



Lexicogrammatical Features in Introduction Sections of Research Articles Published in the International Quality Journals

Faizatul Istiqomah¹, Binti Maunah²

^{1,2}Universitas Islam Negeri Sayyid Ali Rahmatullah Tulungagung

Corresponding email: faizatulistiqomah79@gmail.com

ABSTRACT

To participate in global academia, expert and novice researchers are strongly encouraged to publish their academic papers in international journals. Hence, a remarkable ability to produce quality academic writing, which includes knowledge of linguistics features, is a necessity. This paper aimed at exploring lexicogrammatical properties focusing on grammatical intricacy and lexical density deployed within the introduction sections of research articles (RA) published in four international journals in the Asian region. The corpus of 40 introduction sections of RAs was analyzed based on the Systemic Functional Linguistics framework. The results of this study highlighted the structural complexity indicated by the use of paratactic and hypotactic clauses which led to lexically dense academic writing. Some pedagogical implications and recommendations related to the findings were discussed further in this article's final part.

Keywords: Grammatical intricacy; introduction sections; lexical density; research articles

First Receive:

08 November 2022

Revised:

15 January 2023

Accepted:

10 March 2023

Final Proof Received:

01 May 2023

Published:

01 June 2023

How to cite (in APA style):

Istiqomah, F. & Maunah, B. (2022). Lexicogrammatical Features in Introduction Sections of Research Articles Published in the International Quality Journals. *ELITE Journal*, 5 (2), 409-416.

INTRODUCTION

There has been innumerable research on the academic writing genre, including studies on articles published in reputable national or international journals. In the Indonesian context, being able to succeed in publishing academic papers has even recently been compulsory for some graduate students to be eligible for the degree and faculty members to upgrade their careers. However, many researchers, particularly the novices, are still struggling for publication and experienced rejections from journals for various reasons such as linguistic matters. The linguistic problems could be in the failure on presenting ideas in effective ways or conventionally accepted written discourse. Halliday and Matthiessen (2004) suggested that written language tends to have high lexical density where more lexical words are used (Eggins, 2004), while spoken language has more grammatical intricacy or complex sentences with more combinations of various types or clauses within the utterances.

Studies on lexical density and grammatical complexity or intricacy have investigated spoken and written languages. One example of oral language is the current study by Amelia et al.

(2020) exploring both lexical density and grammatical intricacy within the language expressed during a debate contest. They found that debaters' speeches contained a high level of density and grammatical complexity as they chose more lexical words and complex clauses in presenting their arguments. The studies in written discourse, which are found to be more in numbers than the spoken one, include the research on students' academic writing (Ma'mun, 2017; Misuzawa, 2020), students' thesis proposals (Nguyen & Edwards, 2015; Syarif, 2018) students' thesis abstracts (Hanafiah & Yusuf, 2016), national examination texts (Pujiati & Pramudyawardhani, 2021), and English textbooks for tertiary students in Vietnam (To, 2017). From these previous studies, it is found that very scarce investigation is on research articles (RAs), especially in the introduction section, from international and reputable journals. Therefore, this present study is devoted to filling this void since the information on lexical density and grammatical intricacy employed by writers who successfully published their papers in quality journals could help students and beginner researchers to learn and reflect on their writings.

This research belongs to a content analysis study based on the Systemic functional linguistics (SFL) approach. SFL considers language as a social semiotic system used by humans to communicate their intentions. In the SFL study, language is seen as more than groups of words arranged grammatically correctly, but also as a complex structure that is divided into several levels or ranks called rank scale. Butt et al. (2003, p. 29) claim that the understanding on "how a system as intricate as human language works" should be obtained through the concept of a rank scale of language which contains clause complex at the highest level, followed lower levels including, in the sequence order, clause, group or phrase, word, and morpheme. From this rank, the relationship of each clause within a clause complex, traditionally known as a sentence, can be identified using the system of clause complexing which contains the degree of interdependency or TAXIS and the LOGICO SEMANTIC relation with their sub-system as shown in figure 1 below (Halliday & Matthiessen, 2014, p.438).

