Topic-based Revision Tool to Support Academic Writing Skill for Research Students

Harriet Nyanchama Ocharo
School of Information Science
Japan Advanced Institute of Science and Technology
Nomi City, Japan
Email:harriet.ocharo@jaist.ac.jp

Shinobu Hasegawa
Research Center for Advanced Computing Infrastructure
Japan Advanced Institute of Science and Technology
Nomi City, Japan
Email:hasegawa@jaist.ac.jp

Kiyoaki Shirai
School of Information Science
Japan Advanced Institute of Science and Technology
Nomi City, Japan
Email:kshirai@jaist.ac.jp

***Abstract*— When students write academic articles, they undergo a revision process where they receive feedback in the form of comments from their supervisors to improve the quality of the articles. The comments can be broadly classified into three categories: grammatical comments, format-related comments and topic-related comments. Comments related to the topic of the research are the hardest to resolve because students may lack discipline-specific writing skills needed to resolve such comments. This research developed an interactive tool to enable students search an archive of previous students' articles showing the revision history and comments. A machine learning approach was used to automatically classify the comments in the database into the three categories so that only topic-related comments were brought up in the search result. The result of the search was presented to the student in a way that clearly showed the process previous students used to resolve related comments, thereby showing them a similar way they could use to resolve any difficult topic-related comments. As the student's writing skill level increases, the amount of detail presented to the student reduces so as to avoid over-reliance on the tool. In this way, students could improve their academic writing skills.**

Keywords- research support system; academic writing; writing tools; writing skill.

I. INTRODUCTION

Students in higher education and other researchers measure their achievements through the number of quality research articles they publish. Quality writing is therefore important in research in order to convey ideas clearly. It is the final stage of research and a culmination of effort that deserves to be done properly. However, students sometimes face difficulties revising their articles due to various reasons, such as lack of understanding, focusing too much on the formatting rather than the content and an inability to estimate the time it takes to revise an article because of lack of experience in academic writing [1]. In other words, they may not have sufficient academic writing skill.

Academic writing skill is the ability to write logically organized research papers, essays or reports in a well-structured, concise format. It is the ability to present complex ideas objectively while following the academic writing style, such as writing in third-person style, passive writing, proper citations etc. General academic writing skill may be taught in formal language lectures. However, there are variations in the structure and style of research papers in different research fields. Young researchers therefore feel the need to acquire discipline-specific writing style or skill from previous articles by researchers in the same field [2]. The challenge is that usually, the articles they read and learn from are in the final version. If they face a problem during revision of their own articles, they have no way of knowing how the previous students went through the revision process.

This research therefore proposes a support system for revision of articles based on a revision history database. When students write articles, they have to go through a revision process to improve the drafts based on comments from their supervisors. However, students may lack discipline-specific writing and revision skills needed to resolve such comments. We built an archive of previous students' articles and the corresponding comments they received when they were revising their articles. With a revision history database, students can learn revision skills by looking up similar or related comments and see how the other students resolved their comments.

The comments can be broadly classified into three categories: grammatical comments, format-related comments and topic-related comments. There are many tools, commercial or otherwise, to check the grammatical structure of documents. Format-related comments are also easy to resolve by following some standard guidelines set out for academic papers. However, there is not much research into tools or interfaces to help students resolve topic-related comments and that is why our focus is on comments related to content and meaning – where students take the longest time during revision. A machine learning approach was used to automatically classify the comments in the database into the three categories, so that only topic-related comments appeared when students looked up the comments database. The student would then be able to easily focus on resolving the topic-related comments. The result of the archives database search was presented to the student in a way that clearly showed the changes in subsequent drafts, thereby showing them a similar process they could use to resolve any difficult topic-related comments. If students resolve topic-

 ISBN: 978-1-61208-538-8

related comments quickly, then the duration of the revision process is shortened.

As the student's writing skill level increases, the amount of detail presented to the student reduces so as to avoid over-reliance on the tool. To improve academic writing skill, there is a need to reduce the cognitive support for the students as their level of skill increases. This approach is called fading, where the functions of the supporting tool can be *fadable* according to the student's meta-cognitive skill [3]. This raises the issue of measuring the student's skill level, which we can estimate by the number of comments raised in each article the student is revising.

