writing development. Grammatical complexity developed more prominently than the other aspects of writing. This counter-intuitive result points to the significance of the writers' reflective consciousness towards their own writing.

## **Introduction**

This study investigated the longitudinal effects of repeating a timed writing activity on the EFL (English as a Foreign Language) student's L2 (second language) writing development. Studies on task repetition have shown that the quality of their spoken language improves when the same task is repeated, because learners can pay focal attention to form the second time (Bygate & Samuda, 2005). On the other hand, it is unclear what effects repetition of a writing task brings about, assuming that learners are less pressured in writing.

Unlike most previous studies, this exploratory study throws light on individual developmental patterns, focusing on the non-linearity of writing development. We assessed which of the three principal aspects of L2 writing (fluency, syntactic complexity, and lexical complexity) develop in one year. The textual features that changed significantly during the period in terms of dynamic developmental patterns are examined with a graphical method.

### **Longitudinal Research on L2 Writing**

The significance of longitudinal research on L2 learning is often emphasized (Ortega & Ibarra-Shea, 2005), but there is still a dearth of such research. Research on L2 writing development is no exception, although many studies have identified various characteristics of L2 writing at different developmental stages (Hinkel, 2002; Polio, 2001; Wolfe-Quintero, et al., 1998). With the development of sophisticated computer programs, notable features of L2 writing have been illuminated with a large corpus of L2 texts (e.g., Crossley & McNamara, 2009; Jarvis et al., 2003; McCarthy et al., 2007). However, scant attention has been paid to how these features change over time.

There have been a few longitudinal studies on L2 writing, but they were often either cross-sectional (Henry, 1996), did not look at a variety of text features (Bardovi-Harig, 2002), or focused on a small number of writers (Larsen-Freeman, 2006; Verspoor, et al., 2008). In addition, most studies have dealt with “L2 writers” as a whole, as if there were an average L2 writer at different developmental stages. For educational purposes, however, it is no less important to know how each individual develops than to know how L2 writers in general develop. This is because it is unlikely that every L2 writer develops the same writing features at the same speed. Focusing simultaneously on individual learners and on learners as a whole is an approach that is congruent with the recently advocated complexity systems perspective, which sees both the forest and the tree (see below).

To date, no L2 studies have looked at both group trends and individual developmental patterns, except that a series of longitudinal studies of L2 learner’s spontaneous speech have been conducted by Crossley and his colleagues (Crossley, et al., 2008, 2009, 2010, 2011, in press). For example, Crossley et al. (2009) investigated the development of hypernymic lexical knowledge in the speech of six learners, who were interviewed every two weeks over one year. In their analysis, they examined not only whether there were significant differences between times for this group of learners, but also whether each of the six learners fit the group trend.

We carried Crossley and his colleagues’ idea one step further in this study. We analyzed the developmental trajectories of individuals in order to contribute to an understanding of the dynamic nature of L2 writing development, not simply to “control for random effects” (Crossley et al, 2009, p. 320).

### **The Dynamic Nature of L2 Writing Development**

Two-wave research designs have often been used to investigate the effects of a certain task or teaching method on L2 writing. Usually, a pretest and a posttest are conducted, and a t-test or ANOVA assesses the difference between the two time points (e.g., Polio et al., 1998; Shaw & Liu, 1998). However, while such an approach reveals

*whether* L2 writers as a whole have developed or not, it is impossible to know *how* they developed their writing. This is because the two-wave research design assumes linear development, which is often not the case with language learning.

To explore the *how* of L2 writing development, this study adopted a complex systems view, in which the non-linearity of language development is underscored (Larsen-Freeman & Cameron, 2008). The exploration of non-linear language development requires multiwave data (Willet, 1994). That is, data is collected from each participant multiple times, which enables us to plot individual growth trajectories.

Furthermore, with the complex systems approach, intra-individual variability has come to the fore. The intra-individual variability is disregarded or controlled for in two-wave statistical analysis that presumes linearity. However, it is a fruitful source of information from a complex systems perspective. To analyze the intra-individual variability, various (especially visual) tools have been proposed, such as a polynomial trendline, moving min-max graph, and detrended representation of L2 development (Verspoor, et al., 2008).

