Corpus-based Research on Tense Analysis and Rhetorical Structure in Journal Article Abstracts

Pin-ning Tu
National Taiwan University of Science and Technology, Taiwan/ #43, Sec. 4, Keelung Rd., Da'an Dist., Taipei 106, Taiwan (R.O.C.)
promise523@hotmail.com

Shih-Ping Wang
National Taiwan University of Science and Technology, Taiwan/ #43, Sec. 4, Keelung Rd., Da'an Dist., Taipei 106, Taiwan (R.O.C.)
spwang2005@yahoo.com.tw

Abstract

There has long been a growing interest in journal articles (JA) abstract writing, and this pervading interest has boosted the exigency for further instructive research. This current study aims to investigate both the variant application of the verb tense as well as the rhetorical structure within JA abstracts. A 9.9 million word corpus of 1000 JAs was collected based on four prestigious journals, i.e., *Journal of Pragmatics, Journal of Research in Reading, Journal of Second Language Writing,* and *Reading and Writing*, respectively. The quantitative analysis indicates the tendency of tense shown in the commonly applied reporting verbs. On the other hand, the qualitative analysis shows the prevailing adoption of three-, four-, and five-move theories in terms of the CARS model, the IMRD structure, and the IPMPrC structure. The results not only reveal the explicit tendency of the variance within reporting verbs but also suggest a distinct pervasiveness of the IMRD structure over the other models. These findings not only present a more systematic pattern within JA abstracts, but also show potentials for enlightening further pedagogy-oriented composition instruction for JA abstract.

1. Introduction

Previous studies have highlighted the indispensable importance of JA abstract in the contemporary flow. The pivotal role of JA abstract has received considerable attention in academic written genre among the international community. Swales (1990) appeals to the academia, claiming that the research in JA abstracts ought not to be ignored inasmuch of its influential significance upon the genre investigation and disciplinary discourse communities.

As the knowledge of proficient preferences for language choice as well as rhetorical structure has a great influence on academic written genre, many investigators have recently turned to the relevant research in relation to genre analysis, thematic organization, formulaic language, rhetorical structure, etc. (Cortes, 2004; Hyland, 2008a; Lorés, 2004; Martín, 2002; Swales, 1990; Wang & Chan, 2011; Wang & Kao, 2012). Furthermore, research in terms of corpora decoding for rhetorical structures such as moves and steps is also regarded as one of the recommendations for further research expansion by Flowerdew (2010).

Taking the contribution of the previous studies, this current reserach sets out to explore the variation of tense within the reporting verbs among the transitions of moves via the structural analysis in JA abstracts.

2. Literature review

In the respect of genre analysis, move analysis has been always considered to be one of the most influential elements. A move is a rhetorical element which serves the function of correlating and cohering within the written or spoken context (Lorés , 2004; Swales, 2004). However, it is not a definite unit which is constraint to perform in a fixed pattern because it is able to vary along with the context. In other words, move functions as a communicative role between each transition of the rhetorical structure.

A brief elaboration of the most pervasive move theory in terms of three-, four-, and five-moves is described in sequence.

2.1. Three-move theory

Create a Research Space (CARS) model, proposed by Swales (1990), has been widely used by scholars to outshine their publication in this competitive academia (Cheng, 2006). CARS model is divided into three moves,

including establishing a territory, establishing a niche, and occupying the niche (Swales, 1990).

In other words, Move 1: establishing a territory can be commensurate with "goal" and "current capacity," synthesizing the research aim with the previous research (Swales, 1990, p. 142). As a consequence, Move 2: establishing a niche functions as offering a space for research gap and possible research questions (Swales, 1990). Under this circumstance, Move 3: occupying the niche will provide a "solution of criteria of evaluation" that taps into the intricacies which came up with in Move 2 section (Swales, 1990, p. 142).

2.2. Four-move theory

The most well-known and considerably applied structure in academic writing is the IMRD structure (i.e., Introduction, Methods, Results, and Discussion) (Golebiowski, 2009). It was first proposed by Ventola (1994).

To illustrate the content in depth, the introduction segment would cover the further elaborations of the purpose and objective of the current research. Lorés (2004) additionally comments that any other questions that could possibly bring out further open discussion might also be included in this Introduction segment. When it comes to the second stage – Method, a clarification of the scheme adopted in the research will be described (Lorés, 2004). As the lines progress, the Result segment is expected to offer critical information in relation to the findings from the implement of the research (Lorés, 2004). The final Discussion section is required to contain a further discussion of the findings, an exploration of possible research space and practical application (Lorés, 2004).

