# AN ANALYSIS OF LEXICAL DENSITY AND GRAMMATICAL INTRICACY ON SCIENTIFIC ARTICLES IN JOURNAL OF ENGLISH EDUCATION VOLUME 4, NUMBER 1, 2019, PUBLISHED BY USN SCIENTIFIC JOURNAL, INDONESIA

**[1]Nur Eka Sari**
nurekasari366@gmail.com


**Universitas Muhammadiyah Kotabumi**

**Abstract:** Writing scientific articles for university students is an obligation. However, in the process of writing they found some difficulties. Their writing tends to use a lot of words that are not appropriate and the sentence they made eventually became too long and difficult to understand. There are two indicators for written language that can help students: lexical density and grammatical intricacy. Lexical density shows how much lexical items are in the text and grammatical intricacy shows how much clauses are in the sentence. Both of those show that the article from is in written language. Written language has high percentage of lexical density that is up tp 40% and low ratio of grammatical intricacy. Thus, the researcher decided to conduct an analysis of lexical density and grammatical intricacy on scientific articles in Journal of English Education Volume 4, Number 1, 2019, Published by USN Scientific Journal, Indonesia. This study aims to investigate: (1) how high the percentage of lexical density of those scientific articles, (2) how high the ratio of grammatical intricacy on those scientific articles, (3) the scientific articles can be classified into spoken language or written language. Qualitative descriptive method was used in this research. From the analysis it can be seen that the percentage of lexical density in six articles are 51.93%, 51.27%, 51.2%, 50.1%, 49%, and 52.16%. The ratios of grammatical intricacy in six articles are 1.59, 1.72, 2.29, 1.77, 1.99, and 1.97. The percentages are up to 40% and considered as high, while the ratios are under 3 and considered as low. Based on the criteria, those six articles used written language. Thus, it can be concluded that the writers of those articles have good knowledge of the use of language in writing scientific articles.

**Keywords:** analysis, lexical density, grammatical intricacy, scientific articles


***Abstrak: Menulis artikel ilmiah bagi mahasiswa merupakan kewajiban. Namun dalam prosesnya mereka menemui beberapa kesulitan. Tulisan mereka cenderung menggunakan banyak kata yang tidak sesuai dan kalimat yang mereka buat akhirnya menjadi terlalu panjang dan sulit untuk dipahami. Ada dua indikator bahasa tulis yang dapat membantu siswa: kepadatan leksikal dan kerumitan grammatikal. Kepadatan leksikal menunjukkan berapa banyak item leksikal dalam teks dan kerumitan grammatikal menunjukkan berapa banyak klausa dalam kalimat. Keduanya menunjukkan bahwa artikel ditulis dengan bahasa tertulis. Bahasa tulis memiliki persentase kepadatan leksikal yang tinggi yaitu di atas 40% dan rasio kerumitan grammatikal yang rendah. Maka dari itu peneliti memutuskan untuk melakukan analisis kepadatan leksikal dan kerumitan grammatikal pada artikel ilmiah di Journal of English Education Volume 4, Nomor 1, 2019, diterbitkan oleh USN Scientific Journal, Indonesia. Penelitian ini bertujuan untuk menginvestigasi: (1) seberapa tinggi***

[1)]Mahasiswa Universitas Muhammadiyah Kotabumi

*persentase kepadatan leksikal artikel ilmiah tersebut, (2) seberapa tinggi rasio kerumitan gramatikal pada artikel ilmiah tersebut, (3) apakah artikel ilmiah tersebut dapat diklasifikasikan ke dalam bahasa lisan atau bahasa tertulis. Metode deskriptif kualitatif digunakan dalam penelitian ini. Dari hasil analisis dapat diketahui bahwa persentase kepadatan leksikal pada enam artikel adalah 51.93%, 51.27%, 51.2%, 50.1%, 49%, dan 52.16%. Rasio kerumitan grammatikal pada enam artikel adalah 1.59, 1.72, 2.29, 1.77, 1.99, dan 1.97. Persentase artikel-artikel tersebut di atas 40% dan cenderung tinggi, sedangkan rasio artikel-artikel tersebut di bawah 3 dan cenderung rendah. Berdasarkan kriterianya, artikel-artikel tersebut telah menggunakan bahasa tertulis. Sehingga dapat disimpulkan bahwa penulis-penulis artikel tersebut memiliki pengetahuan yang baik tentang penggunaan bahasa dalam menulis artikel ilmiah.*

