Generic Structure of Research Article Abstracts in Sciences

Budsaba Kanoksilapatham

Abstract

This paper has the objective of illustrating how genre analysis, a discourse analysis approach, can be used to explore organizational patterns prevalent in a set of scientific abstracts from multiple scientific disciplines. Results show that, within this single genre, variations in the use of certain moves across disciplines can be observed. However, these abstracts seem to share resemblance, using the same set of moves forming a particular sequence. While highlighting the integral role of genre analysis in analyzing texts of other disciplines, this study provides pedagogical implications, facilitating novice scientists to disseminate their research work internationally.

Key words: abstract, genre analysis, research article, rhetorical structure, science

Generic Structure of Research Article Abstracts in Sciences

Budsaba Kanoksilapatham

INTRODUCTION

Among a number of academic discourse genres, be they written or spoken (e.g., research articles, short reports, conference presentations, classroom lectures), abstracts have succinctly been recognized as one of the most common means of communication. Specifically in the academic field of sciences, the necessity for rapid dissemination of scientific knowledge across the globe contributes to increasing paramount importance of abstracts. By reading abstracts, scientists can stay connected, sharing results with peer colleagues and enhancing the quality of research in their disciplines. Since abstracts typically provide readers with a brief summary of the study being reported, they also play a crucial role impacting the acceptance or rejection of submitted research articles for publication, and consequently determining scientists' professional growth and success to a certain extent.

Currently, English has been promoted to be an international language in business and academic contexts. As a result, possessing knowledge of how scientific abstracts in English are constructed becomes essential in scientific academia. As a result, to survive, thrive, and prosper in their respective academic disciplines, scientists need to be able to write their English abstracts in a manner that is acceptable and conforming to the expectations of the target journal.

A discourse analysis approach of genre analysis has been known as an effective tool to provide insights onto how texts are constructed. According to Swales (1990, 2004), the creator of this approach, a genre is a recognizable communicative event characterized by a set of communicative purposes which are identified and mutually understood by the members of the professional or academic discourse community in which it regularly occurs. Usually, a genre is highly structured and conventionalized. The goal of genre analysis is to identify the rhetorical organization of texts belonging to a given genre. Based on this notion, the terms 'move' and 'step' are invented to refer to textual units of analysis. That is, a 'move' refers to a text segment that performs a communicative function, and a 'step' is a subunit of a move that, in turn, contributes to the move's communicative function. In addition to its communicative function, a move can be recognized by a certain set of co-occurring linguistic features. For instance, the implication move is usually signalized by the linguistic features of hedging and *that* clause. For more information, please refer to Kanoksilapatham (2003, 2007).

Genre analysis has been used as an analytical framework for a wide variety of discourses including academic discourse (e.g., Kanoksilapatham 2005, 2007; Ozturk, 2007; Pho, 2008) and professional discourse (e.g., Bhatia, 1993; van Mulken & van der Meer, 2005). In the abstract genre, a number of studies were conducted (e.g, Hyland & Tse, 2005; Martín, 2003; Melander et al., 1997; Samraj, 2005). Samraj (2005), for instance, compared the rhetorical organizations of introductions and abstracts in two disciplines of conservative biology and wildlife behavior. She found that

research article introductions and abstracts were overlapping in their rhetorical organization. Specifically, they were more alike in conservative biology than those in wildlife behavior. Martín (2003) investigated the abstracts written in two languages of English and Spanish in the field of social sciences. The comparison showed that, unlike English abstracts, Spanish abstracts tend to omit the result move. These previous genre-based studies seem to indicate that discipline and language variations in the abstract genre are discernible.

The study reported here aims to explore the rhetorical organization of abstracts written in English from a number of scientific disciplines using the analytical framework of genre analysis. The analysis reveals that the abstracts analyzed seem to display variations in rhetorical organization. Nevertheless, in scrutiny, the same set of moves is used in the abstracts across scientific disciplines. Furthermore, within the same set of moves, these moves tend to occupy a relatively fixed position in the abstracts. The findings suggest that the abstract genre is unique; it is formed by a limited set of five moves, and these moves interact with each other in a cohort pattern. The study is deemed to be beneficial for several stakeholders, including scientists in general and novice scientists in particular, graduate students in sciences, and ESP teachers who would like to enhance their students' competence in writing abstracts.

METHODS

The paper focuses on the analysis of a small set of research article abstracts from four academic disciplines namely, biochemistry, microbiology, civil engineering, and software engineering. These abstracts were randomly selected from their respective bigger pools of data or corpora that were systematically compiled at different points in time, for different research projects, and for different objectives. In a nutshell, these original corpora represent the most prestigious research articles in their respective disciplines. For more detailed descriptions of these corpora, please refer to Kanoksilapatham 2003, 2007. Given the limited length of abstracts in general (as opposed to that of research articles, for example), the abstracts were analyzed by genre analysis to delineate the textual organizational structure of texts in terms of moves only, revealing the types of information that are usually presented in the abstracts of scientific research articles, and the order in which the information is sequenced is also identified.

