



Title of Project:

Product and Process in TOEFL iBT Independent and Integrated Writing Tasks: A Validation Study

Researcher:

Liang Guo
Georgia State University
guoguo18@gmail.com



Liang Guo

Research Supervisor:

Dr. Sara C. Weigle

Project Summary:

This study focuses on comparing text based integrated writing tasks (writing based on source materials) with more traditional independent writing tasks. Despite their many potential benefits, there have been relatively few studies of integrated writing tasks in the literature of second language (L2) writing assessment, especially when compared with the abundance of research on independent writing tasks. The very limited studies that have examined integrated writing tasks have focused either on the product or the process of the writing performance to validate such tasks used in a testing context. To clarify the construct inherent in text based integrated writing and to verify the previous statements that have been made about the task as a promising task type, it is necessary to conduct quantitative textual analysis of the essays composed by test takers and obtain qualitative information on the writing processes as well (Bachman, 2004). The central issue addressed in the current study, therefore, is concerned with whether writing products and processes elicited by TOEFL iBT integrated and independent writing tasks vary with task type, perceived writing quality, and the academic experiences of test takers in accordance with theoretical expectations.

In order to address the issue of the influence of task type, writing quality, and test taker experiences on writing products and processes, quantitative analyses were conducted on the integrated and independent essays. In addition, qualitative methods were used to explore the writing processes generated by the two tasks respectively. The quantitative section focused on whether textual features vary along with the task type, the perceived writing quality, and the academic experiences of test takers. The qualitative section investigated how writing processes are related to the same three factors. In the following paragraphs, I explain how the data were collected and analyzed, provide a summary of the findings, and suggest possible implications.

To answer the question whether the textual features vary along with the task type, the perceived writing quality, and the academic experience of test takers, 480 TOEFL iBT essays (from fall 2007) provided by Educational Testing Service (ETS) were collected. Two hundred forty test takers from different linguistic backgrounds responded to an integrated task and then an



independent writing task. For the integrated writing task, the test takers read a passage and a listening passage which addressed the same topic but offered a different perspective from the reading material. Then, the test takers were required to summarize the listening passage and state how it challenged the points made in the reading passage. For the independent writing task, the test takers were required to write an argumentative essay on a given topic, where they supported an opinion on a given topic using their prior knowledge and/or experiences. Co-Metrix, a computational tool, was used to analyze textual features of the collected essays.

To explore whether writing processes change with the task type, the perceived writing quality, and the academic experience of test takers, 20 English as a second language (ESL) writers were recruited to participate in think aloud sessions to articulate their thoughts while composing. They were given the same tasks as in the quantitative section. Prior to conducting the think aloud writing tasks, training based on Ericsson and Simons (1993) was provided to help the participants to become familiar with the think aloud protocols. Think aloud data were fully transcribed, and writing behaviors were then identified and analyzed by the researcher and re-examined by another experienced researcher.

To investigate whether textual features vary along with the task type, discriminant functional analysis was conducted and the results suggest that linguistic features, mainly lexical features, can predict essay membership with 100% accuracy. The study found that as compared to the independent essays, the integrated writing was characterized by the more frequent use of verbs in 3rd person singular present tense and a larger number of modifiers per noun phrase, which indicated a more detached way of writing and an informational prose style (Biber, 1988). At the lexical levels, the integrated writing was also marked by heavy use of concrete words and meaningful words. This finding might be an artifact of integrated writing as the writing content was highly controlled. The fact that many more concrete words were used also suggested that the integrated writing tends to be more context-independent than the independent writing, another characteristic of formal, academic writing (Crossley & McNamara, 2009).

To study whether textual features vary with the perceived writing quality (as determined by the essay scores) within each task type, a regression analysis was conducted using the scores as the dependent variable and Coh-Metrix indices as the independent variables. The results illustrated that the linguistic features can at least partially predict evaluation of writing quality for either the integrated or the independent essays.