***Kata kunci:*** *analisis, lexical density, grammatical intricacy, artikel ilmiah*

## I.  INTRODUCTION

Writing scientific articles for university students is an obligation. University students will be assigned to certain subjects to write scientific articles. They will be asked to write scientific articles to be published in a journal in their field. Publishing is useful to ensure the quality of their own writing and also to improve and maintain the quality of the place where they study.

Students who come from the English department certainly have to make scientific articles in English. The process of making scientific articles is begun with the choosing of the topic, making the conception, finding the related theory, and writing it. However, in this process, they will find some difficulties. Weisman (as cited in Setiani, 2019, p. 29) states that amongs the four domains of language, writing can be the most challenging. That is because the students have to understand the whole scientific article and have to write it accurately. They have difficulties in understanding scientific articles because they still have limited knowledge of the vocabulary used in the scientific articles.

Vocabulary used in scientific articles use scientific language and terms. The scientific language is less familiar to students. Thus in writing, they tend to use a lot of words that are not appropriate. The sentences they made eventually became too long and difficult to understand. Supposedly, scientific articles that have been published in journals have rich information about a research. Therefore, it is assumed that scientific articles not only have a lot of information but also are written effectively and not convoluted. The lack of vocabulary mastery and convoluted sentence arrangement can be seen from lexical density and grammatical intricacy.

Lexical density and grammatical intricacy distinguish between written and spoken language. Written language has a higher lexical item ratio to the total number of words and has less grammatical intricacy, while spoken language has lower lexical density and more grammatical intricacy (Halliday, 1990, p.61-62). The higher the lexical density ratio of a text, the more information contained there. That is because lexical items are words that contain meaning or information. While lexical density is dealing with word, grammatical intricacy is dealing with clause.

Grammatical intricacy in the text relates to the number of clauses per sentence, and can be calculated by displaying the number of clauses as a proportion of the total number of sentences in the texts (Eggins, 2004, p.97). While the higher the grammatical intricacy ratio, the more difficult the text is understood by readers because of many clauses that exist in a text. While spoken language such as interviews or dialogues will have a high grammatical intricacy ratio. If it is in practice speaking directly, it does not make the listener find it difficult to grasp its meaning due to the context of the accompanying situation. Thus, the researcher decided to analyze lexical density and grammatical intricacy in scientific articles.

Scientific articles are analyzed as ways of adding insight into how to analyze and to know what the results of lexical density and grammatical intricacy analysis look like in these articles. The scientific articles reviewed are as follow. Each scientific article in the Journal of English Education Volume 4, Number 1, 2019, published by the USN Scientific Journal, Indonesia will be analyzed. The part analyzed is every part of the articles, there are: title, abstract, introduction, discussion and conclusion. For the exception are tables, diagrams, and references. After calculating the percentage of lexical density and the ratio of grammatical intricacy, the researcher described whether the articles are considered into spoken language or written language.

Based on the explanation above, the essential of problems that would be described in this research can be expressed by following questions:

1. How high is the percentage of lexical density on scientific articles in Journal of English Education Volume 4, Number 1, 2019, published by USN Scientific Journal, Indonesia?

2. How high is the ratio of grammatical intricacy on scientific articles in Journal of English Education Volume 4, Number 1, 2019, published by USN Scientific Journal, Indonesia?

3. Are the scientific articles in Journal of English Education Volume 4, Number 1, 2019, published by USN Scientific Journal, Indonesia can be classified into spoken or written language?

## II.   THEORY REVIEW

**Systemic Functional Linguistic (SFL)**

The language that people use, whether oral or written, can be understood by others, is referred to as text. Eggins (2004, p. 1) explains that in contemporary life, people require to produce and react bits of language that make sense; they are required to negotiate text. The text is analyzed in an approach namely Systemic Functional Linguistics (SFL). Halliday (as cited in Eggins, 2004, p.3) argues that the purpose of language as a social semiotic is how people use language with each other in accomplishing everyday social life. Thus, this approach is talking about how someone uses their language and how they can understand it in their life.

