TECHNOLOGY AND CORPUS-BASED FRAMEWORK FOR LANGUAGE EDUCATION IN HIGHER EDUCATION LEVEL

Risa Rumentha Simanjuntak

English Department
Faculty of Humanities
Bina Nusantara University
Jl. Kemanggisan Ilir III No. 45, Palmerah, Jakarta 11480, Indonesia
risarsimanjuntak@binus.edu

ABSTRACT. *Technology-enhanced learning has been widely recognized in the higher education environment, especially to improve the quality of writing for research and publications. This present study investigated the contributions of word-generating technology in facilitating writers with the discipline-specific words, expressions, and common structure for writing good and engaging research abstracts. Data for this study were a corpus of 956 undergraduate students' English abstracts in Computer Science. AntConc. 3.5.8 (2019) was used as the word-generating tool to identify the words and expressions in Computer Science. The result of this study was the model of abstract as an important written genre in Computer Science, including the use of specific words to present objectivity and epistemic truth. Besides, the framework also provided information on the moves or sections in the abstracts, administering the importance of frame of thinking and corpus in learning. Implications from this research included the use of the model for self-learning and further development in technology use for language education in higher education level.*
Keywords: Technology, Writing, Computer Science, Abstract

1. **Introduction.** Composition pedagogy to date has considered writing from the perspectives of product and process. The product-process dichotomy was a result of the development of composition pedagogy mainly in the United States. The product-oriented approach in writing started in the 1950s. During that time colleges in the U.S. were preparing international students to read and write successfully in university courses. These students were trained to write following the universities' guidelines, thus producing texts in the same manner as the provided models. The compliance with the guidelines and the rudimentary rules of grammar became the key to success in writing. In the 1960s, the product-oriented pedagogy found some challenges in that students were making errors in writing due to the writing styles they adopted from their first language. Differences in writing could not be regarded as incompetence but rather as a portrayal of students' cultural backgrounds. Later development in composition pedagogy was aimed at synergizing both product and process approaches.

Genre approach in writing provides both relevant and dynamic approach to composition pedagogy. Genre approach considers writing as a social construction [1] rather than the end product. Social construction in genre refers to the conventions made by the communities [2]. Such communities include publication journals, university cultures, as well as discipline-based rhetoric [3,4]. Recurrent forms as the specific construction appearing in writing are considered as evidence of these conventions. Investigating writing from the genre approach provides both comprehensive and dynamic projections of real and contextual communications.

Studies have identified some variations in the genre of abstracts. These variations included the absence of one or some rhetorical moves and different use of linguistic devices in presenting arguments and persuasion. The different use of linguistic devices, such as the limited use of stance markers, was affecting the epistemic quality of research papers as objective inquiries. The differences also revealed preferences and strategies used by students as writers in presenting their knowledge and opinions. It was concluded that the variations in abstracts revealed the dynamics in the genre [5]. Abstracts as a genre conveyed the conventions made in the scientific communities while allowing some variations as evidence of the dynamics in the community [5].

Studies in research abstracts have also shown the varieties as affected by different socio-cultural backgrounds of the writers [6]. The socio-cultural backgrounds contributed to the lexico-grammatical construction and lexical selections [7]. Abstracts written in Spanish, for example, used more emotive and personal traits compared to the English translation [8]. Abstracts written in Finish were using more genitive construction/forms (x of y) [9]. Meanwhile, Japanese abstracts were identified to use passive and impersonal style [10]. Unfortunately, there were limited studies conducted on the undergraduate level, which is fundamental in higher education.

2. Literature Review.

2.1. **Genre-based teaching of English.** Previous studies in abstracts have investigated the macrostructure and microstructure aspects of abstracts as a genre. Studies in the macrostructure of abstracts identified specific structures or constructions for an abstract, consisting of rhetorical moves and sequences [3]. Rhetorical moves are the building blocks of the genre, consisting of phrases, clauses, or sentences delivering the communicative purposes [3]. The recurrence of rhetorical moves in an abstract identifies the regularities and common construction. Studies in the microstructure of abstracts included the use of linguistic devices to realize the communicative purposes [11]. Findings from these studies suggested specific linguistic devices, such as stance and engagement markers [12] to be used for convincing and engaging abstracts. Stance consisted of hedges, boosters, and attitude markers, which are important in presenting epistemic truth and objectivity in research [13,14]. Engagement, such as self-mentions and engagement markers, increases the persuasiveness of the research.