Comparing the predictor indices of the integrated essays with those of the independent essays, it can be seen that there are both similarities and differences. Both types of essay scores were found to be positively correlated with text length and lexical sophistication and negatively correlated with verbs in base forms. The results indicated that the longer essays and the essays containing sophisticated words tended to be rated higher in both tasks regardless of whether these items were specified in the scoring rubric. Verbs in base forms, through further investigation, were found to be mainly grammatical mistakes (without the necessary morphemes to indicate tense, number or person). The negative correlation indicated that the essays demonstrating higher grammatical accuracy were more likely to be scored higher. The higher



rated integrated essays had a significantly higher score in semantic similarity, which represents conceptual similarities among sentences. No cohesive devices that could predict the essay scores in the independent writing tasks were identified. Verbs in the past participle form were a significant predictor of the integrated essay scores but not for the independent scores.

The last quantitative question focuses on whether the linguistic features vary with the academic experience of the test takers within each task type. Writers with more experience at the tertiary level should gain more familiarity with academic writing because of the exposure and practice they have in understanding and communicating in such activities (Kutoba, 1998). The study, however, found that the majority of the linguistic features did not significantly discriminate the two groups of writers. Academic experience at the tertiary level does not seem to leave a noticeable trace in the linguistic choices made by the writers while constructing the integrated and independent essays. Furthermore, the very few features that showed a significant difference across the two groups of test takers do not lead to a score difference for either of the tasks.

The qualitative section first looked at the relationship between the writing behaviors and the task type. The results of a *t*-test illustrated that the two tasks—the integrated and the independent—provoked similar behaviors in the L2 writers. Wilcoxon Signed-rank test results showed that the participants did not differ significantly across the two tasks in “verbalizing one’s writing,” “planning and rehearsal,” “commenting on one’s writing product,” “commenting on one’s writing process,” and “analyzing the task.” However, the integrated writing task generated significantly lower percentage than the independent writing task in “reading the instruction,” “global planning,” “reading one’s writing,” and “revising and editing.” Due to the different nature of writing across the two tasks, the two tasks also produced writing behaviors that were unique to each of them. The integrated writing task had “summarizing source texts,” “referring to source texts,” “commenting on understanding of source texts,” and “commenting on relationship between source texts.” Meanwhile, the independent writing task had a particular category of “positioning self,” which indicated that the test takers spent time evaluating which side of the argument they chose.

As for whether the writing behaviors varied with the writing scores, two experienced raters using the rubrics provided by ETS scored the essays. The participants were divided into high and low performance groups for the integrated and for the independent tasks. For the integrated task, Mann Whitney test revealed no significant differences in writing behaviors across the two performance groups. A similar finding was reported for the independent task except for the category of “revising and editing.” Namely, the high performance participants revised and edited significantly more frequently than their counter parts.

The very last qualitative question looked at the writing behaviors in relation to the academic experience of the participants within each task. The study found that the integrated writing task did not generate significant differences in terms of the types and the frequency of the writing behaviors. As for the independent writing task, the two groups of participants did not differ significantly from each other in the number and type of writing behaviors. However, in terms of



frequency, Mann Whitney test revealed that the undergraduate participants produced significantly more “analyzing the task” behavior than the graduate.

This study aimed to clarify the issue whether writing products and writing processes elicited by TOEFL iBT integrated and independent writing tasks vary with task type, perceived writing quality, and academic experience of test takers in accordance with theoretical expectations. The study found that, on one hand, the textual features varied with the task type and the perceived writing quality. Certain textual features were reported to be able to predict essay membership with 100% accuracy. For each task type, textual features were also found to be able to predict the essay scores (the perceived writing quality). However, as for how textual features are related to the academic writing experience of the test takers, no significant differences were identified within each of the tasks. On the other hand, the majority of the writing behaviors were not found to vary with the task type, the perceived writing quality, or the academic experiences of the test takers.

The findings have significant implications for both L2 writing assessment and L2 writing instruction. The rich empirical evidence revealed that the integrated writing and the independent writing tasks did elicit different writing performances and, thus, affirmed the proposed rationale for the combined use of the two tasks, namely broadening “representation of the domain of academic writing on the test” (Huff et al., 2008). Therefore, the question whether and why the two test items should be used simultaneously in assessing academic writing is answered. The study also investigated whether test performance varied with the writing quality perceived and the academic experience of the test takers. The results not only helped to clarify the score meaning in each of the writing tasks but also to validate the scoring rubrics used. Taken together, all this information helps to clarify the link between the observed score and the underlying writing ability being assessed (Chapelle, Enright, & Jamieson, 2008), thus building a more comprehensive picture of L2 writing assessment, especially in regards to integrated writing tasks.