A text both spoken and written has a context that influences how a person understands it. The context is divided into two, namely the context of the situation or called register theory and cultural context or called genre. Register theory illustrates how it can interpret the text directly based on the context. This is supported by Eggins (2004, p. 9) who argues that register theory describes the impact of dimensions of the immediate context of the situation of a language event on the way the language is used.

There are three dimensions included in register theory (Eggins, 2004, p. 9). The first is mode, which is amount of feedback and role of language: spoken language or written language (mode variation). The second is tenor, which are the role relations of power and solidarity: how to talk to their boss as to talk to their lover (tenor variation). The third is field; which is topic or focus of the activity: how to talk about linguistics as to talk about jogging (field variation).

Genre is about how culture influences someone to understand the text. Some texts that are similar to each other, or have the same genre, will feel more familiar to both producers and recipients. The concept of genre is used to describe the impact of the context of culture on language, by exploring the staged, step-by-step structure culture institutionalizing as ways of achieving goals (Eggins, 2004, p.9).

From the explanation above it can be concluded that a text is analyzed in SFL. SFL is a concept that understands the using of language in social life. How the text is produced and understood based on the context, namely the context of the situation

or called register theory and cultural context or called genre.

## Spoken and Written Language

Based on mode variation in register theory, there are spoken and written language. Halliday (1990, p. 76) explains that spoken language is characterized by the absence of certain features that are found in writing. He also explains that spoken language has short silences (unfilled paused) and all other types of hesitation—false starts, repetitions, and filled paused. While written language is going through drafting, editing, and rewriting also using formal language (Eggins, 2004, p. 92). That makes writing using more effective sentences. However, there are also generalizations of characteristic that can distinguish between spoken and written language, and the features are in both.

The distinguishing characters are lexical density in written language and grammatical intricacy in spoken language. Halliday (1990, p. 62) said that the difference between written and spoken languages is one of density: the density of which information is presented. Written language is dense, spoken language is sparse. He also said that another difference is one of intricacy, the intricacy with which the information is organized. Spoken language is more intricate than written language. In this case it can be concluded that the written language has a high lexical density and has a low grammatical intricacy, while the spoken language has a low lexical density and a high grammatical intricacy.

From the explanation above, it can be concluded that the difference between written and spoken language is in its features, where what are in written language are not in spoken language and the opposite of it, what are not in written language are in spoken language. However, there is also a single generalization that distinguishes written and spoken language but exists in both, namely lexical density and grammatical intricacy.

## Lexical Density

Lexical density is a measure that can be used to see how much information is contained in text. As the name implies, the information referred to here is information contained in lexical items in the text. According to Halliday and Martin (2005, p. 83), lexical density is a measure of information in any passage of text, based on how high the lexical items (content words) have been packed into the grammatical structure.

Ure (as cited in To, Fan, and Thomas, 2013, p. 62) also said that lexical

density should be treated as the proportion of the number of lexical items per the number of running words. This is also in line with the opinion of Eggins (2004, p. 97), lexical density in text or sentence can be calculated by displaying the numbers of content carrying words in the text as a proportion of the total number of words in the text or sentence. Ure (as cited in Hanafiah, 2018, p. 131) the written language has lexical density 40% or higher. Therefore, in Ure and Eggins's opinion, lexical density is also the number of lexical items in a text, compared to the total number of words and written language has minimum 40% of lexical density. Lexical items are content words, while items other than lexical are grammatical items or also called function words. There are divisions of part of speech that which are included in lexical items and which are included in grammatical items.

