

Using Multidimensional Analysis to Investigate the Discussion Sections of Research Articles in Chemical Engineering

Bixi Jin (The University of Hong Kong, Hong Kong)

Introduction

Driven by the 'publish-or-perish' academic culture, there is increasing pressure on scientists and postgraduate science students to publish in prestigious high-impact journals, either for the dissemination of research recognized by international academic community or for professional advancement and promotion. In the process of writing up research articles, previous studies have shown that writing the discussion of results section can be a particular challenging undertaking for EAL novice scientists (e.g. Bitchener and Basturkmen, 2006). This ongoing study investigates the variation in underlying communicative functions in Discussion sections of high- and low- impact research articles in engineering. It is attested in the literature that no comparisons have been made in this aspect. The comparison in this sense is important as it can directly make less experienced scientists familiarize with and aware of the prevalent discourse patterns of discussing research findings in these 'successful' and 'less successful' models of discussion of findings, so as to navigate their participation in international academic discourse community.

Corpus compilation and analysis

A total of 213 RA Discussion sections in chemical engineering were searched and compiled for the analysis based on citations and journal five-year impact factors. One sub-corpus consists of 113 discussion sections coming from the highly cited research articles of high-impact journals published over 2005 to 2015. The selection was to ensure that the texts in this corpus have good quality both in terms of language and content, and thus can be regarded as "situationally effective and the results of expert performance" (Bazerman, 1994, cited in Yang, Zheng and Ge, 2015). The other corpus was made up of 103 discussion sections of research articles with few citations in the less recognized peripheral journals. It should be noted that the selected texts are limited to a single discipline (i.e., chemical engineering in this study), as there might be potential differences arising from the different disciplines.

Multi-dimensional (MD) analysis developed by Biber (1988) was used as the major approach for comparison by looking at the interaction among a range of linguistic features, as single linguistic features cannot reliably account for the description of variation in language use (Biber and Conrad, 2001). Using MD to investigate variations of part-genre of research article has been rare, with the exception of Xiao and Cao's (2013) study of abstract written by NS and NNS writers. *Multi-dimensional Analysis Tagger* (Nini, 2014) was adopted for automatic tagging. The tagged features were then normalized to per 100 words considering the average length of the individual texts. Principle Component Analysis (PCA) was then conducted on the normalized features using SPSS 22.0. The six-factor solution was determined according to the observation of the scree plot. The first six factors account for 39.284% of the total variance.

In order to look into the significant differences of RA discussion sections in different types of journals, the mean scores of each dimension were computed using z-scores for the two corpora and then subjected to an Independent Sample T-test for analysis.

Preliminary findings

The PCA analysis reveals the six underlying communicative functions of discussion sections in the Corpus of High-impact articles (**CHA**) and Corpus of Low-impact Research Articles (**CLA**). The six dimensions with proposed descriptive labels are:

Table 1 Composition of the positive and negative features of Dimension 1-6

Dimension	Linguistic features and factor loadings
Dim 1: Involvement and interactivity	Split auxiliaries(0.746), total adverbs (0.725), amplifiers (0.485), modals (0.485), conjuncts (0.428), perfect aspect (0.408)
Dim 2: Narration vs. non-narration	Present tense (0.848), pronoun <i>it</i> (0.449), <i>that</i> -deletion (0.342), place adverbials (0.312), past tense (-0.779)
Dim 3: Further explanation and elaboration on evaluative expressions	Predicative adjectives (0.856), <i>be</i> -verb (0.847), subordinators (0.503), demonstratives (0.402)
Dim 4: Informational density	Average word length (0.686), attributive adjectives (0.683), <i>do</i> (0.486), common nouns (0.454)
Dim 5: Stating results/claims	<i>that</i> -complements (0.698), private verbs (0.615), public verbs (0.458), first-person pronouns (0.4)
Dim 6: Expression of denial relationships in experimental findings	Negation (0.642), existential <i>there</i> (0.496), <i>to</i> -infinitives (0.477), prepositional phrases (-0.408)

