Case Study of the Use of the Interactive Annotation Software Perusall With Mixed-Proficiency EFL Japanese Students

Johnny George, Meiji University, Japan

The IAFOR International Conference on Arts & Humanities in Hawaii 2024
Official Conference Proceedings

Abstract
This study reports on the use of Perusall in a basic university English course in order to understand how such software facilitates second language acquisition, particularly in mixed-proficiency language groups. This study focuses on the effect of text choice on student performance. The question centers on how the reading difficulty of a particular text affects the type and quantity of responses produced. Sixty students enrolled in a freshman level English course used Perusall over the course of a year to read and annotate English media assignments. The students created questions, comments and responses to other students. The student input was taken and categorized based on the reading level of the assignments cross-referenced with the type and length of annotations produced. The reading content affected the type of student responses; longer articles tended to elicit more questions, while shorter content would elicit more comments. This work shows that varying language prompts can meaningfully influence user interaction.

Keywords: TESOL, Reading, Perusall

Introduction

This study reports on the use of Perusall, an interactive annotation software, in a basic university English course to understand how such software facilitates second language acquisition, particularly in mixed-proficiency language groups. This study focuses on the effect of text choice on student performance. The question centers on how the reading difficulty of a particular text affects the type and quantity of responses produced. This study contributes to the general understanding of second-language user performance in an online social learning environment. This work shows, as demonstrated in earlier research (Cecchinato et al. 2020, Tian 2020, Mark 2001), that a collaborative interactive environment such as *Perusall* facilitates peer-to-peer language learning support by successfully leveraging the collective language knowledge of students and making it accessible in an engaging way. The reading content did appear to affect the type of student responses; more difficult articles tended to elicit proportionately more questions, while shorter content would elicit more comments. This work shows that the nature of the language prompts can meaningfully influence the user interaction strategies.

Perusall and Student Interaction

Perusall facilitates classroom interaction by allowing students to watch or read media along with making annotations available for response from their peers. Perusall is an online social media annotation platform that incorporates an AI-based evaluation feature that assigns each student a grade based on interaction with the given media and quality of annotations produced in response to the text. Prior work shows that Perusall can effectively facilitate student interaction and feedback that allows them to better understand, analyze and critically evaluate at text (Cui and Wang 2023, Hanc et al. 2023, Porter 2022, Cecchinato and Foschi 2020, Tian 2020). With respect to the AI evaluation component, students have expressed concerns about the ability of AI to adequately evaluate their work and the potential of manipulating the system to inflate grades; however, one survey finds that students' confidence for the system increases when the instructor combines the AI evaluation process with instructor direct evaluation metrics (Cecchinato and Foschi 2020).

Text Difficulty and Student Performance

The current study centers on the use of Perusall by second language learners of English, especially with respect to increasing student interaction in English and collectivizing their efforts to comprehend a text largely independent of direct instructor guidance. The length of annotations created by students can serve as a measure of English interaction. The reading level of the text would be expected to meaningfully affect the length and type of the student responses. For instance, students can ask each other questions about a text, respond to such questions or even make comments about the text. Perhaps a more difficult text may encourage more back and forth student engagement to understand the text, or alternatively, students stymied by a difficult passage may annotate less.

Method and Results

Sixty students enrolled in a freshman level English course used Perusall over the course of a year to read and annotate four English media assignments. The students created questions, comments and responses to other students. The student input was taken and categorized

based on the reading level of the assignments cross-referenced with the type and length of annotations produced.

The four assignments consisted of three articles and one video with English subtitles. Table 1 below lists the material in order from the easiest to read to the most difficult reading, which has the lowest Flesch Reading Ease score as determined by Word.