Eggins (2004, p. 97) explained that content-carrying words include nouns, the main part of verb, adverbs, and adjectives; non content-carrying words include prepositions, conjunctions, auxiliary verbs, and pronouns. Content-carrying words are another term for lexical items and content words, while non-content-carrying words are another term for grammatical items and function words. Whereas Halliday (1990, p. 61) said that grammatical items are determiners, pronouns, most prepositions,

conjunct-tions, some classes of adverb, and finite verbs. Lexical density is also one of the characteristics that distinguish between spoken and written language. According to Halliday (1990, p. 61), written language displays a much higher ratio of lexical items to total running words. Halliday (1990, p. 62) also explained that the difference between written and spoken languages is one of density: the density with which information is presented and written language is dense while spoken language is sparse.

In calculating lexical density in text, the formula can be seen as follows:

$$\text{Lexical Density} = \frac{\text{number of lexical items}}{\text{total numbers of words}} \times 100\%$$

Source: Ure (as cited in To, Fan, and Thomas, 2013, p. 63)

Halliday (1990, p. 64) stated that to measure lexical density, draw the distinction between lexical items and grammatical items, simplifying it by treating each word as the relevant items, and counting the ratio of lexical to grammatical items. From the explanation above, it can be concluded that lexical density is the number of lexical items in the text. The higher lexical density is the more information in it. The high lexical density also shows its characteristics as a written language because spoken language is sparse.

## Grammatical Intricacy

Grammatical intricacy is a term to explain how many clauses contain in one sentence. According to Eggins (2004, p. 97), grammatical intricacy relates to the number of clauses per sentence. It is supported by Halliday (1990, p. 62) that state the difference between written and spoken language is one of intricacy; the intricacy in which information is organized, spoken language is more intricate than written language. Thus, grammatical intricacy shows how convoluted a sentence and how complex a language; where in one sentence has many clauses. From those statements, spoken language has a high grammatical intricacy than written language.

Spoken language has a high grammatical intricacy because when someone talks they tend to connect ideas that cross their minds. According to Eggins (2004, p. 97), in spoken language people tend to chain clauses together one after another, that makes a sentence very long, while in written language people tend to use relatively few clauses per sentences. Although in actual conversation it is not too difficult to understand—due to the pausing, repetition, and correcting speech—but in writing it will make it difficult for the readers to understand its contents. An independent clause is easier to read than many clauses in one sentence.

While lexical density has formula to calculate it, grammatical intricacy also has one. It is from Halliday (as cited in Sujatna, 2019, p.34), grammatical intricacy can be calculated by dividing the total number of ranking clauses with the total number of clause complexes. The formula is as follows.

$$\text{Grammatical Intricacy} = \frac{\text{Total number of ranking clauses}}{\text{Total number of clause complexes}}$$

Source: Halliday (as cited in Sujatna, 2019, p. 34)

Ranking clause means each clause found in the text, while clause complex has the same meaning with the sentence. Clause complex is an attempt to examine the notion of clause, combination of clauses, and brief distinction between clause complex and traditional notion of sentence (Halliday as cited in Sujatna, 2019, p. 34-35). In other word, clause complex replaces the term sentence in this research to differentiate it from the traditional meaning of a sentence. That is a capital later following a full stop until the next full stop (Halliday, 1990, p. 66). From that formula, grammatical intricacy can be calculated by count how much ranking clauses in a text, and how much clause complexes in a text. After calculating the total number of them, put them into the

formula. Thus, the ratio of grammatical intricacy can be measured.

By the explanation above, it can be concluded that grammatical intricacy shows how much clauses in a sentence. If the grammatical intricacy is high, then it will be characterized into spoken language. It can be measured with calculating the total number of ranking clauses divided by the total number of clause complexes in the text.

## III.  RESEARCH METHOD

This research conducted by using qualitative method. Even though this research used number, the numbers were not the result of an experiment but only show the data. According to Bogdan & Taylor (as cited in Moleong, 2017, p. 4) qualitative method is a research procedure that produces descriptive data in the form of written or spoken words from people and behavior that can be observed and directed at individual setting holistically or as a whole. Therefore, the data obtained in this study described in the form of description and tables.

Based on the purpose of this study, to identify the percentage of lexical density and the ratio of grammatical intricacy, this research is included in qualitative descriptive research. This kind of research is not yet a perfect qualitative study

because it is still influenced by quantitative traditions, especially in applying the theory to the data obtained (Bungin, 2011, p. 68).