2.3. Five-move theory

Differing from the above discussed structures, the five-move theory – IPMPrC structure, proposed by Hyland (2004), is especially designed to access the RA abstracts. Nevertheless, it is clarified in the first place that the aim of setting this five move structure lies in providing an assertion in relation to JA abstract conducting as well as an inter-textual projection in terms of the significance of each research instead of addressing definite move steps (Hyland, 2004).

In an attempt to provide a clearer framework of the main characteristics of the IPMPrC structure, Table 1 elaborates the primary functions of each move in the five-move theory (Hyland, 2004, p. 67).

Move	Function
Introduction	Establishes context of the paper and motivates the study or discussion.
Purpose	Indicates purpose, thesis or hypothesis, outlines the intention behind the paper.
Method	Provides information on design, procedures, assumptions, approach, data, etc.
Product	States main findings or results, the argument, or what was accomplished.
Conclusion	Interprets or extends results beyond scope of paper, draws inferences, points to applications or wider implications.

Table 1: IPMPrC structure

3. Methodology

In an attempt to shed lights on the various dimensions that are possibly exposed from the academic written genre, a total of 1000 journal articles, which comprises 9,983,482 tokens out of 117,855 types of distinct words, were extracted evenly from four prestigious academic journals: *Journal of Pragmatics* (JOP), *Journal of Research in Reading* (JRR), *Journal of Second Language Writing* (JSLW), and *Reading and Writing* (R&W).

In accordance with the principle aim of this current research, that is, to specify the variation of verb tense and rhetorical structure in JA abstracts, 1000 abstracts were additionally extracted from the retrieved research materials, and constructed as the primary research corpora. Table 2 compares the tokens and types of the five primary corpora.

Type of Corpora	No.	Tokens	Types
JOP abstracts corpus	250	47,074	6,068
JRR abstracts corpus	250	39,699	4,001
JSWL abstracts corpus	250	45,520	4,265
R&W abstracts corpus	250	45,700	3,679
1,000 abstracts corpus	1,000	177,945	9,711

Table 2: Comparison among research corpora

The data analysis is twofold. On the one hand, the quantitative analysis focuses on the investigation of verb tense, especially set out for reporting verbs, by manipulating the analytical instruments such as *MonoConc Pro* and *WordSmith version 5.0*.

On the other hand, the qualitative analysis of consists in the assessment of the rhetorical structure in accordance with CARS model, the

IMRD structure, and the IPMPrC structure. Table 3 below illustrates the comparison of different transitions among the applied theories from three to five moves.

Three moves	Four moves	Five moves
CARS model	IMRD structure	IPMPrC structure
Context	Introduction	Introduction
Gap	Methods	Purpose
Present study	Results	Methods
	Discussion	Product
		Conclusion

Table 3: Comparison of move theories

4. Results

The current study reports on two dimensions in JA abstracts: verb tense and rhetorical structure. Firstly, the analysis of verb has uncovered a prevailing application of be-verbs, such as *is*, *are*, *was*, *were*, as well as reporting verbs such as *show*, *examine*, *suggest*, *investigate*, and *find*, in applied frequency sequence. It is reasonably assumed that the different application pattern of be-verb contains two possibilities: plain statement as well as passive voice.

Prior to tackling the various findings on the different tense of verbs, a clearer comparison of how be-verb is applied in each of the research corpora is shown in Table 4.

Form	JOP	JRR	JSLW	R&W	ALL
is	**526**	200	275	173	1204
are	**329**	206	213	153	901
was	75	**300**	134	**355**	864
were	83	**374**	243	**442**	1142
be	192	110	140	94	536
total	1,205	1,190	1,005	1,217	4,647

Table 4: Frequency of be-verbs

As can be seen from Table 4, it is apparent that JOP has its tendency to use present tense whereas JRR and R&W has similar tendency to apply past tense. This finding also reflects on the results obtained from the reporting verbs, as shown in Table 5.

Form	JOP	JRR	JSLW	R&W	ALL
show	**63**	21	32	19	135
-s	22	8	11	2	43
-ing	3	6	1	4	14
-ed	13	**77**	26	**88**	204
examine	**19**	16	21	40	96
-s	**32**	16	25	5	78
-ing	7	3	12	7	29
-ed	14	**53**	24	**81**	172
suggest	26	31	39	46	**142**
-s	15	23	23	7	**68**
-ing	2	8	9	9	28
-ed	9	9	20	12	50
investigate	**19**	19	18	25	81
-s	**20**	5	16	5	46
-ing	5	4	6	5	20
-ed	13	**32**	28	**46**	119
find	5	15	14	5	39
-s	2	1	----	----	3
-ing	----	----	2	----	2
-ed	**38**	**54**	**51**	**60**	**203**

Table 5: Frequency of reporting verbs

It can be seen from the data shown in Table 5 that the each research corpus has its tendency towards the application of tense as stated in the analysis of be-verbs, except for the single peculiar example "find." When it comes to the reporting verb "find," it was found that this particular word possesses its consistent tendency of past tense or passive voice in all reserach corpora. The practical examples of past tense and passive voice extracted from each research corpus are demonstrated below.