Since the investigation of intra-individual variability has just begun to draw attention in the L2 developmental field, there are only a few studies that have examined intra-individual development in L2 writing. For instance, Larsen-Freeman (2006) studied the development of five immigrant females' oral and written production over six months in terms of their fluency, accuracy, and complexity. As a group, there was overall progress in all aspects of the participants' writing. However, paths to development widely diverged from individual to individual concerning the rate of developmental speed and the aspects that developed. Likewise, Verspoor et al. (2008) conducted a case study of one ESL graduate student's writing development in terms of lexical and grammatical complexities over three years. One of the key findings of their study was the observation that there were three stages in his development. The first stage is stability at a low level, the second stage is a major fluctuation period, where his performance went up and down, and the third stage is stability at a high level.

In this study, the dynamic approach to the study of L2 writing development taken by Larsen-Freeman (2006) and Verspoor et al. (2008) is used and extended. Whereas the previous studies focused on a fairly small number of writers, the aim of this study is to make a slightly broader generalization about individual developmental patterns. As Crossley and his colleagues did with spoken data, we explore the group trends as well as individual trends with a larger number of participants than those in Larsen-Freeman and Verspoor's studies. Thus, the research questions of this study are: (1) Which textual features of students' writing change in one-year EFL courses through repeating the same writing task? and (2) If there are significant changes in their writing, how do individual students develop each aspect of their writing?

## Methods

Since this study is a part of a larger research project, we have collected other data not reported here. Such data was largely qualitative, involving questionnaires, post-task reflection, and interviews. Although the analysis of qualitative data provides us with important findings about the development of individual students' writing, in this paper we focus on linguistic analysis due to limited space.

## Participants

The study was conducted at two Japanese universities in two EFL classrooms, which are called Class One and Class Two. The classes we investigated have similarities and differences. Both courses were given for first-year English major students. Their initial English proficiency levels were similarly low with very little experience producing meaningful output in either speaking or writing. However, Classes One and Two had different course objectives; Class One focused on English writing, while Class Two was oriented more toward integrated skills of speaking, listening, and reading as well as writing. Initial enrollments for Class One and Class Two were 23 and 29, respectively, but as is common in real language classrooms, some students stopped attending in the middle of the course. Finally, 23 students in each class were engaged in the project regularly until the end of the year. Thus, the compositions of a total of 46 students were used for the analysis. The students had little experience writing in English before entering university, and, even in university, they had little chance to write in English outside the classroom. Therefore, any progress in their writing is attributable to the classroom writing activities.

## Task

We used a task called 'freewriting' (Elbow, 1998). As the name suggests, learners simply write without stopping for ten minutes. It is important to emphasize that the primary focus of the task is on meaning rather than form. This is writing, but it may be more appropriate to say that it lies somewhere between speaking and writing, because freewriting 'teaches you to write without thinking about writing', as we usually speak without thinking about speech (Elbow, 1998: 15).

## Data Collecting Schedule

The project was conducted over two academic terms consisting of thirty weeks in total. Although there were different course objectives between Classes One and Two, the exact same procedures were followed in both classes. In every class meeting, the students were given a topic list of three different topics. For example, students could choose from (1) Discuss some recent news event; (2) Do you want to live in a large

city? Explain why or why not; and (3) If the average life span were increased to 150 years, what major changes in society would you expect?

The aim of offering three topic alternatives is to compensate for differences in the students' individual experiences and preferences. Immediately after writing a composition, students were asked to rate the topic difficulty so that we could check the influence of the difficulty on their writing. The students were also requested to write reflective comments on their writing in Japanese (L1) about, for example, what they found difficult to write, what they thought during writing, etc. The collected writing was checked by the teachers and returned with some feedback in the next week. The aim of the feedback was to create a sense of audience and to maintain students' motivation for writing every week. Since the freewriting focuses on meaning rather than form, linguistic correction was not offered.

