The Effect of Writing Task and Task Conditions on Colombian EFL Learners' Language Use

Kim McDonough & César García Fuentes

This classroom study examines whether English L2 writers' language use differs depending on the writing task (operationalized as paragraph type), and task conditions (operationalized as individual or collaborative writing). The texts written by English L2 university students in Colombia (N = 26) in response to problem/solution and cause/effect writing tasks were compared in terms of analytic ratings, use of target grammatical clauses, and accuracy. Approximately half of the students wrote individually while the other half collaborated in pairs. Results indicated that the writing task was a significant factor, with cause/effect paragraphs rated higher and having more target clauses than the problem/solution paragraphs. Task condition was also a significant factor, with collaborative texts more accurate than individual texts. Implications for L2 writing pedagogy are highlighted.

Cette étude s'est déroulée dans une salle de classe et a examiné dans quelle mesure l'utilisation de la langue par des étudiants d'ALS varie selon la tâche d'écriture (types de paragraphe) et les conditions de la tâche (travail individuel ou en groupe). Des étudiants d'ALS dans une université en Colombie (N=26) ont rédigé des textes cause-effet et problème-solution. Nous avons comparé leurs textes en fonction d'éléments analytiques, de l'emploi de propositions grammaticales ciblées et de l'exactitude. Environ la moitié des étudiants ont écrit de façon individuelle; les autres ont rédigé en groupe. Les résultats indiquent que la tâche d'écriture constitue un facteur significatif, les paragraphes cause-effet ayant reçu de meilleures évaluations et comportant plus de propositions ciblées que les paragraphes problème-solution. Les conditions de la tâche se sont aussi avérées être un facteur significatif, les textes collaboratifs étant plus précis que les textes écrits individuellement. Les implications pour l'enseignement de l'écriture en L2 sont exposées.

Reflecting differences in the role that writing plays across L2 instructional settings, researchers have characterized three pedagogical approaches in terms of their objectives and instructional focus: Learning to Write (LW), Writing to Learn Content (WLC), and Writing to Learn Language (WLL). The LW perspective is most closely associated with first-language composition, English for Specific Purposes, and English for Academic Purposes, and focuses on the acquisition of "good" writing and the development of expert writers

(Hyland, 2011; Manchón, 2011). The WLC approach, originating from the use of writing tasks in school-based subjects and associated with Content-based Instruction and Content and Language Integrated Learning, emphasizes the use of writing as a tool to strengthen and extend content knowledge and to promote rethinking, revising, and reformulating of subject-matter knowledge (Hirvela, 2011; Langer, 1986; Langer & Applebee, 1987). Inspired by cognitive and sociocultural approaches to language learning, the WLL perspective emphasizes the role of writing for promoting the acquisition of linguistic knowledge (Cumming, 1990; Manchón, 2009, 2011). WLL focuses on the benefits of writing for "pushing" students to analyze and consolidate their linguistic knowledge (Williams, 2012). It is closely associated with writing instruction for pre-academic or "general" English L2 students, for L2 writers in foreign language contexts where there is little need to use the target language outside the classroom (Ortega, 2009), and in instructional settings where writing serves as a medium for language practice (Sasaki & Hirose, 1996). Situated within the WLL approach, the current study explores whether writing tasks and task conditions differ in their effectiveness at "pushing" English L2 students to deploy their linguistic resources.

Because the goal of writing in WLL is to stimulate the consolidation and extension of students' linguistic knowledge, it is crucial to identify how different writing tasks and task conditions serve this goal (Byrnes & Manchón, 2014). For example, L2 writing research to date has explored whether manipulating the complexity of writing tasks affects L2 writers' language use (e.g., Kormos, 2011; Kuiken & Vedder, 2007; Ong & Zhang, 2010), but fewer studies have compared the paragraph types typically found in "general" EFL and ESL textbooks (McDonough, Crawford, & De Vleeschauwer, in press; Shehadeh, 2011) that complement the WLL emphasis on writing for language learning. The focus on paragraph types is motivated by the prominence given to writing tasks that elicit different styles of paragraphs (descriptive, narrative, argumentative, problem/solution, compare/contrast, etc.) in ESL writing textbooks for pre-academic students (Fitzpatrick, 2011; Hogue, 2013; Ward, 2011), and their widespread use with beginner and low-intermediate writers in both EFL and ESL contexts. In addition to focusing on writing tasks, L2 researchers have also explored whether task conditions, specifically composing texts either individually or collaboratively, affects L2 writers' language use. Although few differences in lexical richness or syntactic complexity have been identified, several studies have reported that collaborative texts are more accurate than individual texts (e.g., Fernandez Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009).