- *"Finally, the study **found** that while the majority of the complaints on TripAdvisor can be considered indirect (or third party) complaints, there were nevertheless some examples that blur the direct/indirect dichotomy."* (JOP – past tense)
- *"It **was found that** although the same discourse of legitimation (the Bible) is used in some of the arguments, addressers apply their own experience to their views of this discourse and thus create opposing arguments."* (JOP – passive voice)
- *"We **found** an association between the HLE and ethnicity/SES, indicating that (Dutch) majority children and children from high SES families had, in general, the most stimulating HLEs."* (JRR – past tense)
- *"It **was found that** letter knowledge was specifically related to the development of phoneme segmentation in pre-literate children."* (JRR – passive voice)

- *"While some **found that** peer comments were viewed with skepticism and induced little revision, others **found that** they did help learners to identify and raise awareness of their strengths and weaknesses in writing."* (JSLW – past tense)
- *"The best measure **was found** to be total words in error-free clauses."* (JSLW – passive voice)
- *"A principal components analysis **found** partial dissociability between higher-level skills including reading comprehension, vocabulary and print exposure, and lower-level skills including decoding and spelling in adult readers."* (R&W – past tense)
- *"Both word reading and comprehension **were found to be** highly stable, and genetic influences were primarily responsible for that stability."* (R&W – passive voice)

Turning to the exploration of the move analysis has revealed a wide application of three-, four-, and five-move theory in JA abstracts. The distribution of the applied rhetorical structure is shown in Table 6.

JA Abstract Corpora	CARS model		IMRD structure		IPMPrC structure	
	N	%	N	%	N	%
JOP	48	19.2	119	47.6	83	33.2
JRR	32	12.8	166	66.4	52	20.8
JSLW	71	**28.4**	95	**38**	84	**33.6**
R&W	22	8.8	159	63.6	69	27.6
ALL	173	17.3	539	**53.9**	288	28.8

Table 6: Application of move theories

As can be seen in Table 6, IMRD structure is the most commonly applied rhetorical structure. There are 53.9% over half of the JA abstracts which were found out to be written in accordance with the IMRD structure. Table 7 illustrates the transitions of the moves in IMRD structure, which was randomly selected from the JRR abstract corpus.

Move	JA abstract from JRR corpus
I	This semi-longitudinal study **examined** the development of narrative writing quality of young Turkish second language learners in mainstream Dutch-only education, and the impact of student-level and classroom-level predictors of narrative writing quality, using hierarchical linear modelling.
M	Writing samples of 106 third graders and 111 fourth graders of seven Flemish primary schools **were** collected at the beginning and at the end of the school year. Measures included one holistic primary trait

	judgement, and six objective indices of text quality. Student-level predictors included age, SES and home language, while the classroom-level predictor focused on the home language pattern of the classroom.
R	There **was** a significant mean growth for each index in each grade, but effect sizes differed from quite large for content and word level indices over moderate for sentence level indices to small for the text level index. Home language (Turkish) had a significant negative effect on all but one variables, particularly in Grade 4, while the negative effect of low SES was much more limited. A supplementary negative effect was found for homogeneity of classroom population.
D	**Implications** of the study highlight the importance of student and classroom characteristics in writing achievement as well as the need to consider the poor performance of Turkish children.

Table 7: Example of IMRD structure

Table 7 shows that the JA abstract was constructed with (I) the specific research focus, (M) the sampling and participants, (R) main findings, and (D) the derived implications. It can be especially observed from the highlighted boldfaced words in Table 7 that the tendency of verb tense and the reprting verb choice were found out to be reflecting the findings, which indicated the tendency of past tense as well as passive voice applied in JRR corpus, stated in the former sections.

The second example, randomly selected from JSLW corpus (see Table 8), is the IPMPrC structure, which was found out to contain 28.8% of the JA abstracts written in this structural pattern.