RESULTS

The analysis of abstracts across disciplines shows that, despite their superficial different manifestations, they are constructed from a common set of five rhetorical moves. They are: Move B (Background information), Move P (Purposive statement), Move M (Methodological description), Move R (Result announcement), and Move D (Discussion, conclusion, implication). The names of the moves are self-explanatory, justifying their communicative functions

in the abstracts. To visualize how move demarcations are identified, the following excerpts from biochemistry, microbiology, civil engineering, and software engineering research articles are illustrated. In each excerpt, two changes were made to the text. First, all of the sentences are itemized to facilitate any reference made to the texts. Second, double slant lines (//) were inserted to mark move boundaries.

Excerpt One: Biochemistry

Move B	**(S1)** Type III secretion systems mediate export of virulence proteins and flagella assembly subunits in Gram-negative bacteria. **(S2)** Chaperones specific to each class of secreted protein are believed to prevent
Move R	degradation of the secreted substrates. // **(S3)** We show that an additional role of chaperones may be to regulate translation of secreted proteins. **(S4)** We show that the chaperone FlgN is required for translation of the flgM gene transcribed from on mRNA transcript, but not from another. **(S5)** FlgM translated from the class 3 transcript is primarily secreted whereas FlgM translated from the class 2 transcript is primarily retained in the
Move D	cytoplasm. // **(S6)** These results suggest FlgM and other type III secretion substrates possess both mRNA and amino acid secretion signals, and supports a new role for type III chaperones in translation/secretion coupling.

Rhetorical structure: B-R-D

The above excerpt of 6 sentences is from biochemistry. Sentences 1 and 2 are classified as Move B (Background information), establishing the topic of the study. Move R (Result announcement) is relatively extensive, including Sentences 4 and 5 in entirety. Finally, Move D (Discussion, conclusion, implication) of Sentence 6 concludes the abstract. Excerpt One shows that move boundaries seems to, at this point, coincide with the syntactic structure of sentences. As will be shown later (Excerpt Five), a move can be brief or extensive as a phrase, a sentence, or a series of sentences. A number of studies were conducted to investigate the linguistic features commonly used in abstracts, the issue beyond the scope of this paper. In this regard, for more information, please refer to, for instance, Hyland and Tse. (2005) and Salager-Meyer (1992).

The following abstracts are from the other three disciplines of microbiology, civil engineering, and software engineering. They are presented to illustrate variations in terms of the presence and the sequence of moves.

Excerpt Two: Microbiology

Move P	**(S1)** We investigated the clinical characteristics and treatment of patients with a distinctive triad of acute infusion related reactions (AIRRs) to liposomal amphotericin B (L-AMB) via single-center and multicenter
Move R	analyses. // **(S2)** AIRRs occurred alone or in combination within 1 of 3 symptom complexes: (1) chest pain, dyspnea, and hypoxia; (2) severe abdomen, flank, or leg pain; and (3) flushing and urticaria. **(S3)** The frequency of AIRRs in the single-center analysis increased over time. **(S4)** Most AIRRs (86%) occurred within the first 5 min of infusion. **(S5)** All patients experienced rapid resolution of symptoms after intravenous diphenhydramine was administered. **(S6)** The multicenter analysis demonstrated a mean overall frequency of 20% (range, 0%–100%) of
Move D	AIRRs among 64 centers. // **(S7)** A triad of severe AIRRs to L-AMB may occur in some centers; most of these reactions may be effectively managed by diphenhydramine administration and interruption of L-AMB infusion.
Rhetorical structure: P-R-D	

Excerpt Two of microbiology consists of three moves. Move P (Purposive statement) covers Sentence 1, whereas Move R (Result announcement) is quite extensive covering Sentences 2-6. Finally, Move D (Discussion, conclusion, implication) concludes the abstract in Sentence 7.

Excerpt Three: Civil Engineering

Move R	**(S1)** The writers' experiments on local scour at vertical cylinders placed in a sand bed show that similitude of large-scale turbulence is an important consideration influencing equilibrium depth of local scour. **(S2)** For the range of cylinder diameters used in their experiments, the writers identify a direct trend between equilibrium scour depth, normalized with cylinder diameter, and the intensity and frequency of large-scale turbulence shed from each cylinder; values of normalized scour depth increased when
Move D	cylinder diameter decreased. // **(S3)** The writers offer a scour-depth adjustment factor to account for this trend, which essentially is a scale effect incurred with experiments involving three independent length scales: cylinder diameter, bed-particle diameter, and flow depth. **(S4)** The consequent similitude consideration, or scale effect, has general significance for laboratory studies of local scour associated with hydraulic structures in sediment beds.
Rhetorical structure: R-D	

The above abstract from civil engineering consists of only two moves: Move R (Result announcement) in Sentences 1-2, and Move D in Sentences 3-4. Both moves are rather extensive.