As for L2 writing instruction, the differences identified between the products and processes across the two tasks suggest that these two types of writing represent at least two different aspects of academic writing ability. Instruction in the more conventional independent argumentative writing by itself might not suffice and fully prepare L2 writers to cope with academic writing tasks. Writing instruction and learning, therefore, should include sourcing texts and the synthesizing of these texts into writing to provide L2 writers with adequate exposure to such writing activities and to develop the corresponding writing ability that is integral to academic activities of higher education (Cohen, 1998; Hirvela, 2004; Bachman, 2004).



References

- Ascencion, Y. (2005). *Validation of reading-to-write assessment tasks performed by second language learners*. Unpublished doctoral dissertation. Northern Arizona University.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
- Bachman, L. (2004). *Statistical analysis for language assessment*. Oxford, UK: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.
- Bereiter, C., & Scardamalia, M. (1985). Cognitive coping strategies and the problem of "inert knowledge." In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (pp. 65-80). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge, UK: Cambridge University Press.
- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge, UK: Cambridge University Press.
- Biber, D. & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9(1), 2-20.
- Brace, N., Kemp, R., & Snelgar, R. (2006). *SPSS for psychologists: A guide to data analysis using SPSS for windows* (3rd edition). London, UK: Palgrave.
- Braine, G. (1995). Writing in the natural sciences and engineering. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 113-134). Norwood, NJ: Ablex Publishing Corporation.
- Brown, A. L. & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22(1), 1-14.
- Brown, J. D., Hilgers, T., & Marsella, J. (1991). Essay prompts and topics: Minimizing the effect of mean differences. *Written Communications*, 8(4), 533-556.



- Camp, R. (1993). Changing the model for the direct writing assessment. In M. M. Williamson & B. A. Huot (Eds.), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations* (pp. 45-78). Cresskill, NJ: Hampton Press, Inc.
- Campbell, C. (1990). Writing with other's words: Using background reading text in academic compositions. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 211-230). Cambridge, UK: Cambridge University Press.
- Carson, J. (2001). A task analysis of reading and writing in academic contexts. In D. Belcher & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 48-83). Ann Arbor, MI: The University of Michigan Press.
- Carlson, S., Bridgeman, B., Camp, R., & Waanders, J. (1985). *Relationship of Admission Test Scores to Writing Performance of Native and Non-native Speakers of English*. (TOEFL Research Rep No. 19). Princeton, NJ: Educational Testing Service.
- Casanave, C. & Hubbard, P. (1992). The writing assignments and writing problems of doctoral students: Faculty perceptions, pedagogical issues, and needed research. *English for Specific Purposes Journal*, 11(1), 33-49.
- Chafe, W. L. (1975). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (Ed.), *Subject and topic* (pp. 26-55). New York, NY: Academic Press.
- Chapelle, C. A., Enright, M. K., & Jamieson, J. M. (2008). Test score interpretation and use. In C. A. Chapelle, M. K. Enright, & J. M. Jamieson (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 145-186). New York, NY: Routledge.
- Charge, N. & Taylor, L. B. (1997). Recent development in IELTS. *English Language Teaching Journal*, 51(4), 374-380.
- Chodorow, M. & Brustein, J. (2004). *Beyond essay length: Evaluating e-rater's performance on TOEFL essays*. (TOEFL Research Report No.73). Princeton, NJ: ETS.
- Church, E. & Bereiter, C. (1984). Reading for style. In J. M. Jensen (Ed.), *Composing and comprehending* (pp. 85-91). Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills.
- Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 82-95). Clevedon, UK: Multilingual Matters.
- Cohen, A. D. (1998). Strategies and processes in test taking and SLA. In M. H. Long & J. C. Richards (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 90-111). Cambridge, UK: Cambridge University Press.



- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33(4), 497-505.
- Connor, U. (1990). Linguistic/rhetorical measures of international persuasive student writing. *Research in the Teaching of English*, 24(1), 67-87.
- Connor, U. & Biber, D. (1988). *Comparing textual features in high school student writing: A crosscultural study*. Unpublished manuscript.
- Connor, U. & Carrell, P. (1993). The interpretation of tasks by writers and readers in holistically rated direct assessment of writing. In J. Carson & I. Leki (Eds.), *Reading in composition classroom* (pp. 141-160). Boston, MA: Heinle and Heinle.
- Crossley, S. A. & McNamara, D. S. (2009). Computationally assessing lexical differences in second language writing. *Journal of Second Language Writing*, 17(2), 119-135.
- Crossley, S. A. & McNamara, D. S. (in press a). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*.
- Crossley, S. A. & McNamara, D. S. (in press b). Detecting the first language of second language writers using automated indices of cohesion, lexical sophistication, syntactic complexity, and conceptual knowledge. In S. Jarvis & S. A. Crossley (Eds.), *Approaching language transfer through text classification explorations in the detect-based approach*. Bristol, UK: Multilingual Matters.
- 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: AAAI Press.
- Crossley, S. A., Louwse, M. M., McCarthy, P.M., & McNamara, D.S. (2007). A linguistic analysis of simplified and authentic texts. *Modern Language Journal*, 91(2), 15-30.
- Cumming, A. (1997). Learning to write in a second language: Two decades of research. *International Journal of English Studies*, 1(2), 1-23.
- Cumming, A., Kantor, R., Powers, D., Santos, T., & Taylor, C. (2000). *TOEFL 2000 writing framework: A working paper* (TOEFL Monograph Series, Report No. 18). Princeton, NJ: ETS.
- Cumming, A., Kantor, R. Baba, K., Erdoosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in writing-only and reading-to-write prototype tasks for next generation TOEFL. *Assessing Writing*, 10(1), 5-43.



- Cumming, A., Kantor, R. Baba, K., Erdoosy, U., Eouanzoui, K., & James, M. (2006). *Analysis of discourse features and verification of scoring levels for independent and integrated tasks for the new TOEFL* (TOEFL Monograph No. MS-30). Princeton, NJ: ETS.
- Currie, P. (1998). Staying out of trouble: Apparent plagiarism and academic survival. *Journal of Second Language Writing*, 7(1), 5-43.
- Delaney, Y. A. (2008). Investigating the reading-to-write construct. *Journal of English for Academic Purposes*, 7(3), 140-150.
- Dulay, H., Burt, M. & Krashen, S. D. (1982). *Language two*. New York, NY: Oxford University Press.
- Durst, R. K. (1987). Cognitive and linguistic demands of analytical writing. *Research in the Teaching of English*, 21(4), 347-376.
- Educational Testing Service. (2008). Validity Evidence Supporting the Interpretation and Use of TOEFL iBT™ Scores. *TOEFL iBT Research Insight*, 1(4). Retrieved May 2, 2011, from http://www.ets.org/s/toefl/pdf/toefl_ibt_insight_s1v4.pdf.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford, UK: Oxford University Press.
- Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4(2), 139-155.
- Enright, M., Bridgeman, B., & Cline, F. (2002, April). Prototyping a test design for a new TOEFL. In D. Eignor (chair), *Research in support of the development of new TOEFL*. Symposium conducted at the meeting of the National Council on Measurement in Education. New Orleans.
- Enright, M., Bridgeman, B., Eignor, D., Lee, Y. W., & Powers, D. E. (2008). Prototyping measures in listening, reading, speaking and writing. In C. A. Chapelle, M. K. Enright, & J. M. Jamieson (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 145-186). New York, NY: Routledge.
- Ericsson, K.A. & Simon, H. (1987). Verbal reports on thinking. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 24-53). Clevedon, UK: Multilingual Matters.
- Ericsson, K.A. & Simon, H. (1993). *Protocol Analysis: Verbal reports as data*. Cambridge, MA: MIT Press.