The rest of this paper is constructed as follows: section II is a review of related work with a view of identifying the research gap in academic writing and the potential impact of this research. In section III, our approach is discussed in detail covering the writing and revision process, the data collection procedure used to gather previous students' articles and the technical details related to the process of automatic classification of the comments. Section IV includes the results and discussion of the comment classification, and section V is the conclusion and future work.

II. RELATED WORK

With the advent of the use of computers in learning, there has been an increase in research on writing tools to aid academic writing. Earlier research into important linguistic aspects of a good writing style such as readability, sentence and word length, sentence type, word usage and sentence openers [4] enhanced the capability of word processors beyond mere spellchecking. In addition to word processors, grammar checking tools are available that can automatically recognize and clean up grammatical errors in writing [5].While these grammatical tools are beneficial in helping researchers clean up errors in their writing, the quality of writing cannot be evaluated by grammatical accuracy alone [6].

This therefore raises the question of whether these tools can also be useful in improving students' competency in academic writing. Students can of course learn directly from language teachers, but research students are often pressed for time and are likely to end up copying from bibliography, or working in a relationship of informal apprenticeship with more experienced members of their team [2].

Online interactive tools offer a promising way for students to improve their grammar skills. A corpus is one way for novice students to learn from experienced researchers. Narita [6] states that a corpus-based tool of previous students' work can be vital for improving second language learners' grammatical knowledge. Aluisio [7] proposed a design for a tool that made explicit the writing skills performed by language expert authors so that novice researchers could develop their academic drafting and revision skills in a foreign language. Alusio [8] further developed a tool to assist non-native novice researchers in achieving a cohesive schematic structure for their articles.

In their research, Hasegawa and Yemane [1] created an article revising support system that facilitates article revision by managing all the comments as tickets, such as in an issue-tracking system. However, the comments are not classified by categories, as is the case in this research where the focus is on topic-related comments and how to solve them.

Once a research student has written an initial draft, he/she will receive feedback from their supervisor to improve the draft. These comments may not only be related to their grammatical errors, but also to the format or structure of the paper. A third type of comments are those related to the topic of research. As described previously, there are a lot of tools to help students improve their grammatical knowledge as well as the structure of their scientific articles. However, there has not been much research into helping students improve their revision skills.

This paper expands the scope of previous research by presenting a way for researchers to improve their topic-related revision skills and hence resolve comments relating to the content of their research articles.

III. OUR APPROACH

In this section, the process a student goes through when writing and revising an academic article is discussed in detail. We also discuss how our revision tool can help the student to shorten the revision process by automatically filtering out non topic-related comments. In addition, we discuss the process of collecting the necessary data and developing the tool.

A. The Writing and Revision Process

After a student writes the first draft of a research article, he/she sends it to the supervisor for feedback. The supervisor inserts comments to help the student improve the draft and sends it back to the student. The student then revises the draft based on the comments and sends it back to the supervisor for feedback and so on, until the final draft is approved. This is illustrated in Figure1.

The main objective of this tool is to increase the efficiency of the revision process resulting from a reduction in the number of comments and number of drafts, and a shorter duration of the revision process. Increased efficiency implies that the writing skill level of the student has improved as they are able to revise their articles much faster. When the student uploads the draft with comments into the tool, the comments are automatically labelled as grammatical, format or topic-related. If the student were to do this manually, it would take too much time. The student can then quickly revise the format and grammatical comments before focusing on the topic-related comments. He/she can look up in the archive for similar comments and see how previous students resolved their comments.

 ISBN: 978-1-61208-538-8

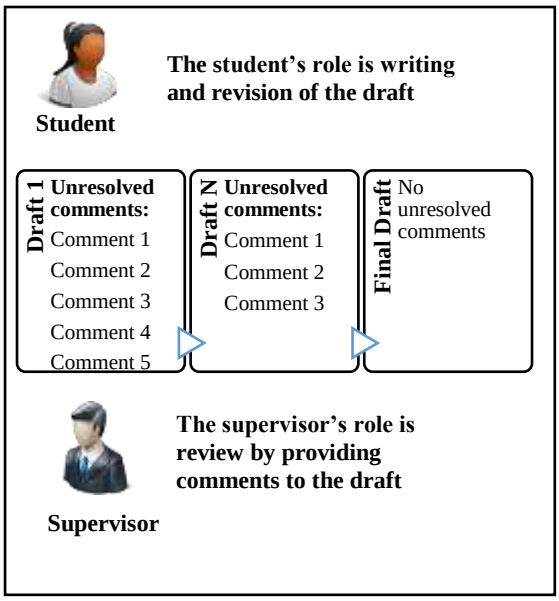

Figure 1. The writing and revision process

B. Data Collection Procedure

The revision history of articles from previous students in the same laboratory was collected. 19 articles were obtained with an average of 6.6 drafts (126 drafts in total). Each draft had an average of 20 comments. The total number of comments was 1,338.