2.2. **Research abstract.** An abstract has been identified as a short text containing all key information in the full research report/paper concisely [15]. In an undergraduate thesis, an abstract is easily identifiable as it appears at the beginning of the thesis and is written in a single paragraph in a justified/block format. A written genre is a specific written text, which directly taps-in to the needs and expectations of the readers. A research abstract as a genre requires important information to be written concisely and orderly to feed the needs of the readers of the important messages in a research paper. A good abstract is aimed at influencing or persuading readers to further read the research paper. In other words, a good abstract would eventually inform readers of the quality of the research and not merely presents summaries of the research [16]. The persuasive quality becomes the main feature and strategic function of an abstract [17]. An abstract is probably the only text people read nowadays because of the large amount of research being made available through the Internet [16,17].

Considering the important and strategic function of an abstract for a research paper, some references provided specific guidelines in writing a good abstract. ANSI (American National Standards Institute) produced a 30-pages guideline entitled *Guidelines for Abstracts* [18]. According to ANSI, an abstract is "a brief and objective representation of a document or an oral presentation" (p.1). Another guideline is provided by IEE, which requires writers to write strong abstract to attract more readers. Besides, it suggests

writers produce, "a concise but comprehensive abstract that communicates the content of the full article". Different organizations, journals, or universities may have different guidelines for writing abstracts. In general, an abstract contains 100-250 words written in a single paragraph. However, there may be different formats and word-lengths following the guidelines. Regardless of the variations, an abstract is considered an important window on the value of the research.

Thesis abstracts as a genre in academic discourse have been studied since the 1960s. Early studies of abstracts have been conducted in an individual discipline, such as applied linguistics [19], engineering [20], education [21], and technology [22]. In the 2000s, comparative studies have been conducted to further investigate the common elements in abstracts. One prominent study was conducted across 16 disciplines, using a corpus of 236 abstracts of published articles from reputable journals [23]. The result of this study was the rhetorical move model of I-P-M-Pr-C (Introduction-Purpose-Method-Product-Conclusion) to be used in all 16 disciplines. It was further assumed that this model could be applied to all abstracts in general.

This present study aims at providing a model for language education in undergraduate level by way of scrutinizing the macrostructure and microstructure for genre-based education. The goal of such investigation would be a technological and corpus-based model for language education in higher education in general.

3. **Research Methods.**

3.1. **Corpus.** The corpus for this present study is 956 undergraduate thesis abstracts in Computer Science. The abstracts were obtained through Bina Nusantara University Library's website. All abstracts were transferred from .pdf format to .txt files for data collection.