- Esmaeili, H. (2002). Integrated reading and writing tasks and ESL students' reading and writing performance in an English language test. *The Canadian Modern Language Review*, 58(4), 599-622.
- Faerch, C. & Kasper, G. (1987). From product to process: Introspective methods in second language research. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 5-23). Clevedon, UK: Multilingual Matters.
- Feak, C. & Dobson, B. (1996). Building on the impromptu: A source-based academic writing assessment. *College ESL*, 6(1), 73-84.
- Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Ferris, D. (1993). The design of an automatic analysis program for L2 text research: Necessity and feasibility. *Journal of Second Language Writing*, 2(2), 119-129.
- Ferris, D. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly*, 28(2), 414-420.
- Field, A. (2005). *Discovering statistics using SPSS*. London, UK: Sage Publications.
- Flower, L. & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Frase, L., Faletti, J., Ginther, A., & Grant, L. (1999). *Computer Analysis of the TOEFL Test of Written English*. (TOEFL Research Report No. 64). Princeton, NJ: ETS.
- Friend, R. (2001). Effects of strategy instruction on summary writing of college student. *Contemporary Educational Psychology*, 26(1), 3-24.
- Gebril, A. (2006). *Writing-only and reading-to-write academic writing tasks: A study in generalizability and test method*. Unpublished doctoral dissertation. The University of Iowa.
- Gebril, A. & Plakans, L. (2009). Investigating source use, discourse features, and process in integrated writing tests. *Spaan Fellow Working Papers in Second or Foreign Language Assessment*, 7, 47-84. Retrieved February, 2011, from http://www.cambridgemichigan.org/sites/default/files/resources/SpaanPapers/Spaan_V7_GebrilPlakans.pdf
- Gernsbacher, M. A., & Faust, M. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(2), 245-262.



- Grabe, W. (2001). Reading-writing relations: Theoretical perspectives and instructional practices. In D. Belcher & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 15-47). Ann Arbor, MI: University of Michigan Press.
- Grabe, W. & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. London, UK: Longman.
- Grant, L. & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9(2), 123-145.
- Graesser, A. C., Jeon, M., Yang, Y., & Cai, Z. (2007). Discourse cohesion in text and tutorial dialogue. *Information Design Journal*, 15(3), 199-213.
- Graesser, A. C., McNamara, D. S., & Louwrese, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82-98). New York, NY: Guilford.
- Graesser, A. C., McNamara, D. S., Louwrese, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers*, 36(2), 193-202.
- Green, A. (1998). *Verbal protocol analysis in language testing research: A handbook*. Cambridge, UK: Cambridge University Press.
- Greene, S. (1992). Mining texts in reading to write. *Journal of Advanced Composition*, 12(1), 151-170.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis*. New York, NY: Macmillan.
- Hale, G. A., Taylor, C., Bridgeman, B., Carson, J., Kroll, B., & Kantor, R. (1996). *A study of writing tasks assigned in academic degree programs* (TOEFL Research Report No. 54). Princeton, NJ: Educational Testing Service.
- Halliday, M.A.K. (1994). Foreword. *Functions of Language*, 1(1), 1-5.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London, UK: Longman.
- Hamp-Lyons, L. (1991). Introduction. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic context* (pp. 1-4). Norwood, NJ: Ablex.
- Hamp-Lyons, L. & Kroll, B. (1996). Issues in ESL writing assessment: An overview. *College ESL*, 6(1), 52-72.



- Hayes, J. R. & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. Steinberg (Eds.), *The science of writing* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Mahwah, NJ: Lawrence Erlbaum.
- Hirvela, A. (2004). *Connecting Reading and Writing in Second Language Writing Instruction*. Ann Arbor, MI: The University of Michigan Press.
- Homburg, T. J. (1984). Holistic evaluation of ESL compositions: Can it be validated objectively? *TESOL Quarterly*, 18(1), 87-107.
- Horowitz, D. (1991). ESL writing assessments: Contradictions and resolutions. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic context* (pp. 71-86). Norwood, NJ: Ablex.
- Huff, K., Powers, D. E., Kantor, R. N., Mollaun, P., Nissan, S., & Schedl, M. (2008). Prototyping a new test. In C. A. Chapelle, M. K. Enright, & J. M. Jamieson (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 145-186). New York, NY: Routledge.
- Jamieson, J. M., Eignor, D., Grabe, W., & Kunnan, A. J. (2008). Framework for a new TOEFL. In C. A. Chapelle, M. K. Enright, & J. M. Jamieson (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 145-186). New York, NY: Routledge.
- Jarvis, S. (in press). Data mining with learner corpora: Choosing classifiers for L1 detection. In F. Meunier, S. De Cock, G. Gilquin, M. Paquot (Eds.), *A Taste for Corpora: In honour of Sylviane Granger*. Amsterdam & Philadelphia: John Benjamins.
- Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19(1), 57-84.
- Jennings, M., Fox, J., Graves, B., & Shohamy, E. (1999). The test taker's choice: An investigation of the effect of topic on language test performance. *Language Testing*, 16(4), 426-456.
- Johns, A. & Mayes, P. (1990). An analysis of summary protocols of University ESL students. *Applied Linguistics*, 11(3), 253-271.
- Jurafsky, D. & Martin, J. H. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice Hall.