Excerpt Four: Software Engineering

Move M	**(S1)** In partition analysis we divide the input domain to form subdomains
Move R	on which the system's behaviour should be uniform. // **(S2)** Boundary value analysis produces test inputs near each subdomain's boundaries to find failures caused by incorrect implementation of the boundaries. **(S3)** However, boundary value analysis can be adversely affected by coincidental correctness-the system produces the expected output, but for
Move D	the wrong reason. // **(S4)** This article shows how boundary value analysis can be adapted in order to reduce the likelihood of coincidental correctness. **(S5)** The main contribution is to cases of automated test data generation in which we cannot rely on the expertise of a tester.
Rhetorical structure: M-R-D	

In the above excerpt of software engineering, the first sentence is classified as Move M (Methodological description). Sentences 2 and 3 represent Move R (Result announcement), and Sentences 4 and 5 for Move D (Discussion, conclusion, implication).

Frequency of moves

It can be observed from the excerpts above that not all of the five moves delineated are always found in a single abstract. Interestingly, some moves are more regular than the others. The results show that among the five moves, Moves R and D tend to co-occur, and based on their

frequencies, they are the most conventional, demonstrating the integral role of these two moves across the four scientific disciplines.

Sequence of moves

The examination of the abstracts shows that the sequence of moves varies a great deal. However, focusing on the two common moves of Moves R (Result announcement) and D (Discussion, conclusion, implication), they not only are prevalent in abstracts but also form a unique sequence of occurrence. That is, if both Moves R (Result announcement) and D (Discussion, conclusion, implication) are used, Move R (Result announcement) usually precedes Move D (Discussion, conclusion, implication). In addition, Move D (Discussion, conclusion, implication), if present, usually concludes an abstract.

Even though the relationship between, and the sequence of, Moves R (Result announcement) and D (Discussion, conclusion, implication) is quite established, those of the other three moves (Moves B, P and M) remained to be elucidated. To address this inquiry, further examination of the following two abstracts in Excerpts Five and Six from biochemistry and microbiology can be quite revealing.

Excerpt Five: Biochemistry

Move B	**(S1)** The Ras GTPase links extracellular mitogens to intracellular
Move P	mechanisms that control cell proliferation. // **(S2)** To understand how Ras
Move M	regulates proliferation in vivo, // we activated or inactivated Ras in cell
Move R	clones in the developing *Drosophila* wing. // **(S3)** Cells lacking Ras were smaller, had reduced growth rates, accumulated in G1, and underwent apoptosis due to cell competition. **(S4)** Conversely, activation of Ras increased cell size and growth rates and promoted G1/S transitions. **(S5)** Ras unregulated the growth driver dMYc, and both Ras and dMyc
Move D	increased levels of cylin E posttranscriptionally. // **(S6)** We propose that Ras primarily promotes growth and that growth is coupled to G1/S progression via cyclin E.

Rhetorical structure: B-P-M-R-D

Excerpt Six: Microbiology

Move B	**(S1)** Cationic antimicrobial peptides are believed to exert their primary activities on anionic bacterial cell membranes; however, this model does not adequately account for several important structure-activity relationships. **(S2)** These relationships are likely to be influenced by the
Move P	bacterial response to peptide challenge. // **(S3)** In order to characterize the
Move M	genomic aspect of this response, // transcription profiles were examined for *Escherichia coli* isolates treated with sublethal and lethal concentrations of
Move R	the cationic antimicrobial peptide cecropin A. // **(S4)** Transcript levels for 26 genes changed significantly following treatment with sublethal peptide concentrations, and half of the transcripts corresponded to protein products with unknown function. **(S5)**The pattern of response is distinct from that following treatment with lethal concentrations and is also distinct from the bacterial response to nutritional, thermal, osmotic, or oxidative stress. //
Move D	**(S6)** These results demonstrate that cecropin A induces a genomic response in *E. coli* apart from any lethal effects on the membrane and suggest that a complete understanding of its mechanism of action may require a detailed examination of this response.

Rhetorical structure: B-P-M-R-D

As demarcated, the two abstracts above consist of the entire set of five moves. The analysis seems to suggest that the sequence of Moves B –P –M – R – D is highly possible. The emergence of this pattern indicates how each of the five moves interacts with one another in a multi-faceted manner in this genre of abstracts.

