LEXICAL DENSITY AND GRAMMATICAL INTRICACY IN LINGUISTIC THESIS ABSTRACT: A QUALITATIVE CONTENT ANALYSIS

Ridwan Hanafiah[*] and **Muhammad Yusuf**

Sumatera Utara University, Medan, INDONESIA
[*]*Corresponding author: ridwances@yahoo.com*

Abstract
Abstract of a thesis can be investigated from many angles. This study departs to construe the lexical density (LD) and the grammatical intricacy (GI) in linguistic thesis abstract written by undergraduate English department student of University of Sumatera Utara (USU) in order to figure out the characteristics whether those abstract can be classified into spoken or written language. This study applied qualitative content analysis. The data of this study were all the text of 7 linguistic thesis abstracts. Based on the analysis, it was found that the average score of GI and LD successively is 1.84 and LD index 0.57. As a conclusion, referring to the result of analysis, those abstracts are characterized as written language because of having high degree of LD index which is more than 0.4 and the use of simple language represented by low degree of GI index.

Keywords: *Lexical density, grammatical intricacy, content analysis, thesis abstract.*

INTRODUCTION

Writing thesis abstract is not easy. Graetz (1982, p. 23) asserts that the aim of abstract is to give the reader a precise and concise knowledge of the total content of the very much more lengthy original, a factual summary which is both an elaboration of the title and a condensation of the report. In writing thesis abstract, the students are required to produce language to give a brief summary of the whole thesis at glance. Moreover, thesis abstract can be analyzed from many angles. Ventola (1997, p. 349) argues that abstracts should be taken as a serious object of linguistic study and provides a comprehensive overview of the linguistic analysis of abstracts up to then.

To produce a good abstract, it is necessary for the students to know how to organize message in their writing and the characteristics of written language. Written and spoken texts also have their own complexity namely Lexical Density (LD) and Grammatical Intricacy (GI). Halliday and Matthiessen (2004, p. 654) assert that written language typically becomes complex by being lexically dense: it packs a large number of lexical items into each clause, whereas spoken language becomes complex by being grammatically intricate. Halliday and Matthiessen (2004, p. 655) further explain that the lexical items in the written version thus have fewer clauses to accommodate them but obviously they are still part of the overall grammatical structure — what typically happens is that they are incorporated into nominal groups.

Saragih (2006, p. 9) holds that LD describes number of content words (noun, verb, adjective, and adverb) per clause. Then, LD of a text can be calculated by expressing the number of content carrying words in a text/sentence as a proportion of all the words in the text/sentence (Eggins, 2004, p. 97). When a text is lexically dense, the text is characterized to be in a written mode. As the

example quoted in Saragih (2006, p. 9), the clause complex of "Ali arrived late, which worried us but pleased our rival team" is comprised of 3 clauses and 7 lexical items (printed in bold letters). The LD is 7/3 which gives 2.3 or simply 2. In the same way, the clause "Ali's late arrival resulted in our worries and our rival team's pleasure" which is a single clause has LD of 8/12 or 60% (0.6). In addition, Ure (1971, p.445) concludes that a large majority of the spoken texts have a lexical density of under 40%, while a large majority of the written texts have a lexical density of 40% or higher. It means that lexical density is the representation of content words in a text/sentence and high portion of LD characterizes a text to be in a written mode.

Then, Eggins (2004, p. 97) holds that grammatical intricacy relates to the number of clauses per sentence or clause complexes and the means of calculating by expressing the number of clauses in a text as a proportion of the number of sentences in the text. To exemplify, *Andini was absent because she was sick* is more complex than *her absence was caused by her sickness*. The high ratio or level of GI is the indicator of a text to be categorized as a spoken language. It means that high ratio of GI can be used to characterize a text to be in spoken mode.

In English department of Faculty of Cultural Sciences University of Sumatera Utara, there are two main areas in writing thesis, namely linguistics and literature areas. This study attempts to investigate the characteristics of linguistic abstract in terms of lexical density and grammatical intricacy and to characterize if those abstracts can be categorized into written or spoken language. And the findings of this study can be used as a suggestion for students to write a good abstract.

METHOD

This study applied qualitative content analysis method. Singh (2006, p. 150) states that content analysis, sometimes known as document analysis, deals with the systematic examination of current records or documents as sources of data. So, content analysis is a method used to analyze the data in the written forms (documents) which is the focus of the analysis is the content of it and it is suitable for this study.

Purposive sampling technique was employed in this study. The source of the data was 7 linguistic thesis abstracts by undergraduate students of English department, Faculty of Cultural Sciences, University of Sumatera Utara who graduated during March until October 2014. The data of this study are whole sentences from the abstracts.