Comparative L2 writing studies of task types and task conditions have relied largely upon global measures of L2 writers' language use, particularly t-unit based measures of complexity (e.g., clauses per t-unit) and accuracy (e.g., errors per t-unit). However, the t-unit (i.e., an independent clause and all its associated dependent clauses) has been criticized for failing to distin-

guish among different grammatical forms (Biber, Gray, & Poonpon, 2011) and obscuring variation among lower-proficiency levels (Norris & Ortega, 2009). Previous research has also revealed a lack of consistency in the classification of dependent and subordinate clauses, specifically the status of nonfinite clauses. Furthermore, the ecological validity of t-unit–based measures can be questioned in that L2 writing instructors are unlikely to assess student texts through t-unit measures; instead they tend to use analytic or holistic rubrics, evaluate student's use of specific linguistic forms, or consider the number of errors students make. For example, in the Colombian university EFL setting investigated here, instructors typically evaluate student texts by using a rubric and identifying whether students produced recently taught grammatical forms.

In order to contribute to ongoing efforts to identify writing task types and task conditions that encourage "general" ESL/EFL students to deploy their linguistic knowledge through writing, the current study compared two writing tasks (cause/effect versus problem/solution) and two writing conditions (individual versus collaborative) in terms of text ratings, the students' use of target structures, and their accuracy. The first research question was "Does writing task type affect the ratings, target structure use, and accuracy of paragraphs written by Colombian EFL university students?" Because few studies have compared the linguistic features of the paragraph types typically found in pre-academic EFL/ESL textbooks, no predictions about possible differences were made. To address the potential impact of task conditions, the second research question was "Do task conditions affect the ratings, target structure use, and accuracy of paragraphs written by Colombian EFL university students?" In light of previous research that found differences only in accuracy, we predicted that collaborative texts would be more accurate than individual texts.

Method

Participants and Instructional Setting

The participants were 26 undergraduate students (16 women, 10 men) enrolled in two required, integrated-skills EFL classes at a large university in Bogotá, Colombia, that were taught by the same instructor. They ranged in age from 17 to 23 years, with a mean age of 19.7 years (SD = 1.7). Their previous amount of English instruction varied widely, ranging from 1 to 15 years, with a mean of 5.8 years (SD = 3.9) of prior instruction in EFL classes in primary and secondary schools. However, despite variation in the number of years of previous formal instruction, the students had similar prior experience using spoken English. The majority of the students (20 of 26) had never travelled to a country where English was spoken, while 6 students reported visits ranging in length from 8 to 90 days. Furthermore, in terms of their pro-

ficiency, the students had never taken any international standardized tests, but based on the university placement exam they were all equivalent to B1 in the Common European Framework of Reference. The validation information provided with the students' course textbook (*Touchstone 4*; McCarthy, McCarten, & Sandiford, 2009), which was designed for B1-level learners, indicated that the textbook was appropriate for students with IELTS scores of 4.0–5.0. In other words, the students were pre-academic English learners who would not be admitted to Canadian university degree programs, but would be placed in full-time intensive English programs.

The students' EFL class was a theme-based, integrated skills course that met for two, 100-minute classes per week. Each unit had a content theme or topic, target vocabulary items and grammatical structures, and provided a variety of aural and written texts, along with oral and written activities that created opportunities to practice the target forms in the context of the unit themes. In terms of writing practice specifically, the textbook provided sentence and paragraph activities in which students were asked to convey information and ideas about concrete and abstract topics related to the unit content (e.g., careers, fame, law and order). Reflecting the WLL orientation, the course did not provide instruction about academic writing conventions (e.g., thesis statements, topic sentences, introductions, or conclusions) or teach the rhetorical organization of different genres (e.g., narration or argumentation). Furthermore, stages in the writing process, such as planning, drafting, and revising, were neither explicitly taught nor emphasized. Instead, the course focused on writing for the purpose of developing competence with the vocabulary and grammar in each unit. Prior to completing the paragraphs analyzed here, the students had written argumentative (week 8), descriptive (week 9), and narrative (week 11) paragraphs.