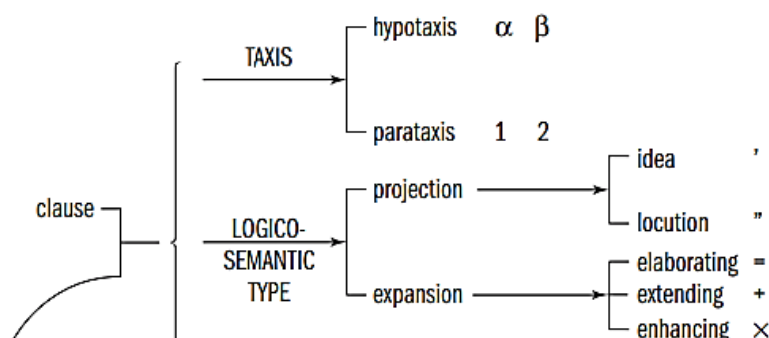


Figure 1 The system of clause complexing (adapted from Halliday & Matthiessen, 2014, p.438)

The focus of this study is on clause interdependency or TAXI which is classified into two: hypotaxis (symbolized with α , β , ...) and parataxis (symbolized with 1, 2, ...). Hypotaxis is an unequal status relation of the dependent and independent (dominant) clauses, whereas parataxis is an equal status relation between independent clauses in which one of them is "initiating and the other continuing" (Halliday & Matthiessen, 2004). Look at the examples of two clause complexes below for more explanation and the meanings of the symbols within the clauses can be seen at the notational conventions table which follows:

- (a) ||| (α) ...the development of benchmark assessments are of little value || (β) if they do not take the teacher into account. ||| (extracted from Text 023I)
- (b) |||...(1) studies [[examining L2 use]] reported a strong effect of L1 frequency, || (2) and those [[adopting paired-associate learning format]] obtained a small effect. ||| (extracted from Text 007A)

The first example (a) is a hypotaxis clause with the first clause as the dominant/dependent clause (α) and the second as the dependent one (β). The second example is a parataxis clause with two independent clauses connected by the coordinating conjunction *and*.

Table 1 SFL notational conventions and meaning (adapted from To, 2017)

SFL notational conventions	Meanings
	clause complex boundary
	ranking clause boundary
[[]]	embedded clause boundary
$\alpha, \beta, \gamma, \dots$	hypotactic structure
1, 2, 3, ...	paratactic structure

In addition to the discussion on clauses relationship, this present study also concerns the level of complexity of the clauses within the texts being analyzed in terms of their grammatical intricacy. Eggins' (2004 p. 97) calculation of the grammatical intricacy of a text is by dividing the number of sentences in the text by the number of clauses per sentence. The explanation of the calculation will be discussed further in the following method session.

METHOD

This study is qualitative in nature by employing content analysis of corpus-assisted research on the introduction sections of research articles (RAs) from four international reputable journals in Asian regions. The introduction sections were taken from 40 RAs with a corpus of 26068 tokens in total. The articles were selected randomly from the latest volumes of the selected journals, while the journals were chosen based on the data of journal ranking of Scimago and Scopus (accessed in August 2022). The data from the journals are presented in the following table.

Table 2 The Selected International Journals

No.	Journal names	SJR	Country
1	Journal of Asia TEFL	Q1/ 0.375	South Korea
2	JALT CALL Journal	Q1/ 0.346	Japan
3	Indonesian Journal of Applied Linguistics	Q2/ 0.297	Indonesia
4	PASAA	Q2/ 0.289	Thailand

The 40 introduction sections of RAs were chosen from four journals with 10 RAs from each journal. The texts of the introduction sections were coded as described in Table 3 below named with the journal name initials: Q1 A (n=10), Q1 J (n=1), Q2 I (n=1), and Q2 P (n=1).

Table 3 Text codes

No.	Journal Q1 A	Journal Q1 J	Journal Q2 I	Journal Q2 P
-----	--------------	--------------	--------------	--------------

Available online at <https://www.elitejournal.org/index.php/ELITE>

1.	001A	011J	021I	031P
2.	002A	012J	022I	032P
3.	003A	013J	023I	033P
4.	004A	014J	024I	034P
5.	005A	015J	025I	035P
6.	006A	016J	026I	036P
7.	007A	017J	027I	037P
8.	008A	018J	028I	038P
9.	009A	019J	029I	039P
10.	010A	020J	030I	040P

