

A CORPUS STUDY OF THE TOKENS OF NOUN PHRASES, PREPOSITIONAL PHRASES, AND CLAUSES FOUND IN THE ENGLISH WRITTEN TEXTS WRITTEN BY ENGLISH NATIVE SPEAKERS AND INDONESIAN WRITERS

Hertiki

Dosen Prodi. Pend. Bahasa Inggris FKIP Universitas PGRI Adi Buana Surabaya

Abstrak

Media cetak seperti surat kabar, majalah, jurnal, buku merupakan salah satu contoh dari media tulis yang seringkali digunakan untuk berkomunikasi dan berbagi pengalaman maupun pendapat. Saat ini baik artikel maupun jurnal dalam Bahasa Inggris pada umumnya ditulis oleh penutur asli yang memang menggunakan Bahasa Inggris sebagai bahasa pengantar dalam kehidupan sehari-hari, tetapi sekarang orang Indonesia pun berusaha untuk menulis jurnal dalam Bahasa Inggris sekalipun Bahasa Inggris digunakan sebagai bahasa asing dalam kehidupan sehari-hari. Makalah ini membahas penelitian korpus yang ditinjau dari jumlah frasa kata benda, frasa kata depan, dan klausa yang ditemukan di jurnal yang ditulis dalam Bahasa Inggris oleh penutur asli dan orang Indonesia. Data penelitian berupa frasa kata benda, frasa kata depan, dan klausa yang terdapat dalam sumber data. Adapun sumber data yang digunakan adalah tiga artikel ilmiah dari tiga jurnal meliputi jurnal ekonomi, jurnal pendidikan dan jurnal psikologi. Hasil penelitian ini menunjukkan bahwa orang Indonesia pun memiliki kemampuan yang sama dengan penutur asli dalam menulis jurnal dalam Bahasa Inggris ditinjau dari frasa kata benda, frasa kata depan dan klausa.

Key words: *corpus, tokens, noun phrases, prepositional phrases, clauses.*

Background of the Study

People have been communicating a lot using media. Printed media such as: newspapers, magazines, journals, books were the first media to communicate and to share ideas to public. Most of the English written texts are written by English native speakers, but now Indonesian speakers have tried to write English texts too. In this study, the writer uses journals to be analyzed. The writer wants to know whether Indonesian writers can write as well as English native speakers based on noun phrases, prepositional phrases, and clauses. She uses some journals written by English native speakers and Indonesian writers with different fields. There are some fields nowadays, but Economic, Education and Psychology have been selected because they are scientific English written texts that most written by both of the speakers.

The writer's analysis is based on the theory of corpus. A corpus essentially tells us what language is like, and gives information about how a language works. As stated by Hunston (2002:2), "a corpus is described as a

collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written text or tape recordings, which have been collected for linguistic study." There are some key terms used in corpus, the writer is interested to make a study about token. As what Biber in Hunston (2002:17) says, "a token is the figure that word-count function of a word-processing program gives." In this case, the writer emphasizes her study in analyzing noun phrases, prepositional phrases, and clauses using content analysis especially parsing technique which divide the sentences into their constituents. She is interested in examining the noun phrases, prepositional phrases, and clauses of English texts since those parts can make sentences become longer and also occur most frequently in sentences.

Statements of the Problem

Related to the background of the study, this study is designed to compare the corpora of English sentences written by English native speakers and Indonesian speakers. With this

study, the writer would like to answer the questions as follow:

1. Do noun phrases written by English native speakers have more tokens than those written by Indonesian speakers?
2. Do prepositional phrases written by English native speakers have more tokens than those written by Indonesian speakers?
3. Do clauses written by English native speakers have more tokens than those written by Indonesian speakers?

Objectives of the Study

Closely related to the statements of the problems, this study intended:

1. To find out whether the corpora of English sentences which are written by English native speakers have more noun phrases than Indonesian writers or not.
2. To find out whether the corpora of English sentences which are written by English native speakers have more

prepositional phrases than Indonesian writers or not.