In summary, this paper has shown that this academic genre can be captured in terms of rhetorical moves. Even though the manifestations of moves identified by genre analysis can be diverse, certain generalizations can be made with regard to their rhetorical organization in terms of their frequency and typical sequence.

DISCUSSION

This study highlights the role of genre analysis in analyzing scientific abstracts in several academic disciplines. The analysis sheds light onto the structural components or moves making up the abstract texts. However, in light of the exploratory nature of this study and the specific focus on research article abstracts (as opposed to conference presentation abstract, for instance), caveats are in order, and results should be interpreted with caution. First, given the limited number of texts analyzed, the findings of this study remain to be substantiated by subsequent studies analyzing a larger abstract corpus. Next, it is possible that the moves identified might be influenced by the length (number of words) of the abstracts. Had the length of words been controlled, the results might have been different. Finally, as a word of caution, genre analysis represents an attempt to describe,

not to prescribe, how texts are constructed. Therefore, variations are highly expected, as demonstrated in this study.

The genre analysis applied to abstracts of multiple academic disciplines elucidates the textual structural patterns conventionally followed by scientists in their respective disciplines. This body of knowledge, which is likely to be intuitive for native speakers of English but not overt to non-native speakers of English, allows particularly novice scientists or scholars to better understand how scientific abstracts are constructed. The study also sheds light onto the significance of the awareness of conventional patterns of organization preferred in this genre and in these disciplines. Keeping this in mind, these novice article writers will be able to successfully and internationally disseminate and exchange their research work. Finally, the application of discourse analysis in analyzing scientific abstracts elucidates the contribution of the field of applied linguistics to other academic disciplines.

Acknowledgement

The compilation of the four corpora was possible through the following funding sources:

- the biochemistry corpus from the Doctoral Dissertation Improvement Grant # BCS 0213948, the National Science Foundation (NSF) and TOEFL Small Grant for Doctoral Research in Second and Foreign Assessment Program from the Education Testing Service (ETS), USA.

- the microbiology corpus from the Thailand Research Fund under Grant No. MRG4780210
- the civil engineering corpus and the software engineering corpus from the Thailand Research Fund under Grant No. RSA5080005

References

Hyland, K. & Tse, P. (2005). Hooking the Reader: A Corpus Study of Evaluative *that* in Abstracts. *English for Specific Purposes, 24,* 2, 123-139.

Kanoksilapatham, B. (2003). A Corpus-based Investigation of Biochemistry Research Articles: Linking Move Analysis with Multidimensional Analysis. Unpublished Ph.D. Dissertation. Georgetown University.

_____ . (2005). Rhetorical Structure of Biochemistry Research Articles *English for Specific Purposes, 24,* 3, 269-292.

_____ . (2007). Rhetorical Moves in Biochemistry Research Articles. In D. Biber, U. Connor & T. Upton (eds.), *Discourse on the Move* (pp. 73-120). Amsterdam: John Benjamins.

Bhatia, V. K. (1993). *Analysing Genre: Language Use in Professional Settings.* London & New York: Longman.

Martín, P. M. (2003). A Genre Analysis of English and Spanish Research Paper Abstracts in Experimental Social Sciences. *English for Specific Purposes, 22,* 1, 25-43.

Melander, B., Swales, J. M., & Fredrickson, K. M. (1997). Journal Abstracts from Three Academic Fields in the United States and Sweden: National or Disciplinary Proclivities? In A. Duszak (Ed.), *Intellectual Styles and Cross-cultural Communication* (pp. 251–272). Berlin: Mouton De Gruyter.

Ozturk, I. (2007). The Textual Organization of Research Article Introductions in Applied Linguistics: Variability within a Single Discipline. *English for Specific Purposes, 26,* 1, 25-38.

Pho, P. D. (2008). Research Article Abstracts in Applied Linguistics and Educational Technology: A Study of Linguistic Realizations of a Study of Rhetorical Structure and Authorial Stance. *Discourse Studies, 10,* 2, 231-250.

Samraj, B. (2005). An Exploration of a Genre set: Research Article Abstracts and Introductions in Two Disciplines. *English for Specific Purposes, 24,* 2, 141-156.

Salager-Meyer, F. (1992). A Text-type and Move Analysis Study of Verb Tense and Modality Distribution in Medical English Abstracts. *English for Specific Purposes, 11,* 2, 93-113.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings.* Cambridge: Cambridge University Press.

_____. (2004). *Research Genres: Explorations and Applications.* New York: Cambridge University Press.

van Mulken, M. & van der Meer, W. (2005). Are you Being Served?: A Genre Analysis of American and Dutch Company Replies to Customer Inquiries. *English for Specific Purposes, 24,* 1, 93-109.