The analysis of these data was divided into two parts with two different tools of analysis: first is lexical density (LD) analysis and second is Grammatical Intricacy (GI) analysis. The LD analysis was done for all data of 40 texts by using ready-made software called ADELE X Analyser (ADA), a free tool analyzer developed by Granada University to count LD (Nguyen & Edwards, 2015). The link to this tool is provided in the appendix. Later, from the results of LD, the researcher selected the texts with the highest level of LD to be analyzed further on their GI. Thus, one text from each journal, with the highest LD compared to the data from the same journal, was chosen. So, four texts were chosen to be analyzed in detail to identify the employment of GI written in the texts. The second analysis was not applied to all texts due to time constraints since the analysis was done manually. In this analysis, each clause within each clause complex (sentence) was analyzed to see the clause interdependency such as hypotactic and paratactic structures. The number of clauses in each chosen text, then, is divided by the number of clause complexes or sentences. The appearance of embedded clauses also became slightly part of the exploration.

FINDINGS AND DISCUSSION

The Lexical Density Analysis

After gathering the data of the introduction sections of 40 RAs, each text was then copied into Microsoft word documents and pasted into the tool to identify the Lexical Density percentage. The LD shown in the analysis is the ratio between tokens and types of words written in the texts. The results from ADA are presented in Table 4 below. The table includes the numbers of tokens or words, word types, and LD.

Table 4 The Results of Lexical Density (LD) Analysis from ADA software

Text	Tokens	Types	LD	Text	Tokens	Types	LD
001A	1072	358	33.40	021I	788	364	46.19
002A	866	376	43.42	022I	1079	399	36.98
003A	732	319	43.58	023I	614	303	49.35
004A	632	266	42.09	024I	768	333	43.36
005A	403	179	44.42	025I	910	358	39.34
006A	465	222	47.74	026I	1426	520	36.47
007A	570	317	55.61	027I	864	409	47.34
008A	416	218	52.40	028I	703	304	43.24
009A	255	128	50.20	029I	714	337	47.20
010A	835	319	38.20	030I	1315	514	39.09
011J	700	324	46.29	031P	417	202	48.44

012J	362	191	52.76	032P	726	287	39.53
013J	343	181	52.77	033P	645	274	42.48
014J	532	245	46.05	034P	671	321	47.84
015J	275	154	56.00	035P	637	289	45.37
016J	550	239	43.45	036P	718	293	40.81
017J	316	185	58.54	037P	468	234	50.00
018J	289	173	59.86	038P	426	213	50.00
019J	373	192	51.47	039P	615	286	46.50
020J	549	269	49.00	040P	673	290	43.09

From the findings, it is apparent that the LD of the introduction sections of RAs is between 33.40% to 59.86%. The highest and lowest LD texts are highlighted in the table above. The highest are highlighted darker than the lowest LD. Thus, the texts with the highest LD from the Q1 journals are Text 007A with 55.61% LD and Text 018J with 59.86% of LD; while from the Q2 journals are Text 023I with 49.35% LD, and Texts 037P and 038P with 50% LD. In the case of similar LD between two texts of the last journal, the one selected for GI level was Text 037P since this text had more tokens and types. So, it is believed that this text had more lexical diversity within the text.

The Results of Grammatical Intricacy Analysis

The second analysis aims to calculate the GI level from the four texts with high levels of LD. The results of the manual analysis are presented in Table 5 consisting of the total number of clause complexes per text, the total number of clauses per sentence, the total number of embedded clauses per text, and the GI level which is the total of clause complex divided by the total of the clause in the text. The number of embedded clauses mentioned here is to show their employment within the texts, whereas the number of clauses written in the table below includes both hypotactic and paratactic structures/clauses.

Table 5 The Results of grammatical Intricacy (GI) Analysis

No	Text Codes	Total of CC	Total of C	Total of EC	GI level
1	007A	19	37	6	1.94
2	018J	10	21	4	2.1
3	023I	19	30	5	1.57
4	037P	19	40	8	2.1

Note:

CC = Clause complex

C = Clause

EC = Embedded Clause

GI = Grammatical intricacy

Table 5 above presents that, surprisingly, three out of four texts have similar numbers of sentences or clause complexes, even though the numbers of clauses are different. These results showed that texts with a high level of density, in this study, were constructed from more complex than simple sentences. However, the number of words and sentences in the text could further affect the employment of various clauses as it is shown that even though Text 018J has the least number of clause complexes, its GI level is similar to Text 037P which

has an almost double number of sentences. Furthermore, the data also show that embedded clauses are also present in the texts although not in a large numbers.