Move	JA abstract from JOP corpus
I	Hedges and boosters **are** important metadiscursive resources for writers to mark their epistemic stance and position writer–reader relations.
P	Building on previous research that suggests notable cross-cultural and cross-linguistic differences in the use of hedges and boosters in academic discourse, this comparative study **investigates** the use of such discourse markers in academic article abstracts.
M	Based on a **corpus** of 649 abstracts collected from 8 journals of applied linguistics, this study **examines** if hedging and boosting strategies differ (a) between applied linguists publishing in Chinese- and English-medium

	journals and (b) between authors of empirical and non-empirical academic articles.
Pr	Quantitative analyses indicated that abstracts published in English-medium journals featured markedly more hedges than those published in Chinese-medium journals and that abstracts of empirical research articles used significantly more boosters than those of non-empirical academic articles. Textual analyses further revealed that the distinct patterning of hedges and boosters in Chinese and English abstracts had a joint, interactive effect on the authorial certainty and confidence conveyed therein.
C	These results **are discussed** in terms of culturally preferred rhetorical strategies, epistemological beliefs, lack of facility in English as a second/foreign language, and the nature of supporting evidence drawn on for knowledge claims in different types of academic writing.

Table 8: Example of IPMPrC structure

Table 8 presents the trasitions of the five-move theory from (I) establishing the context, (P) staing the main purpose, (M) methodology, (Pr) emerged observations, and (C) discussions. Interestingly, the contrary tendency of present tense is also reflected in the randonly selected example, as shown in Table 8.

Turning now to the findings of the three-move structure. It was revealed that only 17.3% of the JA abstracts were written in accordace with the CARS model; however, the minority does not shatter the irreplacable importance that each move theory possesses. The example randomly selected from the JSLW abstract corpus is demonstrated in Table 9.

Move	JA abstract from JSLW corpus
1	English as an Additional Language (EAL) students' textual borrowing in disciplinary writing **has attracted** wide research interest in recent years.
2	**However**, much of the research was conducted in the regular curriculum setting while the relevance of the issue in a writing-for-publication context has largely been overlooked. In particular, disciplinary experts' perspectives concerning textual borrowing **have not been explored in-depth**.
3	The present study **fills** such a gap in the literature by looking into how an expert writer, a professor of biochemistry in a Chinese university, perceived novices'

	textual borrowing in their initial drafts and eliminated such borrowing as he redrafted novice texts for publication. The study **revealed** that the expert had complete tolerance for his students' copying and that his elimination of it during redrafting was guided by his genre expertise and rhetorical skills for publishing. The paper also **pointed out** that the shortage of explicit teaching from the supervisor to his students as well as the lack of active participation of his students in the writing process was bound to the publication pressure in the local institutional context.

Table 9: Example of CARS model

It is clear that the transition of the three-move theory set out from (1) appealing to readers' attention, (2) indicating the gap, and (3) presenting the current findings. The highlighted boldfaced words imply the impartial selection of tense shown in the JSLW abstract corpus.

5. Concluding remarks

This current research aimed to investigate tense analysis as well as rhetorical structure in JA abstracts. The implemented data analysis has indicated the variant but consistent tendency of verb tense in each discipline. The findings of the verb tense are reflected in the assessment as shown in the transitions of rhetorical structure. It is believed that the variations of verb tense as well as the rhetorical structure are strongly connected and interrelated. Despite the lack of pedagogical experiment design that has led this study to the limitation of not knowing the actual benefit of the research findings for language learners, it is determined that the acquired language use and the structural pattern suggest authentic disciplinary conventions on linguistic productions.

References

Cheng, A. (2006). Analyzing and enacting academic criticism: The case of an L2 graduate learner of academic writing. *Journal of Second Language Writing, 15*(4), 279-306.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397-423.

Flowerdew, L. (2010). Using corpora for writing instruction. In A. O'keeffe & M. McCathy (Eds.), *The routledge handbook of corpus linguistics* (pp. 444-457). London: Routledge.

Golebiowski, Z. (2009). Prominent messages in Education and Applied Linguistics abstracts: How do authors appeal to their prospective readers? *Journal of Pragmatics, 41*(4), 753-769.

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor: The University of Michigan Press.

Lorés, R. (2004). On RA abstracts: from rhetorical structure to thematic organisation. *English for Specific Purposes, 23*(3), 280-302.

Martín, P. M. (2002). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes, 22*(1), 25–43.

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

Swales, J. M. (2004). *Research genres: Exploration and applications*. Cambridge, UK: Cambridge University Press.

Ventola, E. (1994). Abstracts as an object of linguistic study. In S. Cmejrkova, F. Danes & E. Havlova (Eds.) *Writing vs Speaking: Language, Text, Discourse, Communication*.(pp. 333–52). Tubingen: G. Narr.

Wang, S.P., & Chan, C.W. (2011). Research on wordlists, lexical bundles and text structures of journal article abstracts. In R.F. Chung, et al. (Eds.), *Diversity of Languages: Papers in Honor of Professor Feng-fu Tsao on the Occasion of his Retirement* (pp. 369-381). Taipei: Crane Publisher Co.

Wang, S.P., & Kao, C.L. (2012). Wordlists, clusters and structure in research article introductions. *Studies in English Language and Literature*, *30*, 27-43.