The same list of three topics was used for two weeks. The week after a new topic list was given, students were required to write about the same chosen topic to identify the task practice effects. Then, the next week, another new topic list was given.

### **Text Measurements**

There have been four main categories of measures used to study the development in the quality of texts by L2 writers: fluency, accuracy, and lexical and grammatical complexity (Wolf-Quintero et al., 1998). Since we did not focus on accuracy (we neither told our students to pay attention to grammatical and lexical accuracy nor corrected errors in their writing), we looked at the other three aspects of text quality. In particular, we had expected that fluency would improve considerably, as task repetition research has shown with spoken data.

To analyze the approximately 1500 compositions that were written by the students over one year, we used the web-based computational tool *Coh-Metrix* (Graesser et al., 2004) available at <http://cohmetrix.memphis.edu/cohmetrixpr/index.html> (version 2.0). Spelling mistakes in the compositions were carefully corrected manually and also with a spelling checker for analytical purposes. We selected the six most relevant measures of fluency and lexical and grammatical complexity as follows.

**Fluency.** We used three measures of fluency: (1) number of words per composition (text length), (2) average number of words per sentence (sentence length), and (3) latent semantic analysis (LSA). The first two are common measures of fluency in L2 writing research, but LSA needs some explanation. LSA evaluates the similarity of meaning between words, sentences, and passages by analyzing large corpora. What LSA measures may be debatable, but a series of studies conducted by Crossley and his colleagues used it as a measure of lexical proficiency (Crossley et al., 2008, in press). However, we used it as a measure of fluency because similarity of meaning across

sentences and paragraphs is likely to contribute to coherence in text, which is an important aspect of fluency. LSA is thus a significant additional measure of fluency that is different from mere counting of words. There were three LSA measures on Coh-Metrix, but we chose LSA of all combinations of sentences.

**Lexical Complexity.** The operationalization of lexical complexity varies with the focus of study, but it usually denotes either lexical sophistication or lexical diversity in text. Therefore, we used two measures of lexical complexity: (4) word frequency values from the CELEX corpus (Crossley et al., 2010), and (5) type-token ratio (TTR).

**Grammatical Complexity.** To measure grammatical complexity, we used (6) sentence syntax similarity (STRUT, all sentences across paragraphs) (McCarthy et al., 2009). STRUT gauges the degree of similarity in syntactic structures of sentences in a passage by comparing syntactic trees of each pair of sentences. If a variety of sentence structures is used in a passage, its STRUT value will decrease. So a lower STRUT value possibly indicates greater grammatical complexity.

### Results

A one-way multivariate analysis of variance (MANOVA) was conducted to assess whether the six text measurements significantly changed over one year. The MANOVA compared the compositions that were written in the first and the last weeks (Weeks 1 and 30). Significant differences were found among the first and the last compositions on the dependent measures, Wilks's  $\Lambda = .68$ ,  $F(6, 85) = 6.53$ ,  $p < .001$ . The multivariate partial  $\eta^2$  based on Wilks's  $\Lambda$  was strong, .32. Thus, it was shown that there was a significant change in the students' writing over one year.

Analyses of variances (ANOVAs) on each dependent variable were conducted as follow-up tests to the MANOVA (Table 1). To control for Type I error, we used the Holm method (Jaccard & Guilamo-Ramos, 2002). Two measures turned out to be significant: sentence length,  $F(1, 90) = 18.26$ ,  $p < .001$ , and sentence syntax similarity,  $F(1, 90) = 13.98$ ,  $p < .001$ . Text length was not significant, but close to the significance level,  $F(1, 90) = 6.26$ ,  $p = .01$ .