The findings of this study confirm Nguyen and Edwards' (2015) investigation that written language tended to have high LD. The lowest percentage of LD in this study, which is 33.40%, is similar to Nguyen and Edwards' (2015) in the mean of students' proposal LD before the writing training. However, the highest percentage of LD in this study is quite higher compared to the highest LD of students' proposals in their study which was 42.78%. Thus, the texts of the introduction sections of RAs are more lexically dense because the writers could be more experienced and have better writing competence as it is proven that their articles were accepted for publication in reputable journals.

In relation to the level of GI, the results are close to some findings of To (2017) who explored the GI of English textbooks for tertiary students in Vietnam. In his study, the mean calculations of GI from four levels of English textbooks and the present research results are almost the same as for the mean GI of elementary and intermediate books with 1.58 and 1.80. However, the other two texts of this study have a higher level of IG than the textbooks of the previous study. Meanwhile, compared to the study of Syarif (2018), the GI of the introduction sections in this study seems to be lower than the GI of students' thesis proposal introduction sections.

CONCLUSION

From the aforementioned results and discussion above, it can be concluded that the introduction sections of RAs selected from reputable international journals demonstrate average to high lexical density (LD). The range of LD percentage found from the data could indicate that the writers or the researchers did not focus on choosing more economical sentences to express their arguments in the introduction parts of the RAs. It is not impossible that more dense structures may be employed in other parts of the RA. In the case of grammatical intricacy, the level calculated from the introduction sections of RAs of this study is not very high, but it can be obviously recognized that the texts were composed in more ranking clauses, a combination of dependent and independent clauses, than the clause complexes or simple clauses.

The results of this study highlight the importance of more lexical words in academic texts to improve the density of the texts and the complexity of the sentences to increase the interdependency among clauses written in the texts. Thus, it is suggested that writing teachers raise students' awareness of this issue in writing lessons. Besides, future researchers are recommended to analyze a larger corpus and discuss the employment of taxis in more detail to see whether the writers use more hypotaxis or parataxis in their texts.

REFERENCES

- Amelia, P., Sinar, T., & Zein, T. (2020). Lexical density and grammatical intricacy in debaters' speeches. *Language Literacy*, 4(1).
- Butt, D., Fahey, R., Feez, S., Spinks, S., & Yallop, C. (2003). *Using functional grammar: An explorer's guide*.
- Eggs, S. (2004). *Introduction to systemic functional linguistics*. A&c Black.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* third edition. London: Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. (2014). *An introduction to functional grammar*. Routledge.

- Hanafiah, R., & Yusuf, M. (2016). Lexical density and grammatical intricacy in linguistic thesis abstract: A qualitative content analysis. *Proceedings of EEIC*, 1(2), 43-46.
- Ma'mun, N. (2017). Grammatical intricacy on students' writing. *Vision: Journal for Language and Foreign Language Learning*, 6(1), 30-40.
- Mizusawa, Y. (2020). Lexicogrammatical and semantic development in academic writing of EFL learners: A Systemic Functional Approach. *Modern Journal of Studies in English Language Teaching and Literature*, 2(2), 105-117.
- Nguyen, T. H. T., & Edwards, E. C. (2015). An Investigation of Nominalization and Lexical Density in Undergraduate Research Proposals¹. *Language Education in Asia*, 6(1), 17-30.
- Pujiati, R., & Pramudyawardhani, S. R. (2020). Grammatical Intricacy and Lexical Density on Genres in National Examination Texts. *Jurnal Dialektika Program Studi Pendidikan Bahasa Inggris*, 8(2), 202-211.
- Syarif, H. (2019, February). Lexical Density VS Grammatical Intricacy: How Are They Related? In *Sixth of International Conference on English Language and Teaching (ICOELT 2018)* (pp. 16-22). Atlantis Press.
- Syarif, H., & Putri, R. E. (2018). How lexical density reveals students' ability in writing academic text. *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa*, 12(2), 86-94.
- To, V. (2017). Grammatical intricacy in EFL textbooks. *International Journal of English Language Education*, 5(2), 127-140.

Appendix

Link to the analysis tool: ADA Adelex Analyser:

(<http://www.ugr.es/~inped/ada/ada.php?ada=rn6f3fdbcckckgfhb8nd872hg4&lng=English>)