Materials and Procedure

The students completed two paragraph-level writing tasks in weeks 13 and 14 of a 16-week academic semester. Both paragraphs were written when the students were completing the same chapter in their textbook, which was thematically organized around the topic of careers, highlighted vocabulary related to jobs, and targeted nominal clauses, nouns with postnominal modification, and future verb tenses. These paragraphs were selected because they were completed during the same instructional unit and were intended to provide students with opportunities to use the same grammatical structures. The writing tasks complemented the content focus and order of activities in the textbook chapter while eliciting different paragraph types. The first writing task was to describe a problem the students had faced in their educational or professional careers, and to suggest some solutions to that problem. The instructions stated that the students should explicitly describe the problem and two solutions, and gave a recommended paragraph length of 100–150 words.

The second writing task was to describe the causes of a failed job search. The instructions specified that the students should describe at least two causes for why a person might fail to get a job, with a suggested paragraph length of 100–150 words. Both writing prompts were pilot tested with students at the same university enrolled in a similar EFL class. Based on the pilot paragraphs, the format of the writing prompt was modified slightly to emphasize through bold type that two solutions or causes should be provided and to create space where students could brainstorm ideas prior to writing if desired. As mentioned previously, students had no prior instruction in the writing process, such as the benefits of brainstorming and outlining prior to writing, or the discourse structure of academic writing (e.g., main ideas), so the writing prompt revisions drew their attention to these features.

The writing tasks were administered by their instructor during weeks 13 (problem/solution) and 14 (cause/effect) of the students' regularly scheduled EFL classes. The order of writing tasks was determined by the topics in the course textbook, so that each writing task complemented the specific content and activities scheduled for each day. The second researcher was present in the classroom while the students were writing and was available to help the instructor answer any questions. In order to manipulate task conditions, the instructor asked students in one class to write the paragraphs individually, while students in the second class were asked to write collaboratively in pairs. In both classes, the students had 20 minutes to write their paragraphs. They were encouraged to take a few minutes to brainstorm ideas and plan their texts prior to writing, but planning time was not monitored or enforced. While they were writing, the students had access to a monolingual English dictionary and could ask the instructor or second researcher questions.

Data Coding & Analysis

In order to facilitate data sharing between the researchers, the students' hand-written paragraphs were converted into Microsoft Word documents and verified by research assistants. The typed texts were rated by three ESL instructors with experience rating paragraphs written by EFL students. They used an analytic rubric (see Appendix) with three categories (content, organization, and language) and four score bands (*poor*, *needs improvement*, *satisfactory*, and *good*). The rubric was previously used to assess Thai EFL university students' paragraphs (McDonough et al., in press), and was adapted based on feedback from Colombian instructors to ensure that the categories and scoring bands reflected the criteria used in their context. Following training from the first researcher, the three raters independently assessed the paragraphs, and interrater reliability, assessed using a two-way mixed average-measures intraclass correlation coefficient, was .77. The mean scores assigned by the three raters for each text were calculated, and these means were used for subsequent analyses.

To assess the students' use of the target clauses, the paragraphs were coded by the first researcher for the following clause types: (a) tensed nominal *wh*-clauses functioning as subjects or objects; (b) subject and object *that*-clauses either with or without an overt complementizer; and (c) subject and object relative clauses, which included subject relative clauses without an overt relative pronoun (see examples in Table 1). Nominal clauses and nouns with postnominal modification were the focus of the textbook unit, so the writing tasks' effectiveness at encouraging the students to use these structures was examined rather than global, t-unit–based measures of grammatical complexity.

Table 1
Examples of Target Clauses

Clause Type	*Examples*
Nominal *wh*-clauses	If you don't show them *who you really are*, you can lose the opportunity.
	The students don't understand *what they are talking about.*
Nominal *that*-clauses	It means *that they don't have responsibility*.
	One solution is *that the students talk about it with the teacher.*
Relative clauses	Some people put lies and experience *that they don't have*.
	Second solution could be fine the teachers *that come very late to the class*.