Table 1  
*Multivariate and Univariate Analysis of Variance F Ratios for Task Practice Effects for Text Measures*

Variable	ANOVA						
	MANOVA	TextL	SentenceL	LSA	TTR	WFreq	STRUT
	F(6,85)	F(1,90)	F(1,90)	F(1,90)	F(1,90)	F(1,90)	F(1,90)
Time	6.53***	6.26*	18.26***	.02	.69	.09	13.98***

Note. *F* ratios are Wilks's approximation of *F*s. MANOVA = multivariate analysis of variance; ANOVA = univariate analysis of variance; TextL = text length; SentenceL = average sentence length; LSA = latent semantic analysis (all combinations of sentences); TTR = type-token ratio; Wfreq = raw frequency of content words; STRUT = sentence syntax similarity (all sentences across paragraphs).  
\* $p < .05$ . \*\*\* $p < .001$

The results of the ANOVAs suggest that the students improved fluency (sentence length) and syntactic complexity (STRUT) more markedly than the other text features. It was no surprise that they became more fluent in writing; however, it was somewhat unexpected that syntactic complexity, rather than lexical complexity, improved substantially.

While the MANOVA and the ANOVAs examined the general trends for this group of students, a further analysis was conducted to illustrate how each student developed (or did not develop) the two text measures (sentence length and syntactic complexity). First, we drew a line graph for each text measure for all individual students. These graphs simply display the change of the measure in the individual students' writing over one year. Next, we added a polynomial trendline of the third degree (Verspoor et al., 2008) to clarify the general trajectory of each student's development. Last, by closely examining the trendlines, we identified five line patterns of development: (1) Reverse-S, (2) S, (3) Mound, (4) U, and (5) Line. All the patterns except the Line are non-linear (see Figure 1). Since we focused on the line patterns for this analysis, we ignored the differences in the units of analysis and ranges of scores. In addition to the line patterns, we also identified whether each trendline has an upward, stagnant, or downward trend. For example, the upper left graph in Figure 1 has a Reverse-S pattern with an upward trend. Thus, there were 15 developmental patterns (five line patterns times three up-and-down trends). The authors separately categorized all the students into the 15 patterns. Our inter-rater reliabilities were 83% for sentence length and 96% for syntactic similarity. After calculating the inter-rater reliabilities, we reached agreement on all discrepancies.

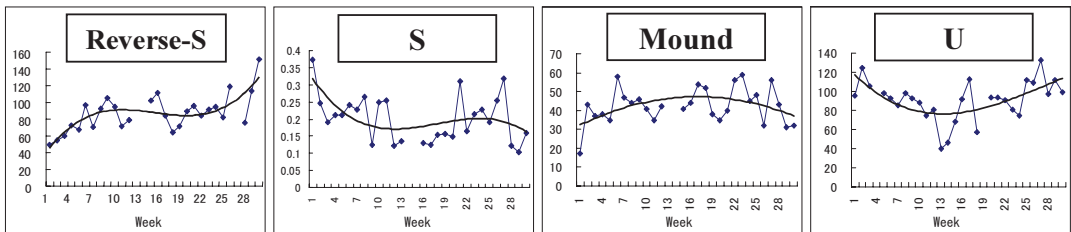


Figure 1. Four non-linear line patterns of development (samples).

The numbers of students who were categorized into each developmental pattern are tabulated for sentence length and syntactic similarity in Tables 2 and 3. To make the two text measures comparable, the up-and-down trends are reversed for syntactic similarity.

The first thing to be noted is that the developmental patterns for most of the students are non-linear, even though we identified the patterns with the trendlines. When we look at the raw line graphs, the fluctuation of the development is more obvious.

Tables 2 and 3 also show that more than 80% of the students made progress in both text measures (37 students for sentence length and 41 students for syntactic similarity). For both measures, the largest number of students displayed a Reverse-S pattern with an upward trend. This may mean that in both measures, most students made progress at first, then there was a stagnant period, and then they progressed further.

Table 2  
*Number of Student Categories in Each Developmental Pattern for Average Number of Words per Sentence*

	RevS	S	Mound	U	Line
+	23	2	3	1	8
0	1		3	1	
-		2	1		1

Table 3  
*Number of Student Categories in Each Developmental Pattern for Sentence Syntax Similarity*

	RevS	S	Mound	U	Line
-	16	3	5	13	4
0	1				
+		3		1	

There were, however, differences between the two measures. First, there were nine students who did not make any progress (with a stagnant and a downward trend) in sentence length, while only four students did not make progress in syntactic similarity. This may suggest that it might have been easier for most students to improve syntactic complexity than writing fluency. Second, the distribution of each developmental pattern converged into one pattern (a Reverse-S pattern with an upward trend) for sentence length, but diverged for syntactic similarity. In particular, approximately 30% of the students (13 students) showed a U pattern with an upward trend. It is assumed that syntactic complexity in these students' writing did not improve constantly, and that there might have been a period when the students tended to use similar syntactic structures before they came to use a wider variety of syntactic structures.