The comments and corresponding comment ranges were extracted from the original Microsoft Word documents and uploaded to a MySQL database. This formed the backend for the web-interface that was developed using Django, a Python framework. The web interface was used to search for matching comments and for viewing a history of the revision process for the documents containing similar comments.

C. Comments Classification Process

The comments were classified into 3 types – format, grammatical, and topic-related. The comments were classified as per the below definition:

- **Format**: comments about font type and size, positioning of figures and tables, page limitations etc. *Example: Change font style for section title*
- **Grammatical**: comments about correction of grammatical or spelling mistakes etc. *Example: Is paragraph structure OK? (Main topic). However, the contrary situation?- On the other hand?*
- **Topic-related**: comments about actual content or topic of the article *Example: I cannot catch the goal of community based learning from this document. What types of knowledge and skill do the community members have through the CBL?*

If a comment belonged to both topic-related and any other category, then that comment was labeled as being topic-related in the training model because it was considered useful to the revision process.

Natural language processing and machine learning techniques were used to automatically classify comments into the three categories. Each comment was analyzed and annotated using the Stanford CoreNLP [9]. The annotation included tokenization, sentence splitting, POS tagging and lemmatization. The machine learning algorithm LIBSVM [10] was used for training because of its simplicity in rapidly obtaining acceptable results, even with texts short in length (the average length of the comments was 90 words). For simplicity reasons, this classification test used only the lemmas of the comment words as features. The comment words included were nouns, verbs, adjectives and adverbs. Pronouns, articles and other parts of speech were not considered relevant features for the classification algorithm.

IV. Results and Discussion

This section details the results of using machine learning to automatically classify comments into the three categories discussed earlier, and a discussion of the implications of the results as well as the practical application of the revision tool in a laboratory setting.

A. Classification Results

After obtaining the features, the number of unique comments was down to 612. The number of features considered was 902. Each comment was manually assigned a label as SVM is a supervised learning method. There was an imbalance in the distribution of the target classes with topic-based comments accounting for 56% of the total number of comments. Grammatical and format comments accounted for 25% and 19% respectively. Stratified k-fold validation was applied, with k=10. Data was split into 10 groups, with each group containing 552 (137 grammar, 105 format, 310 topic-related) training data items and 60 (15 grammar, 11 format, 34 topic-related) testing data items. The ratio of relative class frequencies was approximately preserved in each training and testing fold. An average prediction accuracy rate of 56.21 was obtained using the LIBSVM tool (default parameters).

B. Case Study:

In order to observe the usability of the tool, we obtained a student's article with reviewer comments. There are two stages:

Stage 1: The student uploads the article and the comments are automatically extracted and classified. There were 15 comments, which were automatically extracted and classified as grammatical, format or topic-related. In this case, all the 612 unique comments were used as training data and the resulting model was used to predict the category of the 15 comments. 13 of the 15 comments were accurately predicted as being topic related, having an accuracy rate of 86.67%. In Figure 2, the topic-related comments are presented to the student.

 ISBN: 978-1-61208-538-8

Figure 2. Results of the classification of the case study article presented in the web interface

Stage 2 is looking up the comment in the database of previous work to see how other students resolved similar comments. In this case, clicking "look up" on the first comment presents a list articles whose results closely match the key word "skills". The first step in the search is to look for specific content-related keywords in the comment. The reason keyword searches are used is because the comments are stored in a relational database as string fields. Therefore, the search is a simple database lookup. Only the papers containing comments in the database marked as topic-related by the SVM algorithm were brought up in the lookup results. The search results are shown in Table I.

Selecting one of paper titles presented "HCII2016_Ocharo", the revision history of the article i.e. the changes the phrase in question had gone through various draft versions, was presented as in Table II. From this result, they may notice how to define technical terms in research and also how to focus and narrow down the focus of their research.