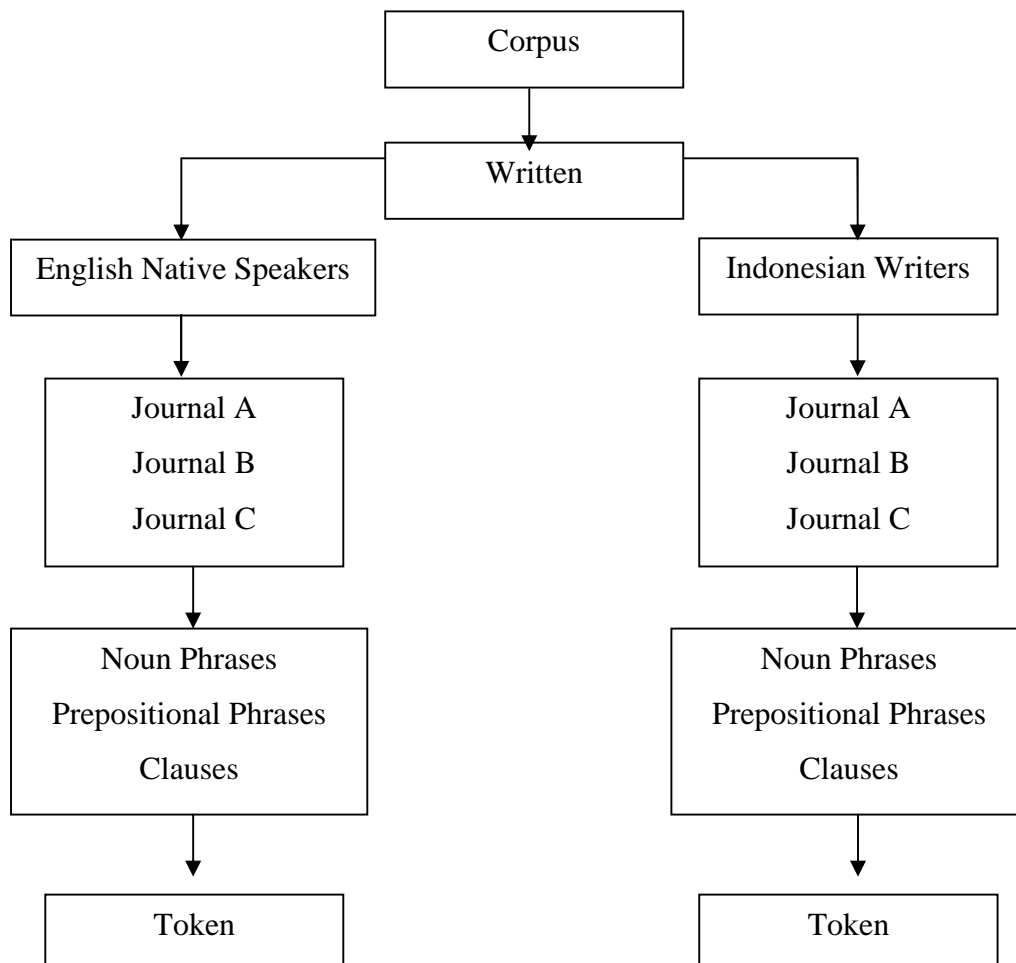
3. To find out whether the corpora of English sentences which are written by English native speakers have more clauses than Indonesian writers or not.

Significance of the Study

The writer hopes that the result of this study can show the ability of Indonesian writer in writing long sentences.

Research Design

In doing her research, the writer uses the qualitative content analysis research. It is qualitative since the analysis of the data based on the understanding toward the context by reviewing the theory sources. Based on the objectives of the study, the writer uses the content analysis research design to carry out this study. Here, the writer analyzes the content of the data from journals. The research design for this study can be seen from the diagram below.



The Subject of the Study

The subjects of this study are journals. There are six journals in different topics. Three journals are written by English native speakers and the other three are written by Indonesian writers. The three fields that are used are:

- a. Economic
For English native speaker, the writer uses the Journal of Accounting Research written by Steven Monahan entitled Discussion of the Value Relevance of Internet Firms: Does Reporting Grossed-up or Barter Revenue Make a Difference?. This journal is published in 2002. On the other hand, the writer uses the VENTURA, Journal Ekonomi Bisnis dan Akuntansi for Indonesian speakers. The journal is written by Lutfi entitled Board Structure, Managerial Ownership, and Shareholder Wealth: A Review of Empirical Evidence. It is published in 2001.
- b. Education
For English native speaker, the writer uses the TESOL JOURNAL, written by James Minor entitled Incorporating Service Learning into ESOL Programs. This journal is published in 2002. While the writer uses the English-EDU, Journal of Language Teaching and Research for Indonesian speakers. This journal is written by Gusti Astika entitled A Discourse Approach to Teaching Structure in Proposal Writing. It is published in 2003.
- c. Psychology
For English native speaker, the writer uses the Anima, Indonesian Psychological Journal for Indonesian speakers. This journal is written by Agnes Maria entitled The Cognitive Model and Interventions for Specific Reading Disorder. It is published in 2002.

The Research Instrument

The research instrument of this study is the writer. Here, the writer collects and analyzes the data from journals. The writer seeks whether the corpora of English sentences written by English native speakers and Indonesian writers have more tokens or not.

The Procedure of Collecting Data

In doing her research, the writer collects the data from some journals. The writer takes the following steps:

1. The writer selects journals written by English native speakers and Indonesian writers that are used in this study.
2. In selecting the journals, the journals should be published in the year of 2001-2003.