3.2. **Data collection methods.** Data for rhetorical moves were collected manually using Hyland's model of rhetorical moves. Data for linguistic devices were collected using AntConc. 3.5.8 (2019) [24] using Hyland's list of metadiscourse (hedges, boosters, attitude markers, self-mentions, and engagement markers) as the source of information for setting the Word List range (Figure 1).
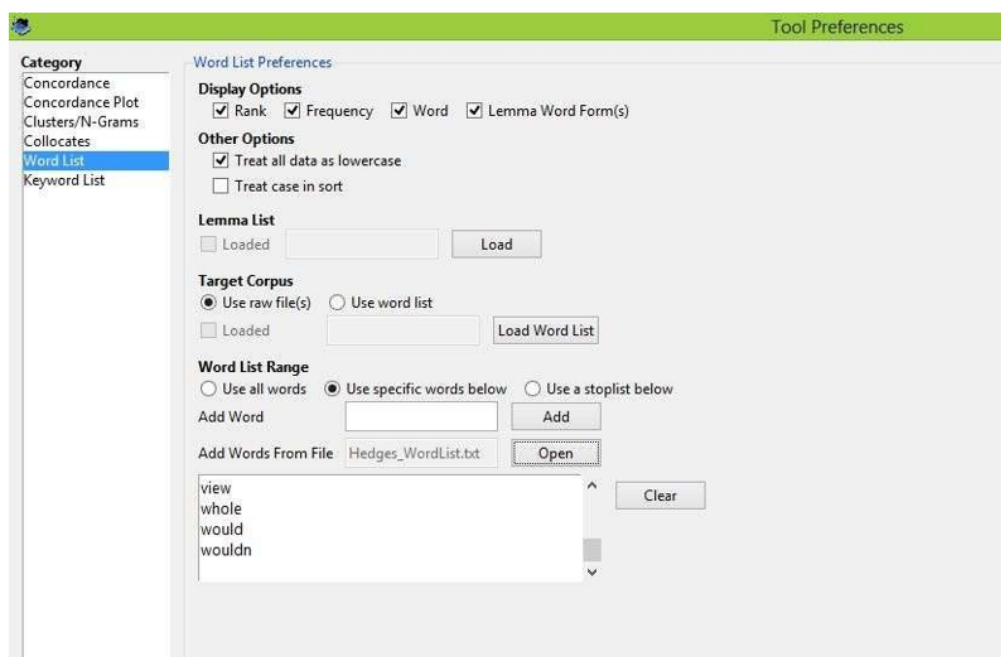

FIGURE 1. The use of a list of hedges to identify key messages in the texts

Generating word list was done by inserting all 956 files of the corpus into AntConc. 3.5.8 (2019), which is an open ware for generating words from the corpus, and limited the word-search inside the range provided by a set of the word list in the linguistic device. Sample for word list generated using hedges as the word list range preference can be seen in Figure 2. Data collected was then tabulated for further analysis.
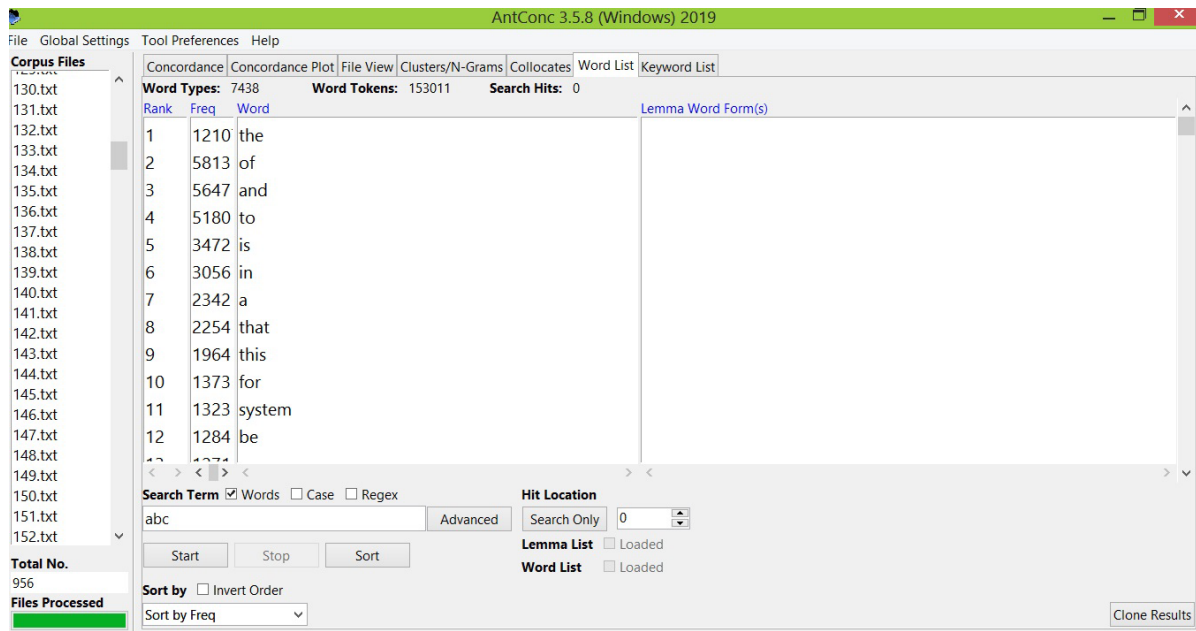