FINDINGS AND DISCUSSION
Findings

In order to find out the findings, the data have been analyzed thoroughly. The first analysis is based on grammatical intricacy (GI). It is found the average score of GI is 1.84. The vivid number of GI from the data is illustrated in the following table.

Table 1. Grammatical intricacy characteristics.

Grammatical Intricacy Characteristics								
	1	*2*	*3*	*4*	*5*	*6*	*7*	*Average*
Total Clauses	12	20	24	14	30	19	16	
Total Sentences	8	11	15	7	13	9	10	
Grammatical Intricacy Score	1.5	1.81	1.6	2	2.30	2.1	1.6	**1.84**

Looking at the finding shown above, these low grammatical intricacy characteristics simply indicate that those abstracts have the tendency to be characterized as written language. To make sure, grammatical intricacy is not enough without checking lexical density score. The average score of LD in the data is 0.57 or 57%. The LD ratio of each linguistic abstract is presented in the following table.

Table 2. Lexical density characteristics.

Lexical Density Characteristics								
	1	2	3	4	5	6	7	Average
Total words	165	194	400	194	259	245	340	
Total lexical words	75	94	293	105	166	128	217	
Total Lexical density	45%	50.5%	73%	54%	64%	52%	64%	**57%**

Discussion

As the presented findings, in terms of lexical density, the average score of those abstract is 57% or 0.57 with none of the abstract comprises less than 40% or 0.4. To exemplify, here is the example of a sentence in the data analyzed based on lexical density calculation.

Thesis entitled "An Analysis of Code Switching and Code Mixing Used by Front Office Department Staffs of Grand Elite Medan" is about code switching and code mixing between Indonesian and English in daily conversation when they were working.
(Taken from Abstract 3)

The data presented above shows that there are 27 content words from 38 words as the total word of the data. The percentage of lexical density is 71.05%. There is also nominalization found in the sentences. The use of nominalization in avoiding using complex or intricate grammar has the impact to make the sentence lexically dense and to pack more information. In line to Ure's statement (1971, p. 445), a large majority of the spoken texts have a lexical density of under 40%, while a large majority of the written texts have a lexical density of 40% or higher.

In terms of grammatical intricacy, the finding shows that the average score of GI is 1.84. It means that the students' tend to use simple language in writing their abstract. The statement is supported by the finding. Here is a simple language represented by simple sentence found in the data.

Advertisement is a kind of communication processes between consumer and producer.
(Taken from Abstract 4)

From the example given, it can be seen that there is only one clause in a sentence. It means that the grammatical intricacy of the data is 1. It means that the use of simple language has the impact on grammatical intricacy of the abstract. It can be inferred that those abstracts can be grouped into written text.

Referring to the findings, as the data present the low score of GI and high score of LD, it means that those abstracts are characterized into written language supported. It means that those abstracts definitely fulfill the requirement to be grouped into written language as characterized by Eggins (2004). Then, it is supported by Ure's statement clarifying that a large majority of the spoken texts have a lexical density of under 40%, while a large majority of the written texts have a lexical density of 40% or higher. In the data, the use of nominalization helps to pack more information and makes the text lexically dense while the use of simple language appears in the data makes the low degree of grammatical intricacy.

CONCLUSION

As the conclusion, Referring to the result of analysis, it is found that the average score of GI and LD successively is 1.84 and LD index 0.57 and those abstracts are characterized as written language because of having high degree of LD index which is more than 0.4 and the use of simple language represented by low degree of GI index. The findings of this study are expected to be useful for students in writing abstract.

REFERENCES

Eggins, S. (2004). *An introduction to systemic functional linguistics.* New York: Continuum.

Graetz, N. (1982). Teaching EFL students to extract structural information from abstracts. *Paper presented at the International Symposium on Language for Special Purposes* (pp. 1–23). August 2-4, Eindhoven, The Netherlands.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3ʳᵈ Ed.). London: Edward Arnold.

Saragih, A. (2006). Metaphorical representations and scientific texts. *Englonesia: An Indonesian Scientific Journal on Linguistics and Literature, 2*(1), 1-11.

Singh, Y.K. (2006). *Fundamental of research methodology and statistics.* New Delhi: New Age International (P) Limited.

Ure, J. (1971). Lexical density and register differentiation. In G. E. Perren & J. L .M. Trim (Eds.). *Applications of linguistics: Selected Papers of the Second International Congress of Applied Linguistics* (pp. 443-452). Cambridge: Cambridge University Press.

Ventola, E. (1997). Abstracts as an object of linguistic study. In F. Danes, E. Havlova & S. Cmejrkova (Eds.), *Writing vs. speaking: Language, text, discourse, communication* (pp. 333-352). *Proceedings of the Conference held at the Czech Language.* Tübingen: Günter Narr.