

Analysis of topic sentences of classification paragraphs and development of a diagnostic function

Kandai ISHIKAWA^{a*}, Hidenobu KUNICHIKA^b

^a*Graduate School of Computer Science and Systems Engineering,
Kyushu Institute of Technology, Japan*

^b*Faculty of Computer Science and Systems Engineering,
Kyushu Institute of Technology, Japan
[*k.ishikawa@minnie.ai.kyutech.ac.jp](mailto:k.ishikawa@minnie.ai.kyutech.ac.jp)*

Abstract: In a paragraph, the topic sentence is an important component that clearly expresses the author's argument and indicates the direction of the entire paragraph. It is difficult for beginners to learn English to write appropriate topic sentences. In order to solve the problem, we have focused on classification paragraphs and developed a diagnostic function of the topic sentences. As the result of an evaluation of the function, we have found that the function correctly diagnosed 70.6% of appropriate topic sentences taken from text books for paragraph writing.

Keywords: Support in English Language Learning, Paragraph Writing, Topic Sentence, Classification Paragraph

1. Introduction

Paragraph writing is one of writing techniques for expressing opinions and arguments. In a paragraph, the topic sentence is the most important component that clearly expresses the author's argument and indicates the direction of the entire paragraph. Moreover, a well-written topic sentence determines whether a paragraph can be written well. There are some conditions for writing topic sentences, and they differ according to types of paragraphs. In order to write paragraphs appropriately, it is necessary to understand such requirements, but beginners of English learning often lack sufficient knowledge. Therefore, we focus on supporting the writing of topic sentences.

There are some studies on assessing writing; e.g., (Su, Lin & Lai 2023). Although such systems will give learners feedback that is somewhat vague and abstract, they will not be able to generate concrete feedback according to types of paragraphs. In order to solve the problem, we are aiming to implement a diagnostic function of topic sentences as one of the functions of an English language paragraph writing support system. In this study, we focus on classification paragraphs, and have implemented a prototype of a diagnostic function of topic sentences (Ishikawa & Kunichika, 2023). This paper describes development and an evaluation of the diagnostic function.

2. Topic Sentences of Classification Paragraphs

Classification paragraphs classify things, that correspond to the topic of the paragraph, into several groups according to their characteristics or features. A topic sentence basically consists of two parts: a topic that indicates what the entire paragraph is about, and an assertion that expresses the author's argument or opinion on the topic. In this study, we have investigated the conditions of assertion in topic sentences of classified paragraphs by referring to textbooks on paragraph writing, and classified the conditions into three types. Also, we have categorized the forms of topic sentences into two types. The

three types of conditions differ in their applicability depending on the form of a topic sentence. The forms of topic sentences and the conditions of assertions are as follows.

(Form 1) The topic is appeared as the subject of a topic sentence.

The subject of a topic sentence corresponds to the topic, and is followed by a verb used for classification; e.g., "Popular music can be divided into three types."

(Form 2) An expletive is used as the subject of a topic sentence.

An expletive such as "there" is used as the subject of a topic sentence, and is followed by a be-verb; e.g., "There are three kinds of skiing."

(Condition 1) The verb of a topic sentence needs to mean classification.

This condition is applied to only Form 1. The topic sentence of a classification paragraph should indicate that the topic is classified into several groups. So, the verb of a topic sentence needs to mean classification. There are some frequent verbs such as "be divided into" and "fall into" (Kitao & Kitao, 1998).

(Condition 2) A criterion for classification needs to be provided.

A criterion for classification needs to be clearly stated in a topic sentence for classifying a topic based on the criterion in a paragraph. There are frequent words which express criteria; e.g., type, kind and category (Kitao & Kitao, 1998).

(Condition 3) The number of classified categories need to be indicated.

The details of each category are presented in supporting sentences of a paragraph.

Therefore, the number of classified categories needs to be indicated in a topic sentence (Donald et al., 1998). There are two ways to express the number of items: one is to indicate both the criteria for classification and the number of categories, e.g., "divided into three types", and the other is to list the name of each category, e.g., "classified as A and B."

3. Development of a Diagnostic Function of Topic Sentences

Our diagnostic function analyzes a topic sentence by referring to the result of word dependency analysis performed by CoreNLP (Stanford, 2022). Figure 1 shows the result of word dependency analysis of "Popular music can be divided into three types." Then, the function diagnoses if a topic sentence satisfies all the conditions. The diagnosis of topic sentences consists of two steps: identifying the form of a topic sentence and judging the conditions of the assertion of the topic sentence. First, for identifying the form of a topic sentence, the function confirms whether or not the main verb, which is shown as a root in the output from CoreNLP, is a be-verb and whether or not the subject is an expletive. Next, the function judges each condition of an assertion. For the judgment, the function extracts important words from the result of word dependency analysis performed by CoreNLP. Important words correspond to the main verb, a criterion for classification, and the number of classified categories. If the topic sentence is "Form 1", the function confirms whether or not the verb means classification by referring to a list of frequent verbs (e.g., Howard & Tracz, 1976; Donald et al., 1998) for condition 1. For condition 2, the function confirms whether or