Dimension 1: Involvement and interactivity

Dimension 2: Non-narration vs. narration

Dimension 3: Further explanation and elaboration on evaluative statement

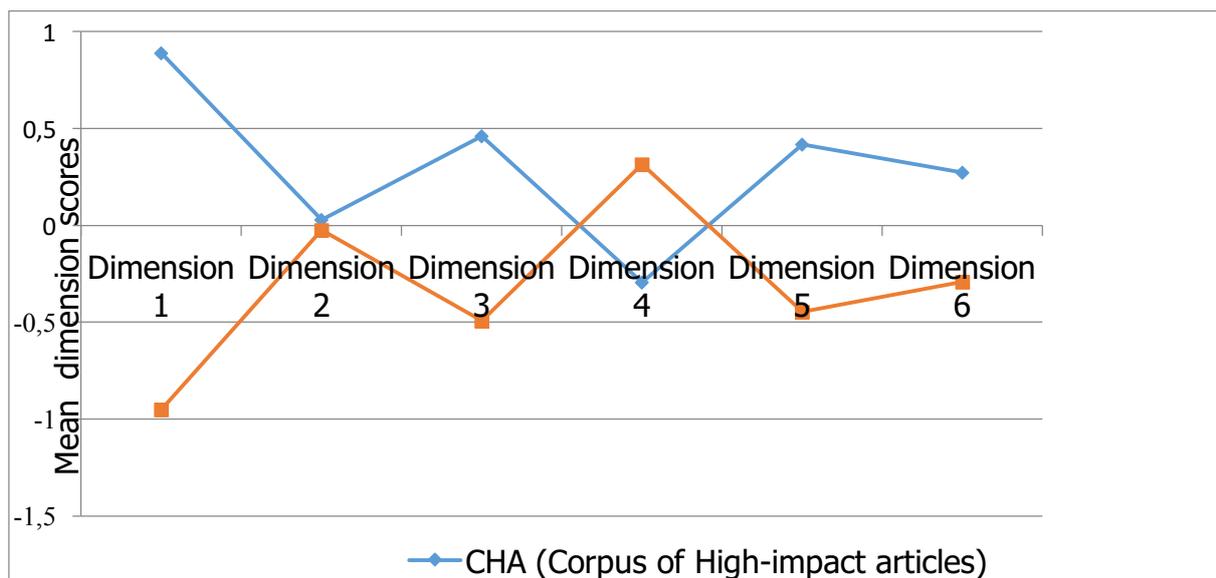
Dimension 4: Information density

Dimension 5: Stating results/claims

Dimension 6: Expression of denial relations towards statement or experimental findings

The follow-up independent sample T-test compared and contrasted the differences of discussion sections in two types of journals along these six dimensions. Dimensions 1, 3, and 5 are found to be significantly different in two corpora.

Figure 1
Mean dimension scores of the RA discussions along six dimensions



CHA and CLA on Dimension 1

The T-tests show that RA discussions in CHA appear to discuss their results by employing a constellation of linguistic features associated with interactivity and involvement, while this is not the norm in the CLA. That is, authors in CHA are capable of using more metadiscourse devices like hedges, modal verbs and boosters to manage a good relationship with readers and make their claims acceptable by their disciplinary communities (see Hyland, 1996; Hyland, 2005; Yang, Zheng, and Ge, 2015), whereas authors in CLA use metadiscourse-related features in a less frequent manner. The contrasting of using hedging devices in CLA is consistent with a series of findings (see ElMalik and Nesi, 2008; Hu and Cao, 2011; Yang, 2013). The possible explanation for this could be related to influence of two different scientific communities (international vs. local), as the majority of the discussion sections of low impact articles included in the subcorpus are national-based English journals (e.g., the *Chinese journal of chemical engineering*, the *Korean journal of chemical engineering*, etc), which attract a fair number of researchers in these countries and regions to publish in them, though the journals aspire to attract more international audiences. The research writers from these local scientific communities, as suggested by Loi, Lim, and Wharton (2016), are inclined to display a less reader-friendly style of writing. That is, they do not invite readers to be actively engaged in research, but require them to participate in and interpret the intention and argument on their own.

CHA and CLA on Dimension 3

In Dimension 3, the result suggests that the preferred communicative styles of discussing results in CHA often constitute a evaluative sentence followed by elaboration and explanation on this statement by making using of adverbial subordinators such as *if*, *because*, *though*, *whereas*, and *since* to indicate the relationship of time, reason, condition, and comparison. On the other hand, discussions in CLA make fewer comments on research findings, and have less concern for details of the use of these corresponding

adverbials that makes their evaluation of findings specific. Similarly, the reason could be partly related to the fact that writers from local scientific communities are less likely to adopt a stance in their work by using value-laden words, making it "more prominent and appealing" (Loi et al., 2016, p.12) due to the less competitive local publishing context.