- Kennedy, M. L. (1985). The composing process of college students writing from sources. *Written Communication*, 2(4), 434-456.
- Kintsch, W. & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Kormos, J. (1998). The use of verbal reports in L2 research: Verbal reports in L2 speech production research. *TESOL Quarterly*, 32(2), 353-358.
- Kroll, B. (1979). A survey of writing needs of foreign and American college freshmen. *English Language Teaching Journal*, 33(3), 219-226.
- Lecocke, M. & Hess, K. (2006). An empirical study of univariate and genetic algorithm-based feature selection in binary classification with microarray data. *Cancer Informatics*, 2(3), 313-327.
- Lee, D. & Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing*, 24(3), 307-330.
- Leki, I. & Carson, J. (1997). Completely different worlds: EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31(1), 36-69.
- Lewkowicz, J. (1994). *Writing from sources: Does source material help or hinder students' performance?* Paper presented at the Annual International Language in Education Conference, Hong Kong. [ERIC Document Reproduction Service No. ED386050].
- Lewkowicz, J. (1997). *Investigating authenticity in language testing*. Unpublished doctoral dissertation. University of Lancaster.
- Lightman, E.J., McCarthy, P.M., Dufty, D.F., & McNamara, D.S. (2007). The structural organization of high school educational texts. In D. Wilson & G. Sutcliffe (Eds.), *Proceedings of the twentieth International Florida Artificial Intelligence Research Society Conference* (pp. 235-240). Menlo Park, California: The AAAI Press.
- Lumley, T. (2005). *Assessing second language writing: The raters' perspective*. Frankfurt, Germany: Peter Lang.
- Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 58-71). Clevedon, UK: Multilingual Matters.
- McCarthy, P. M., & Jarvis, S. (2010). MTL D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392.



- McCarthy, P. M., Lewis, G. A., Dufty, D. F., & McNamara, D. S. (2006). Analyzing writing styles with Coh-Metrix. In G. C. J. Sutcliffe & R. G. Goebel (Eds.), *Proceedings of the 19th Annual Florida Artificial Intelligence Research Society International Conference (FLAIRS)*, pp. 764-770). Melbourne Beach, FL: AAAI Press.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communications*, 27(1), 57-86.
- McNamara, D. S., & Graesser, A. C. (in press). In P. M. McCarthy & C. Boonthum (Eds.), *Applied natural language processing: Identification, investigation, and resolution*. Hershey, PA: IGI Global
- McNamara, D.S., & McDaniel, M. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30(2), 465-482.
- Menard, S. (1995). *Applied logistic regression analysis: Sage University series on quantitative applications in the social sciences*. Thousand Oaks, CA: Sage.
- Messer, S. D. (1997). *Evaluating ESL written summaries: An investigation of the ESL integrated summary profile (ISP) as a measure of the summary writing ability of ESL students*. Unpublished doctoral dissertation. Florida State University.
- Moss, P. A. (1994). Validity in high stakes writing assessment: Problems and possibilities. *Assessing Writing*, 1(1), 109-128.
- Murray, D.H. (1982). *Learning by teaching*. Montclair, NJ: Boynton/Cook.
- Nation, P. (1988). *Word lists*. Victoria, NZ: University of Wellington Press.
- Neter, J., Wasserman, W., & Kutner, M. H. (1989). *Applied linear regression models*. Homewood, IL: Irwin.
- O'Brien, R. T. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690.
- Pearson, P. D. (1974-1974). The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relationships. *Reading Research Quarterly*, 10(2), 155-192.
- Plakans, L. (2007). *Second language writing and reading-to-write assessment tasks: A process study*. Unpublished doctoral dissertation. The University of Iowa.



