

The Effects of Rhetorical Structure on Citing Behavior in Research Articles

Wen Lou¹, Zilong Su¹, Shaodan Zheng¹ and Jiangen He²

¹ *wlou@infor.ecnu.edu.cn, 51204407348@stu.ecnu.edu.cn, daneezsd@163.com*

East China Normal University, Department of information management, Shanghai (China)

² *jiangen@utk.edu*

School of Information Sciences, The University of Tennessee, Knoxville, USA

Abstract

Aiming to explore the distributions of reference location in rhetorical sections of research articles and examine the relationship between the selection of rhetorical structure and the characteristics of references, we conducted a scheme with six types of rhetorical roles of literature review in research articles and applied to 4,931 JASIST research articles. The results showed that the pattern of selecting rhetorical structure of literature review in JASIST have evolved three stages from citing throughout to conventional literature review. Conducting a standard literature review section has become predominant for preparing scientific papers since early 21st century. We also found that the selection of rhetorical structure regarding literature review affect the disciplinary distribution of cited references. We will further investigate whether the rhetorical roles of distribution are related to the individuals' knowledge structure in selecting references.

Introduction

To facilitate scholarly communication in the context of globalization and international collaboration in science, conventional rhetorical structures are used by scientists in the field of their scientific interest (Kanoksilanpatham, 2005). Swales' approach to genre analysis has enhanced our understanding of rhetorical structure of the research articles (RA) that follow the conventional IMRaD (Introduction, Methods, Results, Discussion) macro-structure (Swales, 1990). Many studies have analyzed the rhetorical structure of individual sections of RAs in various disciplines, such as studies of computer science (Posteguillo, 1999), applied linguistics (Yang and Allison, 2004), biochemistry (Kanoksilanpatham, 2005), and comparative studies of multiple disciplines (Lin and Evans, 2012). In a microsense, rhetorical moves in each of the four main sections of RAs have been also well-studied (Bruce 2008; Peacock, 2012). As far as Literature Review section is concerned, it is less considered in the macro-structure in these analyses of RA rhetorical structure, though it is a common section and a necessary argumentative step in RAs. The function of Literature Review is also implemented as rhetorical moves in Introduction section. According to the Swales' CARS model (Swales, 1990), RA introductions contain three obligatory moves: (1) establishing a territory; (2) establishing a niche; and (3) occupying the niche. Move 1 and Move 2 of an RA introduction may fully implement the function of a Literature Review if there is no separate Literature Review section in the RA. The variation and effects of literature review in rhetorical structure of RAs largely remains an unexplored territory.

References cited in rhetorical sections have been found to have different characteristics. For example, the age of references varies by section, with older references in the Methods and more recent references in the Discussion (Bertin, 2016). Highly cited can be mainly selected for Methods purpose (Small, 2018), but also can be largely found in Introduction, Literature Review, and Discussion (Ding, Liu, Guo and Cronin, 2013; Thelwall, 2019). Hu, Chen and Liu (2013) found that citations are more likely to locate in the section of Introduction. Therefore, the selection of rhetorical structure for RA may affect the citing behavior of what kinds of reference would be cited in scientific writing.

Previous studies have found some factors that affect the author's citation behavior. For example, Milojević (2012) concluded that the collaboration level is an important factor related to citing behavior, especially in fields with extensive collaborative practices. A great deal of research has been done on citing behavior is about the reason for citing references, such as citing is guided by self-interest (Merton, 1957) or personal character (Kaplan, 1965). Lipetz (1965) characterized 28 and Duncan, Anderson, and McAleese (1981) concluded 26 different reasons for citing references. There are also many studies on the citation motivation. Garfield (1965) divided citation motivation into 15 categories, such as paying homage to pioneers, giving credit for related work, identifying methods, etc. Later lots of studies have listed various citation motivations from different perspectives. (Jurgens, Kumar, Hoover, McFarland & Jurafsky, 2018).

The selection of rhetorical structure determines how we organize and present a literature review in RAs in significant ways. For example, the distribution of cited references in a RA is greatly affected if the RA has an independent literature review section. Besides, since authors tend to cite different references across rhetorical sections, we believe the rhetorical structure regarding literature review may affect what references are cited and the rhetorical structure of scientific articles may also be a negligible factor for explaining citing behavior. Even the potential implications of rhetorical structure in citation studies, to our best knowledge, we have not found existing studies have examined the impact of the selection of rhetorical structure on citing behavior. In this article, we identify the evolving pattern of rhetorical structure regarding literature review by exploring the main sections for literature review in RAs. We also examine the relationship between the selection of rhetorical structure and the characteristics of references. We aim to answer the following research questions:

- (1) How does the role of literature review evolve in rhetorical structures of RAs?
- (2) How does the selection of rhetorical structures affect what kinds of references being cited in the literature review?