FIGURE 2. A list of words generated from the corpus

3.3. **Data analysis methods.** Rhetorical moves were analyzed using three categories of rhetorical construction: Obligatory [20], Dominant [21], and Optional [21] moves. Obligatory moves become the compulsory elements in a genre because they appear 100% in all abstracts [20]. Dominant moves appear $> 90\%$ [21] and Optional moves appear $< 60\%$ [21]. The organization of the rhetorical move is analyzed to come up with an integrated model of the abstract's macrostructure. The microstructure of abstract is analyzed using Hyland's list of metadiscourse to present the model for linguistic devices in an abstract.

4. **Proposed Framework.** The proposed framework for technological and corpus-based model for language education can be seen in Figure 3 as follows.
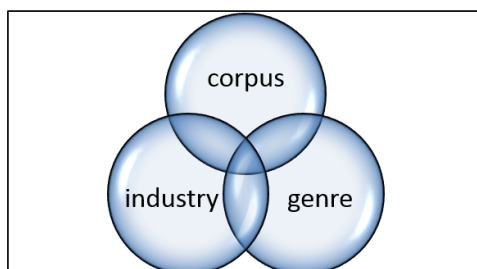

FIGURE 3. Model of language education in higher education

4.1. **Corpus-based.** The corpus-based linguistic framework provides lexical selections as shown in the corpus-based freeware AntConc. 3.5.8 (2019). From the generated data, words were analyzed by identifying the particular word in the context that delivered the intended semantic meaning, using the Concordance lines as seen in Figure 4.
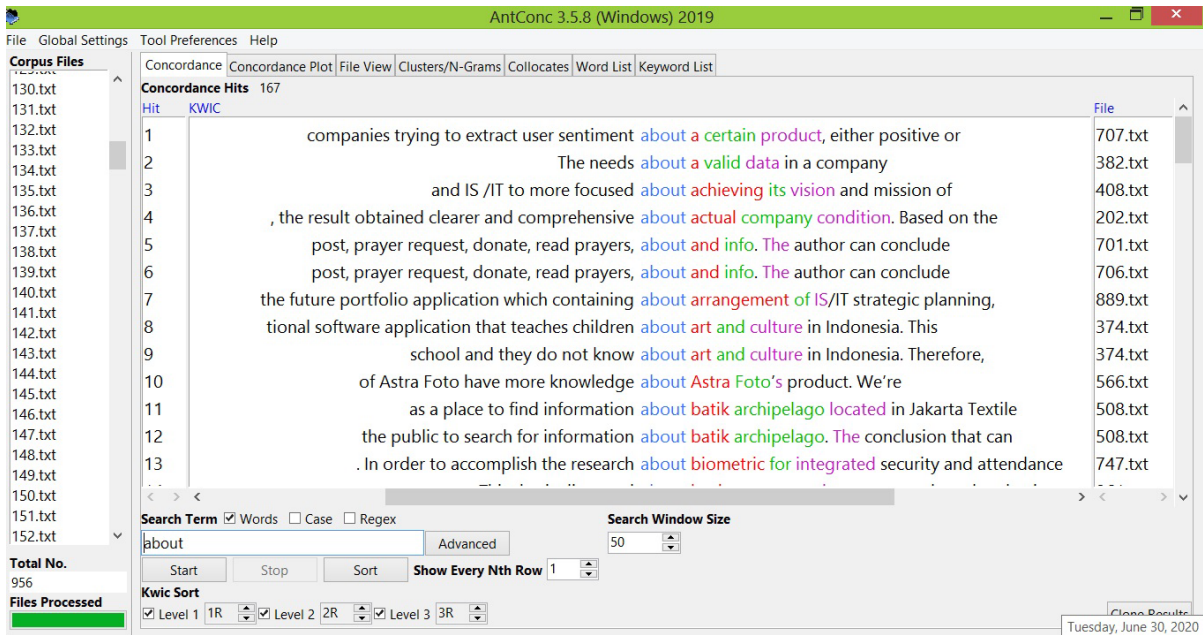
FIGURE 4. Example of the use of the word "about" as appearing in the corpus

