Variability in citation practices of developing L2 writers in first-year writing courses

Jie Gao (Purdue University, USA), Zhaozhe Wang (Purdue University, USA), Ge Lan (Purdue University, USA), Lindsey Macdonald (Purdue University, USA) and Shelley Staples (University of Arizona, USA)

Citation practices and styles are integral to academic writing contexts. Previous research on citation use in academic writing shows important variability across citation form, function, and the stance of writers in relation to sources (Charles, 2006; Hyland, 2004; Swales, 2014). However, most studies have focused on advanced student and professional writing. Novice second language writers' citation practices have not received the same level of attention in corpus-based studies (Keck, 2014; Pecorari, 2006), and they may require the use of different coding schemes. Using a corpus of L2 writing (majority L1 Chinese), we examine whether (1) the L2 writers' citations vary in number across different assignments; (2) the writers show a preference for particular citation styles.

In this study, we consider two assignments from a first year writing course for L2 writers: (1) a literature review (LR) in which students closely examine sources and in (2) a research paper (RP) in which they use the same sources. Students registered for the writing course have TOEFL writing scores ranging from 19 to 27. Our 132 papers (147,000 words) for this study come from a larger L2 corpus comprising 7 million words from multiple sections of the same writing course. We calculated the number of citations and references in each assignment (per 1,000 words), and coded citations for integral/non-integral forms (Swales, 1990; Hyland, 2004).

Our preliminary results show that student writers used the same number of references in both assignments (32/1,000 words, about 3/paper) but more in-text citations in the literature review (LR=67; RP=45). Writers also used more integral than non-integral citations (73% integral: 79% LR; 64% RP). However, citation counts were highly variable, with 44% of writers using more than 10 citations and 33% using less than 3 per paper. The use of integral vs. non-integral citations also varied widely, with some writers showing a greater preference for non-integral citations. Compared to Swales (2014), who investigated upper level undergraduate and graduate student papers, we observed lower citation counts and references and opposite trends for integral/non-integral citation ratios (Swales found 73% non-integral vs. 27% integral in the MICUSP corpus).

These findings show that, on the whole, these L2 writers are still developing their citation practices, and such source work is extremely variable across assignments and individuals. The results have also exposed the need to develop a more detailed scheme to categorize the citation practices, as the classification of integral and non-integral citations does not seem to cover all the complexity and diversity of developing L2 writers' choices. Our future work will draw on Charles (2006), who has labeled a third category as "general reference" in addition to integral and non-integral citation. O'Donnell (2008)

presented an annotation scheme which also provided new ideas for elaborating "non-integral" and "integral" criteria.

Our findings are expected to provide a more comprehensive analytical framework for understanding citation practices of developing L2 writers in the future. In addition, such work will undoubtedly open up greater opportunities for automatic citation identification and application to the teaching of developing L2 writers.

References

Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. *English for Specific Purposes*, *25* (3), 310-331. doi:10.1016/j.esp.2005.053

Hyland, K. (2004). *Disciplinary Discourses: Social Interactions in Academic Writing*. Ann Arbor: University of Michigan Press. 20–40.

Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, *25*, 4–22. doi:10.1016/j.jslw.2014.055

O'Donnell, M. (2008). *Demonstration of the UAM Corpus Tool for text and image annotation.* Paper presented at the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Demo Session. Columbus, OH, USA: Association for Computational Linguistics.

Pecorari, D. (2006). Visible and occluded citation features in postgraduate second-language writing. *English for Specific Purposes*, *25* (1), 4–29. doi:10.1016/j.esp.2005.044

Swales, J. M. (1990). *Genre analysis: English in academic and research settings.* Cambridge, UK: Cambridge University Press.

Swales, J. M. (2014). Variation in citational practice in a corpus of student biology papers from parenthetical plonking to intertextual storytelling. *Written Communication, 31*(1), 118–141. doi:10.1177/0741088313515166