Research design

Data collection

We selected the Journal of the Association for Information Science and Technology (JASIST), a well-known peer review journal in the field of library and information science (LIS), as the data source. Full text and metadata of 5,623 articles in JASIST from 1950 to 2018 were downloaded in June 2019. We excluded several types of articles including brief communication, opinion paper, call for papers, and biographies. 4,931 articles in total were finally considered in our study. The metadata of all the references of 4,931 articles were downloaded as well in Web of Science, including the authors, journals, and publication years. Web of Science only provides JASIST information since 1985. So, 127,356 references since 1985 in total were collected and classified into LIS and Non-LIS categories based on the Journal Citation Report classification.

Coding process

First, to build up a coding scheme in terms of rhetorical structure, we randomly chose 100 articles (six to eight articles in every ten years by a random number generator) as coding samples. We manually read through all the articles and summarized the resemblances of how references distribute in an article based on the literal names of section headings. The resemblances eventually formed into six types of rhetorical structure regarding literature review. Type A and B both contain literature review sections. Type C and D are in line with IMRaD format. Second, we trained the coder to learn the six types among the coding samples.

And the coder passed a validation test via another 50 articles by an accuracy of 92%. Finally, the coder manually classified all the rest of articles.

- A. Articles with specific sections entitled Literature Review, Related Work, etc.;
- B. Articles with specific sections about literature review without entitled as Type A, but as headings with relative topics;
- C. Articles that include literature review in the introduction section. A brief but standard literature review was placed in the end of or spread through introduction;
- D. Articles that include literature review in the background section. Several citations instead of a standard literature review were placed in the beginning of introduction;
- E. Articles without literature review but citations are spread throughout the full text;
- F. Articles without any reference.

Findings

Overview

5.84% of total articles did not cite reference. They are mostly introduction of new concepts, techniques, and case studies. The most popular rhetorical structure in JASIST is surprisingly Type E, which exists the whole publishing life of JASIST. 45.79% of articles have cited reference all over the article. Type E dominates all indicators in Table 1, except the number of total reference and per article. The conventional structure of literature review, Type A, is secondarily most applied in JASIST articles. It showed appearances in most of the years. We observe that literature review in the introduction and background sections can be fairly acceptable in scientific writing. They yielded to almost 25% of all articles. Interestingly, Type B is the least the choice among JASIST authors, yet it occupied the largest number of references per article. It indicates the topical wise of literature review can possess a large proportion of references in individual articles compare to other structure.

Table 1. Statistics of rhetorical roles of literature review

Type	# articles (N)	% total	# appearance year (A)	Ave # per appearance year (N/A)	# total reference (R)	Ref per article (R/N)
A	1,178	23.89%	51	23.10	43,760	37.15
B	83	1.68%	24	3.46	3,788	45.64
C	817	16.57%	48	17.02	31,760	38.87
D	307	6.23%	34	9.03	12,788	41.65
E	2,258	45.79%	69	32.72	35,260	15.62
F	288	5.84%	34	8.47	0	0.00
Total	4,931	100%	69	71.46	127,356	25.83

Note. Appearance year is a year when a type of rhetorical roles appeared.

The distributions of rhetorical roles of literature review in research articles

The annual and proportional distributions certainly demonstrate couples of patterns with dynamic stages' shifts. We divide the total phrase of nearly 70 years into three stages.

Stage 1: Format-less stage (early time to the 1980s). Before scientific articles were published with a globally accepted format, citing behavior was even impossibly measurable and predictable. Although the IMRaD started to be adopted in science in the 1950s, the JASIST authors seemed to separate the roles of writing standard and citing preferences in this stage.

Stage 2: Formatting stage (the 1980s to the early 21st century). Along with the widely accepted standard for preparation of scientific articles (Day & Gastel, 2006), updated APA style (the only format accepted for JASIST all the time) version by version, and the rapidly increasing number of publications, the diversity of citing preferences in JASIST started to make a

difference in this stage. Even though the scattered citing style (D) remained predominant, the emergence of other types of rhetorical roles began blossoming.

Stage 3: Standardized stage (the early 21st century till now). Conducting literature review sections in academic writing overthrew the domination of various cohabit citing behaviors. Citing reference in the conventional sections has become a trend in scientific publishing. Moreover, a norm of citing preferences in the specific literature review section appeared to be fundamental to publish in JASIST.