The data resource in this research was a data document. This research used Journal of English Education volume 4, number 1, 2019, published by USN Scientific Journal, Indonesia. USN is stand for Universitas Sembilanbelas November. The journal contains six scientific articles.

The instrument used in this study was the researcher herself. According to Moleong (2017, p. 168) researcher is instrument of qualitative research because she is a tool for collecting data such as tests on quantitative research. The researcher was used as a tool to collect data while analyzing data and having knowledge about Lexical Density and Grammatical Intricacy.

The data collecting technique in this research is analysis on data document. According to Emzir (2014, p. 17) Data collection in qualitative research, generally includes observation, inter-views, and document analysis. This method is used to obtain data in the form of lexical items, ranking clauses, number of words, and number of clause complexes needed in the analysis of percentage of lexical density and ratio of grammatical intricacy.

In qualitative research, internal validity is replaced by credibility. Credibility means the results of the study

are credible or trustworthy from the perspective of the participants in the study (Emzir, 2014, p. 79). One way to test credibility is by peer debriefing. Bungin (2011, p. 266) stated that discussions with those who understand research problems will provide meaningful information to the researcher as well as an effort to test the validity of research results. This method was done by exposing the provisional results of the research, so that other parties could clarify if there were mistakes and the final results of the research could be exposed.

Reliability in qualitative research is dependability. Dependability is essentially related to whether the research results will be the same if an observation is made for the second time (Emzir, 2014, p. 80). In this study reliability was improved by calculating lexical items, clauses, words, and sentences over and over to avoid errors, and analyzing lexical density and grammatical intricacy carefully.

Data analysis was performed to obtain the results of the study by processing the data. This study used data analysis techniques proposed by Miles and Hubberman. In the analysis of Miles and Huberman data models (Emzir, 2014, p. 129-135) there are three kinds of activities in qualitative data analysis namely:

1) Data reduction

Data reduction is an activity to sharpen, select, focus, dispose and arrange data in a way so that the final conclusions can be drawn and verified. In data reduction, the researcher chose which data used in the study, and which data is not used. In this study the data used are only lexical items, number of ranking clauses, number of words, and number of clause complexes contained in the text.

2) Data Display

The second step of this analysis is data display. The data can be displayed in the form of narratives, charts, graphs, tables and so on. Data display aims to arrange data so that researcher is easy to access data directly. The data display in this research is in the following table.

3) Conclusion drawing and verification

The third step of data analysis is drawing conclusion and verification. The conclusions that have been drawn are then verified by looking back at what has been analyzed briefly.

## IV. RESULTS AND DISCUSSION

In this part, the research results and discussion were presented. The following is a table of the results of lexical density and grammatical intricacy analysis in six scientific articles in the Journal of English Education Volume 4, Number 1, 2019,

published by the USN Scientific Journal, Indonesia.

**TABLE 1**
**Analysis Results of Lexical Density and Grammatical Intricacy on Scientific Articles**

| No. | Title of Article | LD | GI |
|---|---|---|---|
| 1. | The Students and Teacher's Views on the Use of E-textbooks in Reading Classroom | 51.93% | 1.59 |
| 2. | Relevance of Bilingualism in Adult Learners' English Language Development in South-West Nigeria | 51.27% | 1.72 |
| 3. | From Inland to Outland: Experiences of Non-native Expatriate Teachers Teaching in a Foreign Context | 51.2% | 2.29 |
| 4. | Use of Hand Gestures through Digital Video to Improve Phonemic Awareness Among Year 2 Pupils | 50.1% | 1.77 |
| 5. | The Role of Instrumental Motivation among EFL Students in Language Learning Process | 49% | 1.99 |
| 6. | The Difficulties of Writing Scientific Work at the English Education Students | 52.16% | 1.97 |

From the table above, it can be seen that the percentage of lexical density in six articles are 51.93%, 51.27%, 51.2%, 50.1%, 49%, and 52.16%. Based on the opinion of Ure (as cited in Hanafiah, 2018, p. 131) the written language has lexical density 40% or higher. From the result of analysis above, one of six articles had 49% of lexical density and the others are higher

than that, thus it qualifies to conclude that this article uses written language. From these results it can also be seen that the number of lexical words is dense, so there is also a lot of information contained in the article. Grammatical intricacy also found good results.