The Procedure of Data Analysis

The writer analyzed the data by these following steps:

1. The writer looks for the noun phrases, prepositional phrases, and clauses from three fields using content analysis especially parsing technique which divide the sentences into their constituents.
2. When the writer finishes the analysis, there are some points to be considered:
 - a. For the resources (name and year) will not be considered because the competence of the speakers is not directly connected with long sentences.
 - b. For percentage, like 10 % considered as 2 tokens.
 - c. For the abbreviation, like EFL considered as 1 token.
 - d. Noun phrases in prepositional phrases will not be counted. For example: “to the results”.
 - e. Finally, noun phrase in noun phrase also will not be counted, e.g. Evidence regarding **the value relevance** of **the revenues**.

NP

NP

3. After that, she puts them on the lists based on the two categories whether it belongs to English native speakers or Indonesian writers.
4. The writer counts how many tokens of noun phrases, prepositional phrases, and clauses from those three topics and average them. The example of noun phrases written by English native speakers:
 - a. The analysis of the sentences in the proposal = 8 tokens
 - b. Grammar instruction = 2 tokens

- c. This papers = 2 tokens
 - d. These studies = 2 tokens
 - e. My flight = 2 tokens
 - f. The reason for this = 4 tokens
 - g. The semantics of the verb = 5 tokens
 - h. Teachers = 1 token
- TOTAL = 26 tokens

$$\frac{\text{The average of length of noun phrases} = \frac{\text{The total number of token}}{\text{The total number of noun phrases}}}{8}$$

$$= \frac{26}{8} = 3.3 \text{ tokens}$$

So, it can be concluded that the average of length of noun phrases written by English native speakers from three fields is 3.3 tokens. For the other parts, it is also done as the example above.

- 5. The writer clarifies which English sentences have more tokens, whether English native speakers or Indonesian writers.

Findings

It is dealing with the analysis of noun phrases, prepositional phrases, and clauses from the six journals of three fields – economics, education and psychology. The journals are written by Indonesian writers and English Native Speakers. In order to have a clear description, the writer would like to give an example of the analysis of noun phrases from economics fields written by English Native Speakers.

- 1. Discussion of the Value Relevance of Revenue for Internet Firms = 10 tokens
- 2. Reporting Grossed-up = 3 tokens

- 3. Barter Revenue = 2 tokens
 - 4. A Difference = 2 tokens
 - 5. Introduction = 1 token
 - 6. Davis = 1 token
 - 7. The primary results = 3 tokens
- TOTAL = 23 tokens

$$\text{Average} = \frac{\text{The total number of tokens}}{\text{The total of noun phrases}}$$

$$= \frac{23}{8} = 2.87 \text{ tokens}$$

So, it can be concluded that the average of length of noun phrases written by English native speakers for economic field is 2.87 tokens. After analyzing the data from the text in the journals, the writer makes an average for noun phrases, prepositional phrases, and clauses with the formula of calculation as follows:

$$\text{Average} = \frac{\text{The total number of tokens}}{\text{The total of noun phrases}}$$

$$\text{Average} = \frac{\text{The total number of tokens}}{\text{The total of prepositional phrases}}$$

$$\text{Average} = \frac{\text{The total number of tokens}}{\text{The total of clauses}}$$

To give a clear description of all types' analysis in each unit, a complete table is given as a summary of all the categories.

Table 1.

Journals	ECONOMICS	EDUCATION	PSYCHOLOGY	Average of each unit
English Native Speakers	Noun phrases 5.44	Noun phrases 5.12	Noun phrases 6.19	$= \frac{5.44 + 5.12 + 6.19}{3}$ = 5.58
	Prepositional phrases 7.45	Prepositional phrases 6.88	Prepositional phrases 8.33	$= \frac{7.45 + 6.88 + 8.33}{3}$ = 7.55
	Clauses 16.31	Clauses 16.65	Clauses 14	$= \frac{16.31 + 16.65 + 14}{3}$ = 15.65
Average of each field	$= \frac{5.44 + 7.45 + 16.31}{3}$ = 9.73	$= \frac{5.12 + 6.88 + 16.65}{3}$ = 9.55	$= \frac{6.19 + 8.33 + 14}{3}$ = 9.51	9.59

Table 2.