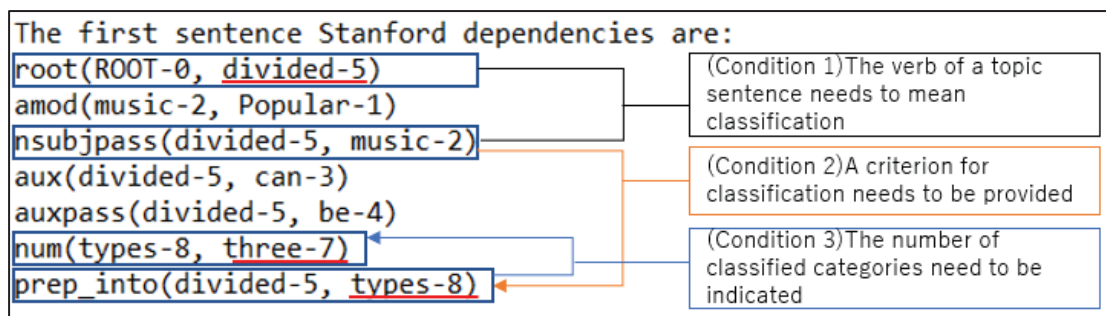


Figure 1. An Example of the Result of Word Dependency Analysis performed by CoreNLP

not a word corresponding to the object of the topic sentence can be used for a criterion of classification by referring to a list of frequent words which express criteria (e.g., Howard & Tracz, 1976; Donald et al., 1998). For condition 3, the function confirms whether or not there are numeral modifiers for the words which express criteria or whether or not the names of categories are listed. Finally, the function outputs the result of the diagnosis.

When an input sentence satisfies all the conditions, the function shows the format of the topic sentence, the verb which means classification, the criteria for the classification, and the number of classified categories. Figure 2 shows an example of the result of an acceptable sentence. If an input sentence does not satisfy the conditions, the function shows the conditions that are not satisfied.

```
[INPUT]
Popular music can be divided into three types.

[DIAGNOSTIC RESULT]
•The topic sentence entered is appropriate
  Topic Sentence Form: Topic as the subject

Verbs that imply classification      : divide
Criteria for classification         : type
Number of classification            : three
```

Figure 2. An Example of the Result of an Acceptable Sentence (Ishikawa & Kunichika, 2023)

4. An Evaluation of the Diagnostic Function of Topic Sentences

As an evaluation, we prepared 17 topic sentences of classified paragraphs taken from eight textbooks (e.g., Howard & Tracz, 1976; Donald et al., 1998; Oshima & Hogue 2017) on paragraph writing, and used the sentences as input for the diagnostic function. As the result, the system correctly diagnosed 12 (70.6%) of the 17 sentences as appropriate, while the other 5 sentences were judged as unacceptable sentences. We have found two patterns of sentences that were misdiagnosed: one is omitting the criteria for classification and the other is listing categories in clauses or sentences connected by a conjunction.

5. Conclusion

In this study, we have developed a diagnostic function of the topic sentences for classification paragraphs. Also, this paper has presented an evaluation of the function. As the result of the evaluation, we have found that the function correctly diagnosed 70.6% of appropriate topic sentences taken from textbooks for paragraph writing. In the future, we plan to evaluate the function by using sentences including inappropriate topic sentences.

References

- Donald, B. R., Moore, D. J., Moorow, R. B., Wargetz, L. G., & Werner, K., (1998) *Writing Clear Paragraphs* (6th Edition). Prentice-Hall.
- Howard, C. J., & Tracz, R. F., (1976). *Writing Effective Paragraphs*. Winthrop Publishers.
- Ishikawa, K., & Kunichika, H., (2023). Development of a Diagnostic Function of the Topic Sentence for a Classified Paragraph. *Annual Student Meeting 2023*, Japanese Society for Information and Systems in Education. Retrieved March 7, 2023, from https://www.jsise.org/wp-content/uploads/2023/01/2022_kyusyu_okinawa_a03.pdf (In Japanese)
- Kitao, S. K., & Kitao K, (1998) *Basic English Paragraphs -Improving Reading and Writing Skills-* Eichosha Co., Ltd.
- Oshima, A., & Hogue, A., (2017) *Longman Academic Writing Series* (3rd Edition). Pearson Education Inc.
- Stanford NLP Group, (2022) Overview – CoreNLP. <https://stanfordnlp.github.io/CoreNLP/>
- Su, Y., Lin, Y., Lai, C., (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57(2). <https://doi.org/10.1016/j.asw.2023.100752>