For example, the hedge "about" was used to reduce the certainty of a claim. The use of the word "about" is provided in some examples of the texts, resulting in alternatives for more original and authentic sentences to use. When learning these examples, students will gain more insights on what word to use and how such words behave in sentence environments. Also, the corpus can be used to present the use of the particular lexical selection in the real text as can be seen in the example (Figure 5).
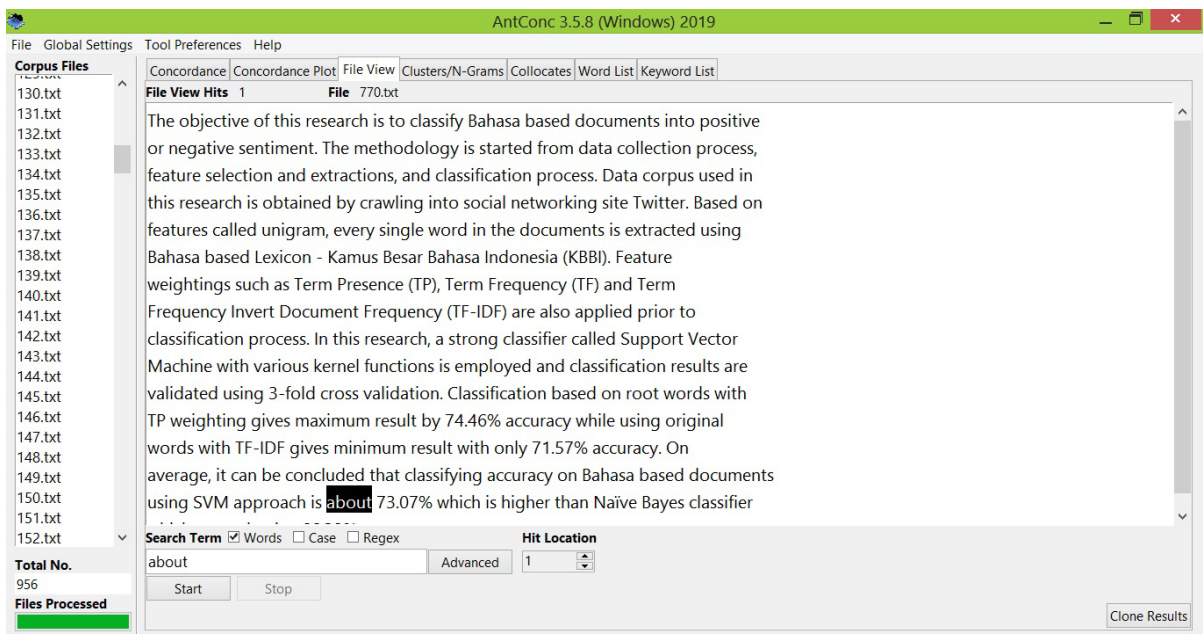

FIGURE 5. Example of the use of "about" in full text

In effect, the use of a particular word can be understood from its definitive meaning and in strategic and contextual meaning. Such comprehensive information is made possible by the use of corpus-generated technology using freeware such as AntConc. 3.5.8 (2019).

4.2. **Genre-based.** From the corpus, it can be identified that a good abstract requires certain sections or rhetorical moves as the building blocks of a genre. The data show the moves as shown in Table 1.