To assess the accuracy of the students' texts, the number of errors per text was calculated. Although errors in spelling and punctuation have not been considered in some collaborative writing research studies (e.g., Storch, 2005; Storch & Wigglesworth, 2007), they were included here because these features of student writing are assessed by instructors in this context. Also coded were errors in syntax (e.g., missing elements, incorrect word order), morphology (e.g., verb forms, plural forms, articles, subject-verb agreement), and word choice errors that impeded meaning. However, errors in paragraph formatting (such as indenting and spacing) were not considered. Frequency counts were obtained for the total number of clauses and errors per text. To account for variation in text length, proportion scores were obtained by dividing the total number of clauses by words and the number of errors by words.

The reliability of the coding was assessed by having an independent rater code a subset of the data (25%) for clause types and accuracy, following training by the first researcher. Interrater reliability was obtained using Pearson correlations (r): Clauses $r = .95$ and Errors $r = .82$. To identify task type differences, the ratings, clause rate, and error rate for the cause/effect and problem/solution paragraphs were compared using paired-samples t-tests. To identify

differences in task conditions, the ratings, clause rate, and error rate for the individual and collaborative paragraphs of each type were compared using independent-samples *t*-tests. The initial alpha level was set at .05.

Results

The data set consisted of 36 paragraphs with an equal number of problem/solution and cause/effect paragraphs. There were 10 paragraphs of each type written by individual students and 8 paragraphs of each type written collaboratively in pairs. The first research question asked whether writing task type affected the ratings, target structure use, and accuracy of paragraphs written by Colombian EFL university students. As shown in Table 2, the cause/effect paragraphs received higher ratings than the problem/solution paragraphs. In addition, the students used a greater proportion of target clauses in the cause/effect paragraphs. Error rates were similar for both paragraph types. Paired-samples *t*-tests using an adjusted alpha level of .017 to correct for multiple comparisons (i.e., .05/3) indicated that there were significant differences between the writing tasks for text ratings and clauses. Based on benchmarks for applied linguistics research (Plonsky & Oswald, 2014), the effect sizes (Cohen's *d*) for paragraph rating and clauses/word fell within the range of values between small (.60) and medium (1.00).

Table 2
Features by Task Type

	Problem & Solution		*Cause & Effect*		*Comparison*		
	M	*SD*	*M*	*SD*	*t*	*p*	*d*
Paragraph rating	18.33	2.40	20.06	2.48	2.97	.009	.71
Target clauses/word	.02	.01	.03	.02	3.16	.006	.63
Errors/word	.14	.05	.16	.07	.88	.389	.33

The second research question asked whether task conditions affected the ratings, target structure use, and accuracy of paragraphs written by Colombian EFL university students. The paragraph ratings, target clause use, and accuracy for the individual and collaborative texts are provided in Table 3. Because the findings for the first research question revealed significant differences for writing task type, the results for task condition are presented separately for each task type.

Although the paragraph ratings were slightly higher for the collaborative texts for both task types, the differences were not significant: problem/solution $t(16) = 1.00$, $p = .333$, $d = .48$; and cause/effect $t(16) = 1.23$, $p = .220$, $d = .62$. There were also no significant differences in the students' use of the target

clauses based on task conditions for either the problem/solution paragraphs, $t(16) = 1.12$, $p = .280$, $d = 0$, or the cause/effect paragraphs, $t(16) = 1.16$, $p = .263$, $d = .63$. However, as predicted based on previous research, the collaborative texts were more accurate (i.e., had lower error rates) than the individual texts, for both task types. The difference in accuracy rates was significant for both problem/solution paragraphs, $t(16) = 3.06$, $p = .007$, $d = 1.70$, and cause/effect paragraphs, $t(16) = 2.81$, $p = .012$, $d = 1.30$. The effect sizes were large (problem/solution) and approaching large (cause/effect).