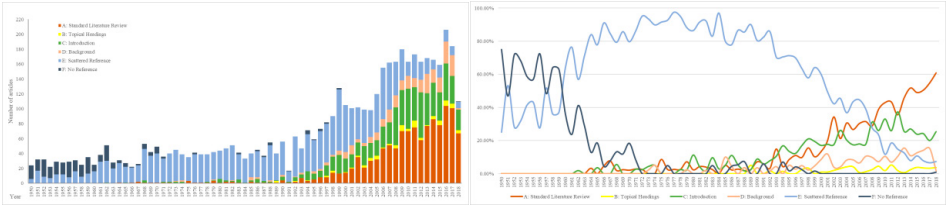


Figure 1. The annual distribution of rhetorical roles of literature review (left)
Figure 2. The annually proportional distribution of rhetorical roles of literature review (right)

Figure 3 compares the overtime changes of articles with literature review sections (Types A, B, C, D) and articles without literature review sections (Type E). Obviously, the proportion of articles with literature review sections has been increasing continuously. Although Type E dominated for the most of the time, especially before 2001, the turning point occurred in the early 1980s. This is when the American national standard for the preparation of scientific papers for written or oral presentation (ANSI Z39.16-1972) was published in 1972 and again 1979 (Day and Gastel, 2006). After 2001, the proportion of articles with literature review sections has begun to be higher than that of articles without literature review sections. The latter was below 50% for the first time. Presumably, the community of library and information science attracted and required researchers to submit manuscripts according to standard writing manual. Therefore, literature review has become an indispensable part in terms of scholarly communication standard since 2001. The growth rate increased from 2001 to 2010 and slowed down from 2010 to 2018 (a current peak has appeared). This slightly stability may result from the words limitation in the recent submitting instruction from Wiley.

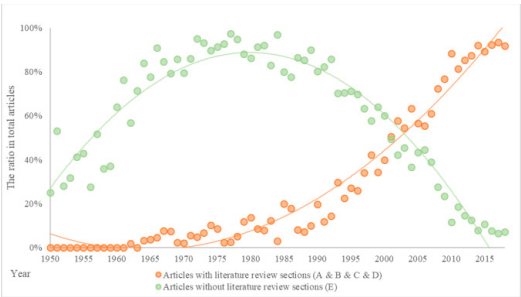


Figure 3. Comparison of articles with and without literature review sections

Note. The ratio is calculated from the total number of different types of rhetorical roles in one category dividing the total number of articles. Hereinafter as Figure 4 and 5.

Type A, B, C, and D all represent articles which contain literature reviews to some degree. Figure 4 shows the comparison between inner groups. The two trends are similar (Spearman correlation test result is 0.832, $P < 0.01$) and generally show an upward trend. However, the slopes of them illustrate that the conventional literature reviews have drawn more attention than the other, especially after the early 21st century. The growth rate of those articles with specific

literature review sections increased even more since 2012 with a current peak of 64.55% in 2018. This interests us. JASIST has started to limit word count currently. We presume the new policy would have affected the writing pattern regarding the selection of rhetorical structure in terms of literature review as literature review is normally considered as a word-demanding section. The results failed our assumption.

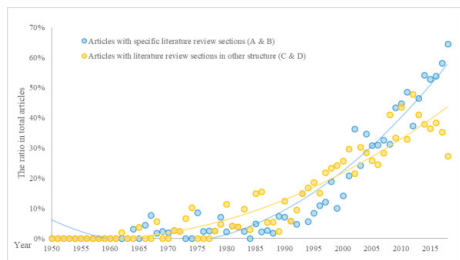


Figure 4. Comparison of articles with specific literature review sections and within other parts of manuscripts (Left)

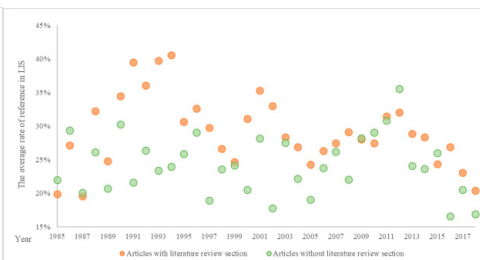


Figure 5. Comparison of reference in articles with and without literature review sections (right)