In summary, there were positive search results for other comments containing key topic-related words such as "testing, analysis, abstract, methodology". The results of the matching comments and revision history of corresponding articles were displayed as expected. In future, the system will be tested by the target students to evaluate its actual effectiveness in reducing the average number of drafts and comments, thus reducing the time it takes to revise academic articles.

TABLE I. THE SEARCH RESULTS IN TABLE FORM

Paper Title	Version no	Matching Comment
GLS2014_Didin	2	These sentences are similar to the ones in Abstract?. Basically, it is OK. But you can add some examples. ?such as, decision-making, team-working, and communication skills?
SIG-ALST2012_Didin	2	Why should the volunteers improve their skills independently?
HCII2016_Ocharo	3	Can you describe a couple of examples of the research skills?
HCII2016_Ocharo	3	What is the research skills in this context? Maybe you define it at section 2 or later. But it would be better if you explain a simple example of the skill in this section so that the audience can easily understand the concept.
ICCE2013_Didin	4	Magnitude which enables the novice volunteers to develop their ethical decision-making skills at all times during official disaster management training inside and outside of class, and expect them to improve their performance in disaster response activities.?

C. *Discussion*

Considering the prediction accuracy of the classification algorithm, more comment data is needed. In machine learning, a large amount of data is needed in order to improve the accuracy of the prediction algorithm but in this case, there was only an initial number of 1,338 comments. A lot of the grammatical and format related comments were misclassified as being topic-related. Even with the limited data set, the factors below could have affected the accuracy.

- Even after using stratified k-fold cross validation, that more than half of the comments were topic-related could have introduced bias in the training model. In addition, the large number of features (902) relative to the data set (612) may have also had an impact on the performance of the prediction model.
- In addition, a single model with three outputs (grammar, format, topic-related) was trained and used for prediction. Instead, three different models each predicting whether a comment belonged to the group or not, may improve accuracy.

 ISBN: 978-1-61208-538-8

TABLE II. COMMENT REVISION HISTORY

Paper Title: HCII2016_Ocharo	**Version 1**	**Version 2**	**Version 3**
Matching comment	But it would be better if you explain a simple example of the skill in this section so that the audience can easily understand the concept	Can you describe a couple of examples of the research skills?	Describe specific examples of the research skills?
Comment range	Research skills	Research skills can be widely categorized into two: discipline specific and general research skills.	Research skills include such generic skills as planning and scheduling, communication and presentation; and specific skills such as trend analysis, problem definition and data analysis

- Thirdly, the classification algorithm only used the content words (nouns, adjectives, verbs and adverbs) of the comments without considering their meaning or context. For example, topic-related comments contain keywords such as 'abstract', 'originality', 'design', 'develop', 'usability' etc. Topic-related comments also apply to certain sections of the article, such as the title, section headers etc. Grammatical comments contained keywords or phrases such as 'redundant', 'sentences', 'misspell' etc. while format comments typically contained keywords such as 'font', 'Calibri', 'style', 'move figure', 'change order' etc. Including other information like the document version, author, and comment metadata such as comment author, comment replies etc. may improve the prediction accuracy.
- Fourthly, the comment range – the words in the document that are covered by the comment – was not considered either and this combined with the comment text could also improve the accuracy of classification.
- Fifthly, the comment length was not considered for convenience purposes as it would require scaling between a range of (0,1). However, it is an important feature to consider as topic-related comments were typically longer than any other, while grammatical comments' length could be as short as a single word.
- Lastly, the kernel and parameter selection of the SVM algorithm also affect the accuracy of the results. In this research, the default parameters were applied. It would require several trials to discover the best kernel-parameter combination to produce the highest accuracy. However, in this research, we focused on rapidly obtaining acceptable results.

D. Practical Application in Laboratory Setting

The revision tool can be applied in a laboratory setting so that the student can look up similar comments in the archive of previous students, since the results are more likely to be relevant if all the students belong to the same laboratory. For the search results to be more useful, the database of revision histories also needs to be large enough to allow more informative searches. The current database is limited to only 19 articles. However, it is difficult to build a large database as even with an average of 10 students publishing 3 times each year, that would only amount to 30 articles in a year. Therefore, finding ways of improving accuracy even with a limited data set is the most important element of future research. Students could help with the annotation, by manually correcting misclassified comments and this feedback would in turn be used to improve future prediction.