Table 3
Features by Task Condition and Task Type

	Problem/Solution		Cause/Effect	
	Individual *M (SD)*	Collaborative *M (SD)*	Individual *M (SD)*	Collaborative *M (SD)*
Paragraph rating	17.82 (2.67)	18.96 (2.02)	19.40 (2.83)	20.88 (1.80)
Target clauses/word	.02 (.01)	.02 (.01)	.03 (.02)	.04 (.01)
Errors/word	.17 (.04)	.11 (.03)	.19 (.07)	.12 (.03)

Discussion

The results indicated that whereas the students' cause/effect essays were rated higher than the problem/solution paragraphs and contained more target clauses, accuracy did not differ by task type. The findings provide additional support for claims that writing task type plays an important role when assessing the quality and generalizability of written texts (e.g., Bouwer, Béguin, Sanders, & van den Bergh, 2015), and sheds further light on potential differences in L2 writing associated with paragraph and essay types (e.g., McDonough et al., in press; Shi, 1998). In terms of task conditions, whether students wrote individually or collaboratively did not affect their paragraph ratings or clause use for either task type. However, students wrote more accurately when they collaborated, which confirms the findings of previous collaborative writing studies (Fernandez Dobao, 2012; Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). Taken together, the findings suggest that task type affected these EFL students' use of more complex linguistic forms, specifically nominal and relative clauses, while task conditions affected their accuracy. Interpreted from the WLL's emphasis on the use of writing as a tool for promoting language learning, both task type and task conditions may play important roles in "pushing" EFL learners to produce complex structures and to use language more accurately, respectively.

The findings raise some potential implications for L2 pedagogy. Although paragraph type has been more extensively researched in first-language writing studies in order to identify interactions between text type and language

development (e.g., Beers & Nagy, 2011), L2 studies to date have not widely investigated writing tasks that elicit different paragraph types. Particularly for the WLL paradigm, where writing is a vehicle for language development, more extensive research to identify how specific text types (paragraphs or essays) encourage L2 writers to deploy their linguistic resources is needed. In instructional settings where WLL is the predominant perspective, such as in EFL contexts or intensive ESL programs for pre-academic students, instructors may need to consider how paragraph type interacts with language use. Particularly in settings where written texts are used to evaluate whether students produce target structures accurately, instructors should be aware that task type and conditions may influence their use of those structures. For example, although both the cause/effect and problem/solution writing tasks were intended to provide practice with nominal and relative clauses as part of the same instructional unit, their effectiveness at eliciting those structures varied considerably. More extensive pilot testing of writing task prompts may be necessary in contexts where writing is being used to assess students' language use.

The association between collaborative writing and accuracy suggests that it may be beneficial for instructors and students to be aware of advantages that writing collaboratively might bring for language development. In some EFL contexts, such as the one reported here, writing is considered to be an individual task, in part due to concerns with timing (i.e., collaborative writing often takes longer) and equal participation in the task. However, instructors might include collaborative writing in their lessons as a way to promote accuracy. Besides positively impacting student texts, the process of writing collaboratively might create opportunities for students to negotiate meaning and discuss language forms that they then incorporate into their texts (e.g., Brooks & Swain, 2009; Watanabe & Swain, 2007). Particularly in EFL settings in which opportunities for target language use may be limited, collaborative writing tasks could be used as a tool for eliciting language production where students work cooperatively instead of in isolation, thereby creating a learning environment that encourages integrated language development (Hinkel, 2006).

Despite their widespread use in EFL and intensive ESL settings, the types of paragraph-level writing tasks investigated in the current study have been criticized because students can complete them by drawing exclusively on their personal experiences and knowledge (Carroll & Dunkelblau, 2011; Leki & Carson, 1997). Unlike writing tasks in students' academic disciplines, which draw upon information students gather from some source texts, the typical paragraph-level writing tasks used in EFL and intensive ESL settings generally do not require students to use information from external sources. Asking students to read and react to source texts may promote language development through writing in ways that differ from tasks that draw exclusively on their personal experiences and knowledge. To complement the WLL

orientation, source texts could be selected based on the occurrence of target grammatical structures, thereby providing students with additional models of the forms they are expected to acquire.