TABLE 1. Rhetorical moves in a thesis abstract

Rhetorical moves	Frequency		Category
	n ($N = 956$)	%	
I (Introduction)	918	96	Dominant
P (Purpose)	956	100	Obligatory
M (Method)	899	94	Obligatory
Pr (Product)	851	89	Dominant
C (Conclusion)	526	55	Optional

Based on the findings, it appears that all rhetorical moves can be categorized following Hyland's rhetorical move model of I-P-M-Pr-C. All phrases, clauses, or sentences could be categorized into the five rhetorical moves. In this present study, the Obligatory move was P (Purpose) rhetorical moves. Data showed all undergraduate thesis abstracts mentioned the purpose of the research. It can be concluded that undergraduate thesis abstracts always use the purpose of research as its rhetorical move. Abstracts also used I (Introduction), M (Method), and Pr (Product) rhetorical moves. Although there were big percentages of use for each rhetorical move, not all abstracts used all three of these moves in presenting their research. The three rhetorical moves are considered as Dominant moves because none of these rhetorical moves always occurred or were used 100% in the data population. The C (Conclusion) rhetorical move appeared to be used in some abstracts. This rhetorical move becomes the Optional move because only a limited number of abstracts used this rhetorical move.

From the findings, it can be concluded that an undergraduate abstract always mentions the purpose of the research. It is also common that introductions to the topic/issue, the method, as well as the result of the research, are mentioned in an abstract. An abstract could draw further implications based on the results of the abstract. However, concluding is not considered as compulsory. Findings in this present study are not in-line with the findings in previous studies of journal article abstracts and post-graduate thesis abstracts [25]. In previous studies, the obligatory rhetorical moves were I (Introduction) and P (Purpose) [25]. One explanation for the difference is because students may consider I (Introduction) and P (Purpose) rhetorical moves as similar and that the most important information to mention is the purpose of the research. The findings also confirm that undergraduate thesis abstracts are different from post-graduate thesis abstracts or journal article abstracts.

4.3. **Industry-oriented.** A pedagogical model of the teaching using a technology-enhanced framework would have to be more exposure to authentic use of certain lexical structures. Moreover, the information is enriched with the use of language in a real and professional context. In other words, industrial oriented genre-based composition pedagogy focuses on relevant and contextual training for English. Applying a genre-based approach in writing classes in universities will prepare students for the real and up-to-date language used in the discipline. By using abstracting activity students are encouraged to think about the bigger picture, or the key information, of the research inquiries. Abstracting as key information literacy skills in the 21st century [26] is aimed at equipping students to communicate using real and effective English.

The industry-oriented framework in this study uses instances of real English in the industries [27] as provided in the corpus. Corpus enables learners to acquire instances from professional or industrial contexts of the common words, expressions, and constructions

used in a particular discipline. A corpus in a particular field, such as Computer Science, could be acquired using freeware word generators, such as AntConc. The generated corpus provides information on the word order, sentence structures, and common expressions used as models for improving writing qualities. Using this framework, in application students can autonomously apply the knowledge in their writing for composing drafts for their research abstracts [28]. Technology-enhanced English learning in higher education collaboratively, students can compose their abstracts together with their peer researchers while negotiating the words selected and the rhetorical moves used. Students can also edit or review their research abstracts and act as peer-review using the model provided in this present study as a point of reference.

5. **Conclusions.** This present study has shown how technology-enhanced framework is mostly suitable for English teaching in higher education level. The undergraduate thesis abstracts as a model for specific language are used in professional and industrial contexts. At the undergraduate level, the teaching of thesis abstracts can be directly applied by identifying the use of rhetorical moves. The technology-enhanced framework enables a model of English used to understand English from the real communication situations in their discipline. In conclusion, an ideal framework for English education at a higher level needs to incorporate the three axial corpus-genre-industry orientated English education framework. Future research can be conducted by doing experiments using a particular set of words to enable students to express stance and engagement in more dynamic and creative ways.

REFERENCES

[1] C. Bazerman, Genre and cognitive development: Beyond writing to learn, in *Genre in a Changing World*, C. Bazerman, A. Bonini and D. Figueiredo (eds.), New York, Parlor Press, 2009.

[2] C. Bazerman et al., *Traditions of Writing Research*, Routledge, New York, 2010.

[3] J. M. Swales, *Genre Analysis: English in Academic and Research Settings*, Cambridge University Press, Cambridge, 1990.

[4] J. M. Swales and C. B. Feak, *Abstracts and the Writing of Abstracts*, The University of Michigan Press, MI, 2009.

[5] K. Hyland and P. Tse, Metadiscourse in academic writing: A reappraisal, *Applied Linguistics*, vol.25, pp.156-177, 2004.

