

Conference Paper Review

Paper ID: 002

Title: Comparing Annotation Guidelines and Annotators: A Preliminary Investigation into the Lexical Complexities Assigned in Two Complex Word Identification Datasets

Author(s): Kai North

Reviewer's Recommendations:

i. Writing

- Minor improvement suggested

ii. Novelty

- Original

iii. Suitability

- Very related

iv. Reviewer's Expertise

- Knowledgeable

v. Recommendation

- Absolute accept

Reviewer's detailed comments: strength, weakness, suggestion

i. Comments for the Authors:

a. Strengths

The paper's objective to compare the effects of annotation guidelines on the quality of complex word identification datasets is made well clear in the abstract. Likewise, the paper's significance has been well stated by giving examples of some application areas like text simplification and lexical simplification.

The related works have been explained very well. It includes the current progress and the changes in methodology due to the introduction of Lexical Complexity Prediction.

The two different annotation tasks and the two datasets and processes involved have been explained and differentiated well. I also liked the structure of introducing the two well-defined hypothesis and validating them through the results later. The results are well presented, as well.

The conclusion covers all the key contributions of the paper. Likewise, the impact and significance of the research have been well stated in the future recommendations section.

b. Weaknesses and Suggestions

I found a few grammatical errors in the paper. For example, the first sentence in the abstract, no space after a period in section 4.1 second sentence, spelling error in "instance" in section 6.3, etc.

Although the disadvantages of annotators' choice have been clearly stated in section 3.1, the advantages are missing. Likewise, removing people not doing their tasks properly has been proposed to eliminate financial incentives in the annotation. However, it is not mentioned clearly how such people are selected.

A suggestion is to make the research question proposed measurable in terms of quantitative or qualitative measurements, if possible.

The formula for Keynes contains the term $(O_k - E_k)^2$. However, the explanation mentions that "the observed frequency values of each complex word: 'O' were then subtracted from their expected values: 'E'". Although the results will be the same, the sentence can be made consistent with the formula.

A concern is that only a sample of words was used to validate the hypothesis, which might not be accurate.

One suggestion is to shorten the title if possible. Although the title is well descriptive, I felt that the title is too long.

ii. Comments for the Program Committee (will be kept confidential and NOT released to the authors)

The paper presents the effects of the annotation guidelines on the quality of complex word datasets quite well. It has possible impact in creation of similar datasets in the future, which can be used in various application areas.

Journal Paper Review

Paper ID: 002

Title: Comparing Annotation Guidelines and Annotators: A Preliminary Investigation into the Lexical Complexities Assigned in Two Complex Word Identification Datasets

Author(s): Kai North

Goals and Contributions:

- i Do the authors clearly state the research goals of the work?

The paper's objective to compare the effects of annotation guidelines on the quality of complex word identification datasets is made well clear in the abstract.

- ii Does the paper clearly indicate what the contributions are?

The paper clearly indicates the contributions in improving the future word complexity dataset by analyzing the effects of annotation guidelines on its quality.

- iii Are the claimed contributions original and significant in terms of

- Novel methodology?
- New applications?

The contributions are original and significant in terms of a new methodology to create effective word complexity datasets for existing application areas like text simplification and lexical simplification.

- iv Does the paper describe the methods in sufficient detail for readers to replicate the work?

The explanations of the methods and datasets are well written in detail, enough for readers to replicate the work. However, the information about the code has not been mentioned in the paper.

Evaluation:

- i. Do the authors carefully evaluate the approach?

The approach has been carefully evaluated by comparing the two experiments and generating a common shared dataset in both methods.

- ii. Does the paper include systematic experiments, a careful theoretical analysis, or give evidence of generality?

Yes, the paper includes a very systematic experiment, conducted by introducing the hypothesis and conducting two different experiments corresponding to two different annotation tasks and datasets. The results obtained were then used to validate the hypothesis. However, the dataset was minimal to reach a conclusion.

Discussion:

- i. Does the paper discuss relevant earlier works, noting similarities, differences, and progress?

The related works have been explained very well. It includes the current progress and the changes in methodology due to the introduction of Lexical Complexity Prediction.

- ii. Does it discuss the limitation of the approach as well as its advantages?

Yes, both the limitations and the advantages of the approach have been mentioned in the paper. Likewise, methods are proposed to overcome these limitations.

- iii. Does it consider the implication of the work and outline direction for future work?

Yes, it considers the work implication and outlines the direction for future work in creating a more effective dataset for word complexity-related tasks.

Presentation:

- i. Is the paper adequately organized and well written?

Yes, the paper is properly organized in terms of structure. It has been divided into standard sections. Also, the paper is well written, with minor improvements required in a few sections.

- ii. Is the paper grammatically correct and free of spelling errors?

I found a few grammatical errors in the paper. For example, the first sentence in the abstract, no space after a period in section 4.1 second sentence, spelling error in "instance" in section 6.3, etc.

- iii. Does it use standard terminology?

Yes, all the terminology used in the paper were standard.

Detailed Comments:

The paper's objective to compare the effects of annotation guidelines on the quality of complex word identification datasets is made well clear in the abstract. Likewise, the paper's significance has been well stated by giving examples of some application areas like text simplification and lexical simplification.

The two different annotation tasks and the two datasets and processes involved have been explained and differentiated well. I also liked the structure of introducing the two well-defined hypotheses and validating them through the results later. The results are well presented, as well.

Although the disadvantages of annotators' choice have been clearly stated in section 3.1, the advantages are missing. Likewise, removing people not doing their tasks properly has been proposed to eliminate financial incentives in the annotation. However, it is not mentioned clearly how such people are selected.

A suggestion is to make the research question proposed measurable in terms of quantitative or qualitative measurements, if possible. One concern is that only a sample of words was used to validate the hypothesis, which might not be accurate.

The formula for Keyness contains the term $(O_k - E_k)^2$. However, the explanation mentions that "the observed frequency values of each complex word: 'O' were then subtracted from their expected values: 'E'." Although the results will be the same, the sentence can be made consistent with the formula.

One suggestion is to shorten the title if possible. Although the title is well descriptive, I felt that the title is too long. The conclusion covers all the critical contributions of the paper. Likewise, the impact and significance of the research have been well stated in the future recommendations section.

Recommendation:

The paper could be published after minor revision, with no further review required.