The relationship between rhetorical structures and references

The analyses above have given the distributions of rhetorical structures from the perspective of publications. We collected the references of all the publications to examine the distributions of the relationship between rhetorical structure and references. We composed a comparison in Figure 5 between Type A and Type E, respectively representing articles with and without literature review sections. We assume that articles with literature review section would contain more LIS references (references were published in LIS journals) than those without literature review section. Mann - Whitney U test was conducted to examine whether articles with literature review section would have relationship with the results of these articles have cited. The results show the significant difference between the articles with literature review sections and the ones without such sections ($U = 274$, $z = -3.738$, $P = 0.000 < 0.005$). Apparently, articles with literature review section have distributed on the higher rate than the other in Figure 5 (logistic regression result shows the difference is 5.83 times). This suggests the conventional literature review clusters inner-disciplinary topics rather comprehensive topics.

Conclusions and future research

In this article, we made a preliminary statistics and analysis on the distribution of the literature review in rhetorical sections. The pattern of selecting rhetorical structure of literature review in JASIST has been three stages from citing throughout to conventional literature review. Conducting a standard literature review section has become predominant for preparing scientific papers since early 21st century. Our preliminary results also show that the rhetorical structure may have effects on the disciplines where cited references are from. Future studies will consider how the selection of rhetorical structure affects selecting the types of cited references, and how these citations in turn impact on the knowledge structure of the essay. Furthermore, we will examine the selection preferences of rhetorical structure from the authors' perspective and disciplinary differences.

Acknowledgements

This research was supported by China Postdoctoral Science Foundation (2019T12035).

References

- Bertin, M., Atanassova, I., Gingras, Y., & Larivière, V. (2016). The invariant distribution of references in scientific articles. *Journal of the Association for Information Science and Technology*, 67(1), 164-177.
- Bruce, I. (2008). 'Cognitive genre structures in methods sections of research articles: A corpus study,' *Journal of English for Academic Purposes* 7/1: 38–54.
- Day, R.A. & Gastel, B. (2006). *How to write and publish a scientific paper (6th ed)*. Greenwood Press, Westport.
- Ding, Y., Liu, X., Guo, C., & Cronin, B. (2013). The distribution of references across texts: Some implications for citation analysis, *Journal of Informetrics*, 7(3),583-592,
- Duncan, E.B., Anderson, F.D., & McAleese, R. (1981). Qualified citation indexing: Its relevance to educational technology. *Proceedings of the First Symposium on Information Retrieval in Educational Technology*, pp, 70–79.
- Garfield, E. (1964). Can Citation Indexing Be Automated? *Symposium on Statistical Assoc Methods for Mechanized Documentation*, 84-90.
- Hu, Z., Chen, C., & Liu, Z. (2013). Where are citations located in the body of scientific articles? A study of the distributions of citation locations. *Journal of Informetrics*, 7(4), 887-896.
- Jurgens, D., Kumar, S., Hoover, R., McFarland, D. & Jurafsky, D. (2018). Measuring the evolution of a scientific field through citation frames. *Transactions of the Association for Computational Linguistics*, 6,391-406.
- Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for specific purposes*, 24(3), 269-292.
- Kaplan, N. (1965). The norms of citation behavior: prolegomena to the footnote. *Journal of the Association for Information science and Technology*, 16(3), 179-184.
- Lin, L. and S. Evans. 2012. 'Structural patterns in empirical research articles: A cross disciplinary study,' *English for Specific Purposes* 31/3: 150–60.
- Lipetz, B. A. (1965). Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators. *Journal of the Association for Information Science and Technology*, 16(2), 81-90.
- Merton, R. K. (1957). Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review*, 22(6), 635-659.
- Milojević, S. (2012). How are academic age, productivity and collaboration related to citing behavior of researchers? *PloS One*, 7(11), e49176.
- Peacock, M. 2002. 'Communicative moves in the discussion section of research articles,' *System* 30/4: 479–97.
- Posteguillo, S. 1999. 'The schematic structure of computer science research articles,' *English for Specific Purposes* 18/2: 139–60.
- Small, H. (2018). Characterizing highly cited method and non-method papers using citation contexts: The role of uncertainty. *Journal of Informetrics*, 12(2), 461-480.
- Swales, J. M. 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Thelwall, M. (2019). Should citations be counted separately from each originating section? *Journal of Informetrics*, 13(2), 658-678.
- Yang, R. and D. Allison. 2004. 'Research articles in applied linguistics: Studies from a functional perspective,' *English for Specific Purposes* 23/3: 264–79.
- Wu, J. (2011). Improving the writing of research papers: IMRaD and beyond. *Landscape Ecology*, 26(10), 1345-1349.