The ratios of grammatical intricacy in six articles are 1.59, 1.72, 2.29, 1.77, 1.99, and 1.97. The six articles have results under 3, and this shows that the number of clauses contained in the sentences in the article is not that much. They are about 2-3 clauses per sentence. Halliday (1990, p. 62) said that spoken language is more intricate than written language. With a low intricacy ratio, it can be concluded that these articles use written language. When the number of clauses used in a sentence is not much, the sentence will be easily understood by the readers.

At the end, it can be concluded that the lexical density and grammatical intricacy in those six articles have the appropriate criteria to be categorized as written language. The highest number of lexical items is nouns, followed by verbs, adjectives, and the least is adverbs. On closer examination, it can be seen that the sentences in these articles sometimes has many clauses in it, for example, eight clauses in one sentence, because this is accompanied by the number of simple sentences (sentence that has one clause),

the analysis results showed an average of two clauses per sentence. Some sentences from interview script are understandable if they have so many clauses in it. Basically, it is because they are from spoken language. Thus, article writers should be careful when explaining long things. They can break a long sentence into several sentences.

## V. CLOSING

The lack of vocabulary mastery and convoluted sentence arrangement from students' writing can be seen from lexical density and grammatical intricacy. Therefore, the researcher investigated six scientific articles written by college students to see how high the lexical density and grammatical intricacy there. After that, the researcher has drawn the conclusion whether the articles can be classified into spoken or written language.

From the results presented, the three questions in the formulations of the problems were answered. These questions are: how high the percentage of lexical density, how high the ratio of grammatical intricacy, and are the scientific articles in Journal of English Education Volume 4, Number 1, 2019, published by USN Scientific Journal, Indonesia can be classified into spoken language or can be classified into written language.

It is found that the lexical density percentages in the six articles sequentially are: 51.93%, 51.27%, 51.2%, 50.1%, 49%, and 52.16%. From the result of analysis above, the six articles have lexical density 49% and higher, so it can be concluded that these articles use written language. The ratios of grammatical intricacy in the six articles sequentially are 1.59, 1.72, 2.29, 1.77, 1.99, and 1.97. The six articles have results below 3, and it can be concluded also from this criterion that these articles use written language.

The scientific articles in the Journal of English Education Volume 4, Number 1, 2019, published by the USN Scientific Journal, Indonesia are using written language. Thus, in the end it can be concluded that the writers of those articles have a good knowledge of the use of language in writing scientific articles.

## REFERENCES

Bungin, B. (2011). *Penelitian Kualitatif.* Kencana Perdana Media Grup.

Eggins, S. (2004). *An Introduction to Systemic Functional Linguistic 2nd Edition.*

Continuum.

Emzir. (2014). *Metodologi Penelitian Kualitatif: Analisis Data.* Rajawali Press.

Halliday, M. A. (1990). *Spoken and Written Language.* Oxford University Press.

Halliday, M. A., & Martin, J. R. (2005). *Writing Science: Literacy and Discursive Power.* Taylor and Francis.

Hanafiah, R. (2018). An Analysis of Lexical Density and Grammatical Intricacy in Thesis Abstract of English Department Students of USU. *Asian EFL Journal*, 20(7), 130-134.

Moleong, L. J. (2017). *Metode Penelitian Kualitatif.* PT Remaja Rosdakarya.

Setiani, R. (2019). Students' Ambiguity in Writing an Essay. *Edukasi Lingua Sastra*, 17(1), 29-39. https://doi.org/10.47637/elsa.v17i1

Sujatna, E. T. (2019). Readibility of the Transactional Text in Bahasa Inggris Textbook for Grade X. *International Journal of Systemic Functional Linguistics*, 2(1), 33-42. http://dx.doi.org/10.22225/ijsfl.2.1.996.33-42

To, V., Fan, S., & Thomas, D. (2013). Lexical Density and Readibility: A Case Study of English Textbooks. *Internet Journal of Language, Culture, and Society*, (37), 61-71.