Title	Author	Source	Flesch Reading Ease	Flesch-Kincaid Grade Level	Word count
I just don't handle stress well (stand-up comedy video/text)	Ron Funches	The Tonight Show w/ Jimmy Fallon (2020)	82.7	5.6	223
A son echoes his father's questions about identity in Japan	Richard Solomon	The Japan Times (2017)	48.6	10.3	1517
The Brilliance and Weirdness of ChatGPT	Kevin Roose	New York Times (2022)	48.5	11.7	1222
On "Exotic" Parenting	Bhakti Shringarpure	Huffington Post (2012)	47.2	11.6	1565

Table 1: Media assignments

The Flesch Reading Ease score was then tested for its correlation to the average number of words written per student and the proportion of questions to comments as shown in Table 2; both metrics serve as proxies for English engagement.

Title	Author	Flesch Reading Ease	Avg # of words	Avg # of annotations per student	Percentage of annotations questions
I just don't handle stress well (stand-up comedy video + text)	Ron Funches	82.7	144.8388	10.0555	26%
A son echoes his father's questions about identity in Japan	Richard Solomon	48.6	128.5234	9.02885	36%
The Brilliance and Weirdness of ChatGPT	Kevin Roose	48.5	119.6357	10.7299	30%
On "Exotic" Parenting	Bhakti Shringarpure	47.2	115.007	9.24815	33%

Table 2: Flesch Reading Ease & Annotation metrics

The results in Figure 1 below show that the average number of words written per student increases as the ease of reading the text increases. However, the average number of annotations produced per student does not correlate with the difficulty of a given text. The proportion of questions asked generally correlates with an increase in the reading difficulty of the texts as shown in Figure 2.