Although the findings indicated that cause/effect paragraphs were more effective at providing these EFL students with opportunities to use more complex grammatical forms while collaborative writing tasks facilitated accuracy, there were several limitations to the generalizability of the study. First, the sample size was relatively small, with only 18 texts per paragraph type, and an unbalanced number of individual and collaborative texts. In order to gain more robust findings about the impact of writing tasks and task conditions on EFL students' language use, additional studies with larger samples are needed to confirm the findings. Second, our focus was on the use of writing in EFL settings where the instructional focus is on writing as a way to practice and learn language. Consequently, the findings may not be applicable in instructional settings where writing tasks are used to promote content learning (such as in content-based or content and language integrated learning classrooms) or to develop students' writing abilities (such as English for Academic Purposes classes). Although the findings are applicable to instructional settings with similar goals, such as intensive English programs for pre-academic ESL students, additional studies in a variety of contexts are needed to verify the findings.

Finally, because the data were collected as part of the instructional routine of the EFL class, the order of the writing tasks was not counterbalanced. As is typical in classroom instruction, the writing tasks were assigned to complement textbook content and activities, which progressed from the beginning to the end of the chapter. Counterbalancing the tasks would have required some students to write on topics before covering the background material, which was not appropriate in our instructional context. Because the tasks were created to reflect the content of the chapter, task features such as formality were not strictly controlled. Although it is unlikely that these EFL students would experience significant language development in the one-week interval between the writing tasks, it is possible that the students produced more target structures because they wrote the cause/effect paragraphs after the problem/solution paragraphs. Future studies could adopt counterbalancing in order to determine whether task order played a role in the students' language use. By addressing these limitations and adopting longitudinal designs to explore the interaction between writing tasks and language development, our future research aims to provide instructors in WLL settings with empirical findings that can help inform their pedagogical decisions.

Acknowledgements

Funding for this project was provided by research grants from the Quebec Ministry of Education (Fonds Québécois de la recherche sur la société et la culture) and the Canada Research Chairs program awarded to the first author, and a Concordia University Mobility Award received by

the second author. We would like to thank the research assistants who helped with data preparation and coding: Jason Bretz, Jonathan Brouillette, Maxime Lavallee, Mathew Lazenby, and Phung Van Dao. We would also like to thank the teacher, Francisco Hernández, and students who participated in this study, along with Jorge Tadeo Lozano University and Centro Colombo Americano for allowing us to carry out the study.

The Authors

Kim McDonough is a Professor and Canada Research Chair in Applied Linguistics at Concordia University. Her research interests include interaction and usage-based approaches to second language acquisition, structural priming, and task-based language teaching.

César García Fuentes recently completed an MA in Applied Linguistics at Concordia University. His research interests include the acquisition of pragmatic competence, task repetition, and the use of technology for second language teaching.

References

Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing, 24*(2), 183–202.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35.

Bouwer, R., Béguin, A., Sanders, T., & van den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing, 32*(1), 83–100.

Brooks, L., & Swain, M. (2009). Languaging in collaborative writing: Creation of and response to expertise. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 58–89). New York, NY: Routledge.

Byrnes, H., & Manchón, R. M. (2014). Insights from and for L2 writing: An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 1–23). Amsterdam, Netherlands: John Benjamins.

Carroll, J., & Dunkelblau, H. (2011). Preparing ESL students for "real" college writing: A glimpse of common writing tasks ESL students encounter at one community college. *Teaching English in the Two-Year College, 38*(3), 271–281.

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication, 7*(4), 482–511.

Fernandez Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing, 21*(1), 40–58.

Fitzpatrick, M. (2011). *Engaging writing 1: Essential skills for academic writing*. London, UK: Pearson.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly, 40*(1), 109–131.

Hirvela, A. (2011). Writing to learn in content areas: Research insights. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 37–59). Amsterdam, Netherlands: John Benjamins.

Hogue, A. (2013). *Longman academic writing series: Level 2* (3rd ed.). London, UK: Pearson Longman.

Hyland, K. (2011). Learning to write: Issues in theory, research, and pedagogy. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 17–35). Amsterdam, Netherlands: John Benjamins.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20*(2), 148–161.

Kuiken, F., & Vedder, I. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 117–135). Clevedon, UK: Multilingual Matters.