Journals	ECONOMICS	EDUCATION	PSYCHOLOGY	Average of each unit
Indonesian Writers	Noun phrases 6.27	Noun phrases 6.04	Noun phrases 5.69	$= \frac{6.27 + 6.04 + 5.69}{3}$ = 6
	Prepositional phrases 6.96	Prepositional phrases 7.25	Prepositional phrases 7.47	$= \frac{6.96 + 7.25 + 7.47}{3}$ = 7.23
	Clauses 16.67	Clauses 16.08	Clauses 13.88	$= \frac{16.67 + 16.08 + 13.88}{3}$ = 15.54
Average of each field	$= \frac{6.27 + 6.96 + 16.67}{3}$ = 9.97	$= \frac{6.04 + 7.25 + 16.08}{3}$ = 9.79	$= \frac{5.69 + 7.47 + 13.88}{3}$ = 9.01	9.59

From the table above, the writer finds that the average of noun phrases written by English native speakers is 5.58 tokens; while the average of noun phrases written by Indonesian writers is 6 tokens. Second, the average of prepositional phrases written by English native speakers is 7.55 tokens and the average for Indonesian writers is 7.23 tokens. Third, the average of clauses written by English native speakers is 15.65 tokens and the average by Indonesian writers is 15.54 tokens. So, the result concludes that Indonesian writers write more noun phrases than English native speakers.

In this research, the writer also finds that the average of economic field written by English native speakers is 9.73 tokens. On the other hand, the average of economic fields written by Indonesian speakers is 9.97 tokens. Then for the second journal, the average of education journals written by English native speakers is 9.55 tokens, but the average of education journals written by Indonesian speakers is 9.79 tokens. Next for the last journal, the average of psychology journal written by English native speakers is 9.51 tokens, and the average for Indonesian writers is 9.01 tokens.

As a result, the writer concludes that Indonesian speakers write more noun phrases whereas English native speakers write more prepositional phrases, and clauses. Even though many people sometimes underestimate of the competence of Indonesian writers; in this study the result shows that they might have the same ability with English native speakers in writing long sentences.

Conclusion

Here in the conclusion, the writer gives the results of the analysis based on the statements of the problems.

1. Do noun phrases written by English native speakers have more tokens that those written by Indonesian speakers?

The result shows that Indonesian speakers write more noun phrases than English native speakers. (English native speakers = 5.58 tokens; Indonesian writers = 6 tokens)

2. Do prepositional phrases written by English native speakers have more tokens than those written by Indonesian speakers?

The result shows English native speakers write more prepositional phrases than Indonesian writers. (English native speakers = 7.55 tokens; Indonesia speakers = 7.23 tokens)

3. Do clauses written by English native speakers have more tokens than those written by Indonesian speakers?

The result shows that English native speakers write more clauses than Indonesian writers. (English native speakers = 15.65 tokens; Indonesian writers = 15.54 tokens)

After analyzing all the journals, the writer concludes that Indonesian speakers write more noun phrases whereas English native speakers write more prepositional phrases and clauses. Many people sometimes underestimate the ability of Indonesian writers because English is only their foreign language and Indonesian people do not use English all the time in their daily activities. However, the result in this study shows that Indonesian writers might have the same ability as English native speakers in writing long sentences. That is why, the writer expects that Indonesian writers want to make more English written texts as many as they can.

Bibliography

- Allen, Harold. (1958). *Reading in Applied English Linguistics*. New York: Alfred. A. Knopf, Inc.
- Astika, Gusti. (2003). *A Discourse Approach to Teaching Structure in Proposal Writing*. English-EDU, Journal of Language Teaching and Research, 2(1), 13-32.
- Azar, Betty. (1989). *Understanding and Using English Grammar*. New Jersey: Prentice Hall, Inc.
- Azar, Betty. (1992). *Fundamentals of English Grammar*. New Jersey: Prentice Hall, Inc.
- Beebe, Leslie. (1988). *Issues in Second Language Acquisition*. Boston: Heinle & Heinle Publishers.

- Bulut, Sencer. (2002). *Counseling in Turkey*. Anima, Indonesian Psychological Journal, 17(3), 227-233.
- Francis, Nelson. (1958). *The Structure of American English*. Ronald Press Company.
- Hunston, Susan. (2002). *Corpora in Applied Linguistics*. Cambridge Press.
- Lutfi. (2002). *Board Structure, Managerial Ownership, and Shareholder Wealth: A Review of Empirical Evidence*. VENTURA, Journal Ekonomi Bisnis dan Akuntansi, 4(2), 32-40.
- Macmillan. (1965). *The Key to English*. New York: Collier Macmillan International Ltd.
- Maria, Agnes. (2002). *The Cognitive Model and Interventions for Specific Reading Disorder*. Anima, Indonesian Psychological Journal, 17(4), 311-317.
- Minor, James. (2002). *Incorporating Service Learning Into ESOL Programs*. TESOL JOURNAL, 11(4), 10-14.
- Monahan, Steven. (2002). *Discussion of the Value Relevance of Internet Firms: Does Reporting Grossed-up or Barter Revenue Make a Difference?* Journal of Accounting Research 40(2), 479-484.
- Quirk, Randolph. (1973). *A University Grammar of English*. London: Longman Group

