

A Corpus Approach to the Disciplinary History of Second Language Writing

Zhaozhe Wang and Ge Lan (Purdue University, US)

The field of second language (L2) writing has grown into a legitimate field of disciplinary inquiry within a relatively short period of time. Yet despite disciplinary historians' endeavor to provide a narrative of the evolution and shape the identity of the field of second language writing (Matsuda, 1998, 1999, 2006, Silva, 1990, 1993, 2004, etc.), historical inquiry is still largely overshadowed by empirical inquiry, as suggested by the latter's dominant presence in the flagship journals in the field (Canagarajah, 2016). Nevertheless, historical inquiry plays a pivotal role in shaping and contextualizing our understanding of the field. The integration of computer technology into corpus linguistics makes feasible the exploitation of massive corpora (McEnery, Xiao, and Tono, 2006). New research methods such as corpus analysis with the aid of digital technology enable researchers to attain a higher level of historical detachment, as large collections of texts can be machine analyzed and interpreted for consistency and reliability (Biber, Conrad, and Reppen, 1998). Therefore, to foreground disciplinary historical study in the digital era, drawing on meta-analysis and corpus linguistics methodologies, we introduce and demonstrate a new approach—a corpus-driven historical study, to examine and complement the existing historical narratives of the field of second language writing. The research questions we pose are (1) How, if possible, can the use of corpus analysis illuminate our understanding of the historical construction of the field of Second Language Writing? (2) Does the change in frequency of certain high-frequency collocates reflect the shift of scholarly attention in the field across different periods of time? If so, how?

Theme	Collocates	Freq_1980s	Freq_1990s	Freq_2000s
Process	Peer review	6.32	189.51	143.04
	Writing task	85.35	98.41	262.34
	Writing process(es)	222.45	376.57	359.12
	Composing process(es)	417.27	159.41	93.84
Text	Academic writing	360.36	412.36	222.78
	Contrastive rhetoric	69.55	346.48	251.99
	Error correction	41.10	83.77	273.90
	Academic discourse	72.71	176.49	88.87
	Discourse analysis	18.97	39.85	118.08
Literacy	Writing proficiency	208.64	102.48	78.52
	Writing ability	56.90	96.79	72.43
	Academic literacy	0.00	62.63	75.48
	Critical thinking	15.81	117.12	28.00
Language	Linguistic features	0.00	12.20	121.74
	Native language	154.90	95.16	110.78
	Language proficiency	98.00	143.96	81.56
Context	Discourse community	12.64	154.53	183.22
	Cross-cultural	34.77	145.59	83.89
Writer	Native speakers	344.57	243.19	153.39
	L2 writers	15.81	238.31	331.73
	ESL writers	123.29	311.51	54.78
	International students	50.58	109.80	49.30
Feedback	Teacher feedback	0.00	181.37	175.91
	Peer feedback	6.32	150.47	123.56
	Corrective feedback	15.81	12.20	157.04
	Peer response	69.55	303.37	156.43
	Error feedback	22.13	3.25	169.82

Table 1. Normalized frequencies of the theme-based collocates

Specifically, we quantitatively and qualitatively investigate a collection of high-frequency keywords and collocates representing frequently referred terminologies in a corpus consisting of 415 journal articles (total token: 3,186,228) that address second language writing issues in 4 leading journals in the interdisciplinary field, including *Journal of Second Language Writing*, *TESOL Quarterly*, *College Composition and Communication*, and *Written Communication*. The frequencies of the keywords and collocates in the corpus are compared across 3 decades (1980s, 1990s, and 2000s) to track the shift of scholarly attention over time. The “shift of scholarly attention” is operationalized as the change of frequency of the high-frequency collocates based on the high-frequency high-keyness keywords in the scholarship during a certain decade. We present, analyze, and contextualize the frequency change of these collocates categorized into 7 thematic groups—“process,” “text,” “literacy,” “language,” “context,” “writer,” and “feedback.” The findings suggest that the terminological trends in the scholarship are generally aligned with the historical narrative, and a plethora of emerging terminologies are replacing the ones that used to display a dominant presence in the literature.

References

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use (Cambridge approaches to linguistics)*. Cambridge & New York: Cambridge University Press.
- Canagarajah, S. (2016). TESOL as a professional community: A half-century of pedagogy, research, and theory. *TESOL Quarterly*, 50(1), 7-42.
- Matsuda, P. K. (1998). Situating ESL writing in a cross-disciplinary context. *Written Communication*, 15(1), 99-121.
- Matsuda, P. K. (1999). Composition studies and ESL writing: A disciplinary division of labor. *College Composition and Communication*, 50(4), 699-721.
- Matsuda, P. K. (2006). Second-language writing in the twentieth century: A situated historical perspective. In P. K. Matsuda et al. (Ed.), *Second language writing in the composition classroom: A critical sourcebook* (pp. 14-30). New York, NY: Bedford/St. Martins.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies (Unit A1)*. Routledge.
- Silva, T. (1990). Second language composition instruction: Developments, issues, and directions. In B. Kroll (Ed.), *Second Language Writing: Research insights for the classroom* (pp. 11-23). New York: Cambridge University Press.
- Silva, T., Leki, I., & Carson, J. (1997). Broadening the perspective of mainstream composition studies: Some thoughts from the disciplinary margins. *Written Communication*, 14, 398-428.
- Silva, T., & Leki, I. (2004). Family matters: The influence of applied linguistics and composition studies on second language writing studies—past, present, and future. *Modern Language Journal*, 88(1), 1-13.