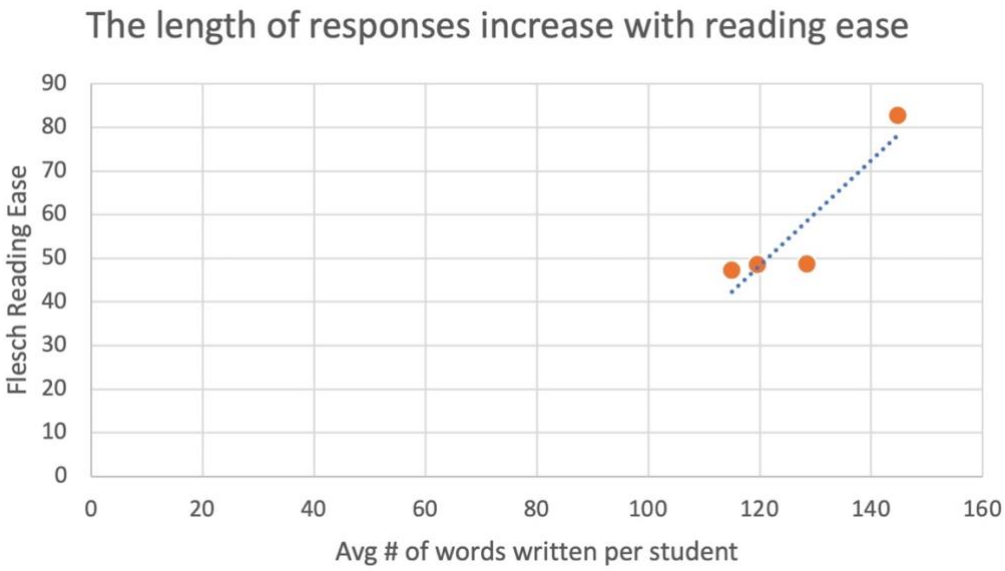

Figure 1: Reading ease correlation to response length

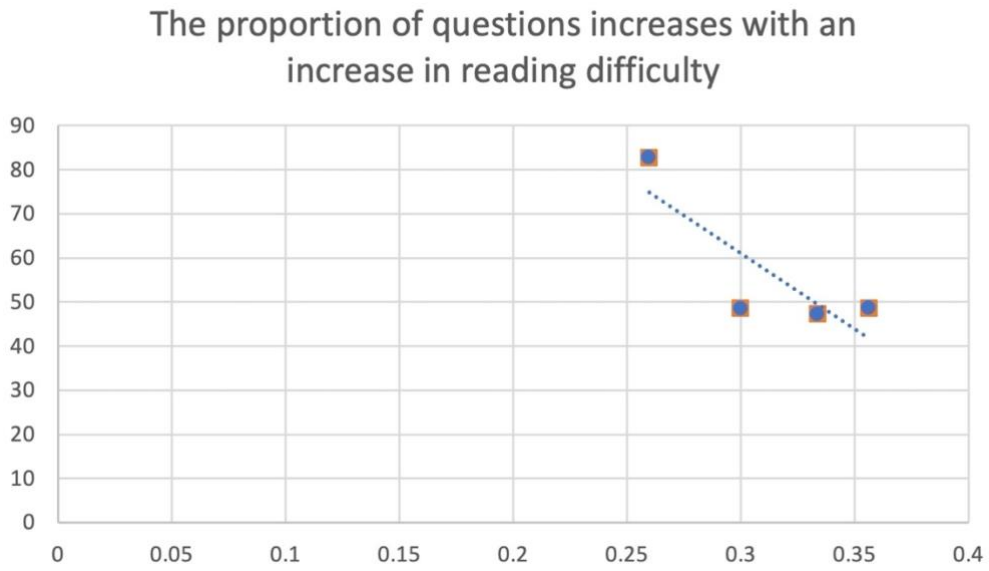

Figure 2: Reading ease correlation to proportion of questions

Conclusions

The effects of text difficulty on reader annotation creation comes as no surprise. The decrease in annotation length with respect to difficulty signals that the students simply tended to have less time to write in response to a difficult assignment. However, the increase in proportion of questions with respect to difficulty reflects the need for students to confer more to work out the meaning of parts of a given text. As pointed out by a conference attendee, the genre of a text served as another factor, as the entertainment-oriented Ron Funches' stand-up comedy increased student engagement, confounding the influence of ease of text versus genre.

Ultimately, the results are rather inconclusive due to the limited date set. There are few data-points and the distribution of difficulty is very small, with only one real outlier with respect to ease of reading, the Ron Funches' stand-up. A larger sample of responses and texts would facilitate a more robust result. Preliminarily we can say that the reading ease and

nature of the text will have some type of effect on the student production of annotations. Future research would include other metrics to compare such as student English proficiency and text engagement.

References

Bakermans, M., Pfeifer, G., San Martín, W., & LeChasseur, K. (2022). Who writes and who responds? Gender and race-based differences in open annotations. Journal for Multicultural Education, 16(5), 508–521. https://doi.org/10.1108/JME-12-2021-0232

Cecchinato, G., & Foschi, L. C. (2020). Perusall: University learning-teaching innovation employing social annotation and machine learning. Qwerty. Open and Interdisciplinary Journal of Technology, Culture and Education, 15(2). https://doi.org/10.30557/QW000030

Cui, T., & Wang, J. (2023). Empowering active learning: A social annotation tool for improving student engagement. British Journal of Educational Technology, bjet.13403. https://doi.org/10.1111/bjet.13403

Hanč, J., Hančová, M., & Borovský, D. (2023). Social Reader Perusall—A Highly Effective Tool and Source of Formative Assessment Data. https://doi.org/10.48550/ARXIV.2308.07188

Mark, K. (2001). A Parallel Learner Corpus: Using Computers in a Humanistic Approach to Langauge Teaching and Research. The Japanese Journal of Language in Society, 4(1), 5–16.

Porter, G. W. (2022). Collaborative Online Annotation: Pedagogy, Assessment and Platform Comparisons. Frontiers in Education, 7, 852849. https://doi.org/10.3389/feduc.2022.852849

Tian, J. (2020). Investigating Students' Use of a Social Annotation Tool in an English for Science and Technology Course. In E. Popescu, T. Hao, T.-C. Hsu, H. Xie, M. Temperini, & W. Chen (Eds.), Emerging Technologies for Education (Vol. 11984, pp. 299–309). Springer International Publishing. https://doi.org/10.1007/978-3-030-38778-5_33