Langer, J. A. (1986). Learning through writing: Study skills in the content areas. *Journal of Reading, 29*(5), 400–406.
Langer, J. A., & Applebee, A. N. (1987). *How writing shapes thinking: A study of teaching and learning.* Urbana, IL: National Council of Teachers of English.
Leki, I., & Carson, J. (1997). Completely different worlds: EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly, 31*(1), 39–69.
Manchón, R. M. (2009). Broadening the perspective of L2 writing scholarship: The contribution of research on foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 1–19). Bristol, UK: Multilingual Matters.
Manchón, R. M. (2011). Writing to learn the language: Issues in theory and research. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 61–82). Amsterdam, Netherlands: John Benjamins.
McCarthy, M., McCarten, J., & Sandiford, H. (2009). *Touchstone 4*. New York, NY: Cambridge University Press.
McDonough, K., Crawford, W., & De Vleeschauwer, J. (in press). Thai EFL learners' interaction during collaborative writing tasks and its relationship to text quality. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda.* Amsterdam, Netherlands: John Benjamins.
Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578.
Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing, 19*(4), 218–233.
Ortega, L. (2009). Studying writing across EFL contexts: Looking back and moving forward. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching and research* (pp. 232–255). Bristol, UK: Multilingual Matters.
Plonsky, L., & Oswald, F. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912.
Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning, 46*(1), 137–168.
Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing, 20*(4), 286–305.
Shi, L. (1998). Effects of prewriting discussions on adult ESL students' compositions. *Journal of Second Language Writing, 7*(3), 319–345.
Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14*(3), 153–173.
Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). Clevedon, UK: Multilingual Matters.
Ward, C. (2011). *Focus on writing: Level 3*. London, UK: Pearson Longman.
Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research, 11*(2), 121–142.
Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing, 26*(3), 445–466.
Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing, 21*(4), 321–331.

Appendix
Rubric

	Content	*Organization*	*Language*
Good (9–10)	The main idea is stated in a clear and interesting way. Supporting information is relevant, elaborated, and fully developed. The ideas are creative, interesting, unique, or unexpected. Paragraph length is appropriate to address the topic specified in the prompt.	Organization is effective. Ideas are linked with varied transitions that are effective and accurate. The paragraph format has a few elements that are not reflective of academic English.	Effective and varied sentence structure. Errors (if any) are due to misspellings or a lack of proofreading. Good vocabulary range and accuracy, with some academic words. Correct and effective word usage.
Satisfactory (7–8)	The main idea is clearly stated. Supporting information is mostly relevant and developed. The ideas are good, but are not creative or unique. Paragraph length leaves some aspects of the topic specified in the prompt underdeveloped.	Organization is acceptable. Ideas are linked with transitions that have some errors or redundancies. The paragraph format does not follow academic English.	Some variation in sentence structure. Some errors in sentence structure. A few word form errors that do not affect readability. Acceptable vocabulary range and accuracy, although academic words may be infrequent.
Needs improvement (5–6)	The main idea is stated, but it is too broad or too specific or difficult to locate. Some supporting information is irrelevant or not fully developed. The ideas are generic, obvious, or uninteresting. The paragraph is short, which leaves several aspects of the topic specified in the prompt underdeveloped.	Organization is weak. Ideas are not clearly linked. Transitions are repetitive or have frequent errors. The format does not follow academic English or resembles lists of sentences as opposed to a cohesive paragraph.	Formulaic or repetitive sentence patterns. Simple sentences used excessively. Frequent errors of sentence structure and word forms. Ordinary vocabulary range with few attempts to use academic vocabulary. Vocabulary errors that do not interfere with readability.
Poor (1–4)	The main idea is confusing, unclear, or missing. Supporting information is not relevant, and details are minimal or random. The ideas are poorly matched to the topic or illogical. The paragraph is so short that the topic specified in the prompt is barely addressed.	Organization is confusing and disjointed. Transitions are missing, inappropriate, or illogical. The format resembles lists of sentences as opposed to a cohesive paragraph.	Multiple and serious errors of sentence structure, i.e., fragments and run-ons. Errors with simple sentences. Limited vocabulary range consisting of general words. Vocabulary usage is inaccurate and interferes with readability.