[6] U. Connor, Mapping multidimensional aspects of research: Reaching to intercultural rhetoric, in *Contrastive Rhetoric Reaching to Intercultural Rhetoric*, U. Connor, E. Nagelhout and W. V. Rozycki (eds.), Amsterdam, John Benjamins, 2008.

[7] S. Arsyad, The discourse structure and linguistic features of research article abstract in English by Indonesian academics, *The Asian ESP Journal*, vol.10, no.2, pp.191-224, 2014.

[8] L. S. Sanz, Lost (and gained) in translation: A contrastive (English/Spanish) analysis of rhetorical and lexicogrammatical patterns in sociology research article abstracts, in *Abstracts in Academic Discourse: Variation and Change*, M. Bondi and L. S. Sanz (eds.), Bern, Peter Lang AG, 2014.

[9] B. Samraj, An exploration of a genre set: Research article abstracts and introductions in two disciplines, *English for Specific Purposes*, vol.24, pp.141-156, 2005.

[10] M. T. Apple, The vocabulary and style of Engineering research abstract writing, *OnCUE Journal*, vol.7, no.2, pp.86-102, 2014.

[11] K. Hyland, *Disciplinary Discourses: Social Interactions in Academic Writing*, Longman, London, 2000.

[12] K. Hyland, *Metadiscourse*, Continuum, London, 2005.

[13] K. Hyland, *Hedging in Scientific Research Articles*, John Benjamins, Amsterdam, 1998.

[14] S. Arsyad, B. K. Purwo, K. E. Sukamto and Z. Adnan, Factors hindering Indonesian lecturers from publishing articles in reputable international journals, *Journal on English as a Foreign Language*, vol.9, no.1, pp.42-70, 2019.

[15] V. Bhatia, *Analysing Genre: Language Use in Professional Settings*, Longman, London, 1993.

[16] R. O'Connor, *Writing Scientific Research Articles: Strategy and Steps*, Wiley-Blackwell, West Sussex, 2009.

[17] P. Gillaerts, Shifting metadiscourse: Looking for diachrony in the abstract genre, in *Abstracts in Academic Discourse: Variation and Change*, M. Bondi and L. S. Sanz (eds.), Bern, Peter Lang AG, 2014.

[18] American National Standards Institute, *Guidelines for Abstracts*, NISO Press, MD, 1996.

[19] S. Can, E. Karabacak and J. Qin, Structure of moves in applied linguistics abstracts, *Publications*, vol.4, no.3, pp.1-16, 2016.

[20] B. Kanoksilapatham, Generic characterisation of civil engineering research article abstracts, *3L: The Southeast Asian Journal of English Language Studies*, vol.19, no.3, pp.1-10, 2013.

[21] P. D. Pho, Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance, *Discourse Studies*, vol.10, no.2, pp.231-250, 2008.

[22] P. Supatranont, Developing a writing template of research article abstracts: A corpus-based method, *Procedia – Social and Behavioral Sciences*, vol.66, pp.144-156, 2012.

[23] M. B. D. Santos, The textual organization of research paper abstracts in applied linguistics, *Text*, vol.16, no.4, pp.481-499, 1996.

[24] A. Lawrence, *AntConc Version 3.5*, 2016.

[25] D. Khansari, C. S. Heng, C. M. Yuit and H. Tan, Regularities and irregularities in rhetorical move structure of linguistics abstracts in research articles, *3L: The Southeast Asian Journal of English Language Studies*, vol.22, no.1, pp.39-54, 2016.

[26] T. Koltay, *Abstracts and Abstracting: A Genre and Set of Skills for the Twenty-First Century*, Chandos, Oxford, 2010.

[27] X. Zhang and Z. Wang, Approaches to multiple attribute group decision making with multi-granularity linguistic term sets, *International Journal of Innovative Computing, Information and Control*, vol.12, no.1, pp.275-291, 2016.

[28] L. Flowerdew, Corpus-based research and pedagogy in EAP: From lexis to genre, *Language Teaching*, vol.26, pp.1-18, 2013.