### Discussion

This study showed that the EFL students improved their writing over one year, and that among the six fluency and complexity measures, they significantly developed sentence length and syntactic complexity. The analysis of individual developmental patterns also shows that more than 80% of the students made some progress in the two measures.

On the basis of findings from previous research in task-repetition effects on speech production, we predicted that writing fluency would improve the most out of the three principal aspects of writing skills (fluency, lexical complexity, and grammatical



complexity). However, there was not notable progress in text length (fluency in terms of speed) or LSA (fluency in terms of meaning coherence), though text length improved somewhat. This finding may suggest that the development of fluency in L2 writing may take more than one year, especially when the learners write only once a week in EFL settings.

Likewise, lexical complexity in the student's writing did not improve over one year. Thus, either more intensive writing training or a longer writing practice may be required to develop this aspect of writing. This result is compatible with the finding from Haastrup and Henriksen (2000), who showed that the building of a semantic network in young children in EFL settings was a time-consuming process, and that 15 out of 17 children in their study did not strengthen their semantic network over three years.

On the other hand, the series of longitudinal studies by Crossley and his colleagues found that adult ESL students studying in an English-speaking country significantly developed more complex semantic relations in their speech. Coupled with the previous findings from research on task type practice on speech, repetition of a task may have a more immediate effect on speech than on writing.

A more important finding of this study is that grammatical complexity was the aspect of writing that developed most over one year. Whereas sentence length, which was labeled as fluency, developed to a similar extent, it is possible that its development was partly due to the increased grammatical complexity. In fact, there was a strong correlation between sentence length and grammatical complexity:  $r = -.58$  for the first compositions, and  $r = -.61$  for the last ones. In addition, more students improved grammatical complexity than sentence length, which may indicate that grammatical complexity developed faster than sentence length.

Why did grammatical complexity develop significantly even though the authors did not give the students corrective feedback on their compositions, nor did we instruct them to heed grammatical structures for this freewriting task? The results of this study do not offer a decisive answer to this question. However, one plausible explanation is that grammatical complexity was easier to consciously improve than was fluency. For instance, it may be difficult for basic writers to consciously write longer, but it may be possible for them to use a wider variety of grammatical structures by intention. This might be why the students could develop grammatical complexity faster.

Examination of the qualitative data is beyond the scope of this paper, but our interviews with the students corroborated this explanation. The students who were interviewed after the one-year freewriting experience recalled that they had become more reflective about their own writing. They said they excursively engaged in this

task, but gradually became aware of the shortcomings of their own writing by repeating the task (e.g., too short, with too many grammatical and lexical errors). Only after that did they deliberately attempt to use newly learned grammatical structures, write longer compositions, and use writing strategies (e.g., compensatory strategies when they could not come up with an exact expression). Thus, the writers' reflective thinking about their own writing may have played an important role in their writing development. This role of reflective consciousness should be further explored in future research (see Conclusions).

### Conclusions

The findings of this study have shown that one-year task-type practice had a general effect on the development of EFL learners' writing. This effect might have been even more noticeable because the students in the study were all basic L2 writers. The strength of the impact seemed to vary with the aspect of writing. Contrary to our prediction, we found that grammatical complexity grew faster than did fluency and lexical complexity. This may imply that the development of fluency and lexical complexity in EFL writing should be investigated over a longer span, or that it may be necessary to offer a more intensive pedagogical intervention to develop these aspects (e.g., corrective feedback).