CHA and CLA on Dimension 5

A final interesting point of difference is in Dimension 5. The discussions in CHA display obvious authorial presence to make explicit claims by making use of the complement clauses controlled by verbs and also the first-person pronouns. Discussions in CLA have a tone of objectivity that avoid direct referring to themselves, but give explicit reference to the research results (e.g., *these results indicate...*, *our results implied...*). The revealed difference partially aligns with Hyland's (2002) finding that novice writers appear to conceal their role due to a lack of confidence, they "consciously avoided the most authoritative functions and sought to deny ownership and responsibility for their views" (p. 1107). In contrast, expert writers are more inclined to make knowledge claims to establish "a personal authority based on confidence and command of their arguments" (p.1104). Although no absolute claims can be made that discussions of high-impact journals are all written by expert writers, these successful models can be considered as 'expert performances' (Bazerman, 1994), and can thus partially allude to Hyland's (2002) findings that suggest the use of first person pronouns to "speak as an authority" is important for successful academic writing (p.1094). Also, the findings may also point to the possibility that maintaining objectivity may still be a rooted mindset for local scientific communities.

CHA and CLA on Dimension 2, 4, 6

In contrast, no significant statistical significance can be found in Dimensions 2, 4 or 6 and thus are not discussed with further details.

Conclusions

The corpus-based comparative study addresses two research questions, i.e., to describe the underlying functional dimensions of 213 RA discussions, and to explore the linguistic variation of discussion sections of high- and low-impact RAs. Six dimensions using Principle Component Analysis (PCA) were generated and three of them were used as the focal aspects for comparison. The significant variation found in Dimension 1, 3, and 5, suggesting that successful models of RA discussions incorporated proportionally more metadiscursive features, first-person pronouns and evaluative statements with further explanation. The differences may be related to the 'expert' and 'non-expert' performances exhibited by two groups of researchers.

The contrastive findings have pedagogical implications for novice scientists and EAP practitioners. Less experienced scientists can adapt their writing of this section to RA discussions of high-impact articles published in high-profile journals, thus enabling their work to achieve greater international visibility. EAP writing instructors can make use of the contrastive findings to develop corpus-informed teaching materials, and through a series of consciousness-raising tasks, help novice writers notice the linguistic characteristics underlying successful RA discussions and skillfully manipulate linguistic patterns to produce convincing and compelling claims in their manuscripts.

References

- Bazerman, C. (1994). *Constructing experience*. Carbondale: Southern Illinois University Press.
- Bitchener, J., & Basturkmen, H. (2006). Perceptions of the difficulties of postgraduate L2 thesis students writing the discussion section. *Journal of English for Academic Purposes, 5*(1), 4–18.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D., & Conrad, S. (2001). Multi-dimensional analysis and the study of register variation. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp.3-12). London. England: Longman.
- Cao, Y., & Xiao, R. (2013). A multi-dimensional contrastive study of English abstracts by native and non-native writers. *Corpora, 8* (2), 209–234.
- ElMalik, A. T., & Nesi, H. (2008). Publishing research in a second language: The case of Sudanese contributors to international medical journals. *Journal of English for Academic Purposes, 7*, 87-96.
- Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics, 43* (11), 2795–2809.
- Hyland, K. (1996). Talking to the academy: Forms of hedging in science research articles. *Written Communication, 13* (2), 251-281.
- Hyland, K. (2002). Authority and invisibility: authorial identity in academic writing. *Journal of Pragmatics, 34*, 1091-1112.
- Hyland, K. (2005). *Metadiscourse: exploring interaction in writing*. London: Continuum.
- Loi, C.K., Lim, J.M-H., & Wharton, S. (2016) Expressing an evaluative stance in English and Malay research article conclusions: International publications versus local publications. *Journal of English for Academic Purposes, 21*, 1-16.
- Yang, A., Zheng, S.Y., & Ge, G.C. (2015). Epistemic modality in English-medium medical research articles: a systemic functional perspective. *English for Specific Purposes, 38*, 1-10.