- Plakans, L. (2008). Comparing composing process in writing-only and reading-to-write test tasks. *Assessing Writing*, 13(2), 111-129.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London, UK: Longman.
- Rayner, K., & Pollatsek, A. (1994). *The psychology of reading*. Mahwah, NJ: Lawrence Erlbaum Associate.
- Reid, J. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), *Second Language Writing: research insights for the classroom* (pp. 191-210). Cambridge: Cambridge University Press.
- Reid, J. (1986). Using the Writer's Workbench in composition teaching and testing. In C. Stansfield (Ed.), *Technology and language testing* (pp. 167-188). Alexandria, VA: TESOL.
- Reid, J. (1992). A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing*, 1(2), 79-107.
- Reppen, R. (1994). *Variation in elementary student language: A multi-dimensional perspective*. Unpublished doctoral dissertation. Northern Arizona University.
- Ruiz-Funes, M. (2001). Task representation in foreign language reading-to-write. *Foreign Language Annals*, 34(3), 226-234.
- Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9(3), 259-291.
- Song, M. Y. (2007). *A correlational study of the holistic measure with the index measure of accuracy and complexity in international English-as-a-second-language (ESL) student writings*. Unpublished doctoral dissertation. University of Mississippi.
- Spivey, N. (1984). *Discourse synthesis: Constructing texts in reading and writing*. Newark, DE: International Reading Association.
- Stratman, J. & Hamp-Lyons, L. (1994). Reactivity in concurrent think-aloud editing protocols: Issues for research. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 89-112). Los Angeles, CA: Sage Publications.
- Swarts, H., Flower, L.S., & Hayes, J. R. (1984). Designing protocol studies of the writing process: An introduction. In R. Beach & L.S. Bridwell (Eds.), *New directions in composition research* (pp. 53-71). New York, NY: Guilford Press.



- Taylor, B. & Beach, R. W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly*, 19(2), 134-146.
- Toglia, M. P. & Battig, W. R. (1978). *Handbook of semantic word norms*. New York, NY: Erlbaum.
- Trites, L. & McGroarty, M. (2005). Reading to learn and reading to integrate: New tasks for reading comprehension test? *Language Testing*, 22(2), 174-210.
- van de Kopple, D. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, 36(1), 82-95.
- van Dijk, T. A. & Kintsch, W. (1977). Cognitive psychology and discourse. In W. U. Dressler (Ed.), *Current trends in text linguistics* (pp. 61-80). Berlin/New York: de Gruyter.
- Wallace, C. (1997). IELTS: Global implications of curriculum and materials design. *ELT Journal*, 51(4), 370-373.
- Watanabe, Y. (2001). *Read-to-write tasks for the assessment of second language academic writing skills: Investigating text features and rater reactions*. Unpublished doctoral dissertation. University of Hawaii.
- Weir, C. (1983). *Identifying the language needs of overseas students in tertiary education in the United Kingdom*. Unpublished PhD Thesis, University of London Institute of Education.
- Weigle, S. C. (2002). *Assessing Writing*. New York, NY: Cambridge University Press.
- Weigle, S. C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing Writing*, 9(1), 27-55.
- White, E. (1994). *Teaching and assessing writing*, 2nd ed. San Francisco, CA: Jossey-Bass Publishers.
- Worden, D. L. (2009). Finding process in product: Prewriting and revision in timed essay responses. *Assessing Writing*, 14(3), 157-177.
- Yang, H. C. (2009). *Exploring the complexity of second language writers' strategy use and performance on an integrated writing test through structural equation modeling and qualitative approaches*. Unpublished doctoral dissertation. The University of Texas at Austin.
- Yang, L. & Shi, L. (2003). Exploring six MBA students' summary writing by introspection. *Journal of English for Academic Purposes*, 2(3), 165-192.



Zipf, G. K. (1945). The meaning-frequency relationship of words. *Journal of General Psychology*, 33, 251–256.