Since this study is only in its initial stage, it has some limitations. First, with regard to the analysis of text, it may be desirable to look at a wider variety of text measures. For example, we used STRUT as a grammatical complexity index, but MED could be used instead, since McCarthy et al. (2009) have shown that MED evaluated paraphrase quality better than STRUT. Second, there are various other ways to analyze developmental patterns than the one used in this study. In particular, other visualization methods such as a min-max graph may be useful in capturing a critical *phase shift* (Larsen-Freeman & Cameron, 2008) in the EFL student's writing behavior. Third, although this study has significant implications for L2 writing research in that it investigated the changes in writing with data that was collected with high frequency over a long period (30 times over one year), one year may not have been long enough to describe the changes in certain aspects of L2 writing, as the results above suggest.

To our knowledge, this study was the first to classify longitudinal developmental patterns in L2 writing by focusing on its non-linear trajectories. Therefore, it is far from describing the system in which learners learn and develop their L2 writing. In addition to overcoming the limitations mentioned above, future research should delve more into agentivity of writers, such as their motivation to write, what the writers are attempting every time they write, and how their perceptions toward L2 writing change over time.

### Acknowledgements

We are grateful that this research was supported by the Ministry of Education, Sports, Culture, Sports, Science, and Technology in Japan (Research Grant #21520643).

### References

- Bardovi-Harlig, K. (2002). A new starting point?: Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition*, 24, 189-198.
- Bygate, M. and Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and task performance in a second language*. Amsterdam: John Benjamins.
- 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.
- Crossley, S. A., & Salsbury, T. (2011). The development of collocational accuracy and production in English second language speakers. *International Review of Applied Linguistics in Teaching*, 49 (1).
- Crossley, S. A., Salsbury, T., & McNamara, D. (2009). Measuring L2 lexical growth using hypernymic relationships. *Language Learning*, 59(2), 307-334.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2010). The development of polysemy and frequency use in English second language speakers. *Language Learning*, 60(3).
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (in press). The development of semantic relations in second language speakers: A case for Latent Semantic Analysis. *Vigo International Journal of Applied Linguistics*.
- Crossley, S. A., Salsbury, T., McCarthy, P. M., & McNamara, D. S. (2008). Using Latent Semantic Analysis to explore second language lexical development. In D. Wilson & G. Sutcliffe (Eds.), *Proceedings of the 21st International Florida Artificial Intelligence Research Society* (pp. 136-141). Menlo Park, CA: The AAAI Press.
- Elbow, P. (1998). *Writing with power: Techniques for mastering the writing process*. New York: Oxford University Press.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202.
- Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, 10(2), 221-240.
- Henry, K. (1996). Early L2 writing development: A study of autobiographical essays by university-level students of Russian. *The Modern Language Journal*, 80(3), 309-326.
- Hinkel, E. (2002). *Second language writers' text*. Mahwah, NJ: Lawrence 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.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- McCarthy, P. M., Cai, Z., & McNamara, D. S. (2009). Computational replication of human paraphrase assessment. In C. H. Lane & H. W. Guesgen (Eds.), *Proceedings of the 22nd*

- International Florida Artificial Intelligence Research Society Conference* (pp. 266-271). Menlo Park, CA: The AAAI Press.
- McCarthy, P. M., Lehenbauer, B. M., Hall, C., Duran, N. D., Fujiwara, Y., & McNamara, D. S. (2007). A Coh-Metrix analysis of discourse variation in the texts of Japanese, American, and British Scientists. *Foreign Languages for Specific Purposes*, 6, 46-77.
- Ortega, L., & Ibarra-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26-45.
- Polio, C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 91-115). Mahwah, NJ: Lawrence Erlbaum.
- Polio, C., Fleck, C., & Leder, N. (1998). "If I only had more time:" ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7(1), 43-68.
- Shaw, P., & Liu, E. T.-K. (1998). What develops in the development of second-language writing? *Applied Linguistics*, 19(2), 225-254.
- Verspoor, M., Lowie, W., & van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal*, 92(2), 214-231.
- Willett, J. B. (1994). Measurement of change. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed., pp. 671-678). Oxford: Pergamon.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy & Complexity*. Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawaii at Manoa.