When it comes to research involving comments, other factors to be considered include whether or not the system will manage comments or leave it to an external application such as text processors. In our case, the tool only provides look up but not comment management. Furthermore, metadata contained in the comments such as author, date, comment replies, comment status (open or closed) could be useful data for students carrying out revision of their academic articles. Some comments are also persistent throughout the revision process, which could mean they are harder to solve, while others occur more frequently. Such an analysis could be combined with search results to push the most relevant comments to the top of the search result.

The student's skill level should be taken into account when presenting the student with the search results. In other words, the system should *adapt* to the student's skill level by presenting a more detailed version of the results to students with low skills, while presenting a less detailed version to students with higher skills. As the student's writing skill level increases throughout the revision process, the amount of detail presented to the student reduces so as to avoid over-reliance on the tool. This would avoid automated writing, which impends student learning.

The student's skill level can be estimated by the number of comments raised in their drafts by their supervisor. More comments show that the student has a lot of revision points to consider, which could mean that the student is lowly skilled. Fewer comments imply the student is highly skilled. This estimate of skill is calculated with each draft to ensure adaptation to the skill level of the student.

 ISBN: 978-1-61208-538-8

V. CONCLUSION AND FUTURE WORK

This paper described an article revision tool that helps students resolve difficult topic-related comments they encounter during the writing process by looking up the revision process of previous students in an archived database. In this way, the students can improve on their knowledge of academic writing. In future, the efficiency of this tool in improving the writing skills of students will be evaluated by the target students in the same laboratory. In this research, the comments were simply classified into grammar, format and content-related. In future, other types of comments such as comments related to the logical structure of the documents will be considered. In addition, if there are many similar corrections in grammatical errors, it should be shared as pre-requisites for paper writing. Such a summarization function would be useful for novices.

Academic writing is an integral part of research in universities and other institutions of higher education, and as such, any computer tools to aid this process can have a significant impact on the quality of output from such institutions. Future work will focus on evaluating the impact of the revision tool discussed in this paper.

REFERENCES

[1] S. Hasegawa and K. Yamane, "An Article/Presentation Revising Support System for Transferring Laboratory Knowledge," 19th International Conference on Computers in Education, Chiang Mai, Thailand, pp.247-254, 2011.

[2] K. Hyland, Disciplinary Discourses: Social Interactions in Academic Writing, Harlow: Person Education, 2000.

[3] S. Hasegawa, K. Mannari, S. Ooiwa, and A. Kashihara, "A Portal Site for Supporting Research Activities," 15th International Conference on Computers in Education, Hiroshima, Japan, pp. 59-60, 2007.

[4] L. Cherry, "Writing Tools," IEEE Transactions on Communications, vol. Volume: 30, no. Issue: 1, pp. 100-105, 1982 .

[5] R. Blake, "Current trends in online language learning," Annual Review of Applied Linguistics, *31*, pp. 19-35, 2011.

[6] M. Narita, "Developing a corpus-based online grammar tutorial prototype," The Language Teacher, pp. 23-29, 2012.

[7] S. M. Aluisio, I. Barcelos, J. Sampaio and N. O. Oliviera Jr., "How to learn the many unwritten 'rules of the game' of the academic discourse: a hybrid approach based on critiques and cases to support scientific writing," Proceedings IEEE International Conference on Advanced Learning Technologies, Madison, WI, USA, pp.257-260, 2001.

[8] S. M. Aluisio and R. E. Gantenbein, "Towards the Application of Systematic Functional Linguistics in Writing Tools," Computers and their applications: Proceedings of the ISCA 12th international conference, Tempe, Arizona, US, pp. 181-185, 1997.

[9] C. D. Manning, et al., "The Stanford CoreNLP Natural Language Processing Toolkit," Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60, 2014. http://stanfordnlp.github.io/CoreNLP/index.html [Retrieved: Aug, 2016]

[10] C. Chih-Chung and L. Chih-Jen, "LIBSVM : a library for support vector machines," ACM Transactions on Intelligent Systems and Technology, p. 27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm [Retrieved: Aug, 2016]

 ISBN: 978-1-61208-538-8