Analyzing user interactions to estimate reading time in web-based L2 reader applications

Nora Hollenstein[1] and Mircea F. Lungu[2]

Abstract. We propose to use reading time as a metric to report progress in language learning applications. As a case study we use a web-based application that enables learners of a foreign language to read texts from the web and practice vocabulary with interactive exercises generated based on their past readings. The application captures generic interactions with the web page (e.g. switching to a different tab) but also interactions directly related to language learning (e.g. clicking on a word to get a translation). We propose two metrics for approximating reading times based on user interactions with the web application. We analyze the correlation between these metrics and other interaction metrics and show that active time is the best metric for estimating the user's actual involvement with the texts and that it can be approximated from interaction metrics.

Keywords: reading time, second language learning, web-based learning, user interactions.

1. Introduction

We propose to use reading time as a metric to report progress to learners and teachers in language learning applications and in particular, in applications that support extensive reading of authentic texts in L2. Given that reading time is a good predictor of learning (Wallot et al., 2014), one simple but meaningful way to provide feedback to the readers is to show them the time they spend reading in the language of their choice. Measuring the reading time in the wild, as opposed to measuring it in controlled experiments in the lab, is however challenging: one must handle the fact that the learner behavior can only be estimated but not fully known.

1. University of Copenhagen, Copenhagen, Denmark; nora.hollenstein@hum.ku.dk; https://orcid.org/0000-0001-7936-4170
2. IT University Copenhagen, Copenhagen, Denmark; mlun@itu.dk

How to cite this article: Hollenstein, N., & Lungu, M. F. (2022). Analyzing user interactions to estimate reading time in web-based L2 reader applications. In B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouësny, & M. J. Whelpton (Eds), *Intelligent CALL, granular systems, and learner data: short papers from EUROCALL 2022* (pp. 168-173). Research-publishing.net. https://doi.org/10.14705/rpnet.2022.61.1453

As a case study, we use a web-based application named Zeeguu[3]; as mentioned in Lungu, van den Brand, Chirtoaca, and Avagyan (2018), this web-based application enables learners of a foreign language to read texts from the web about topics they are interested in and practice vocabulary with interactive exercises based on their past readings (Lungu & van Engen, 2019; Lungu et al., 2018).

We deployed this application in the wild for more than one year and for more than 606 users. The non-controlled environment means multiple challenges to reading speed estimation, e.g. users might start reading an article, but then go to make a coffee, or switch to a different tab to check social media. Sometimes readers will open an article, scroll through it for a bit, and then close it because it is not interesting to them.

To be able to estimate reading time, the web application uses telemetry regarding a variety of user interactions. These interactions include: when an article is opened and closed, when the browser tab is out of focus, and recording when the user provides feedback on the article. The application also captures interactions directly related to language learning, e.g. clicks on words and phrases for getting translations and pronunciations. Finally, the application also presents the user information about their reading time with the help of an activity dashboard.

2. Method

2.1. Data

The data used for analysis can be found in the July 2022 Zeeguu Anonymized DB Dump[4]. The analysis is based on data from 606 highschool and university students from the Netherlands and adult learners from several language schools in Denmark who were studying French, Dutch, and Danish as foreign language. Their data can be found as CSV files in the *reading-times--eurocall22* folder at https://github.com/zeeguu/studies.

2.2. Reading time approximation

Eye tracking technology would provide a perfect measurement of reading time (e.g. Hollenstein, Barrett, & Björnsdóttir, 2022), but in its absence, we have to use

3. Web: https://zeeguu.org; source code: https://github.com/zeeguu
4. Found at https://github.com/zeeguu-ecosystem/Data-Releases

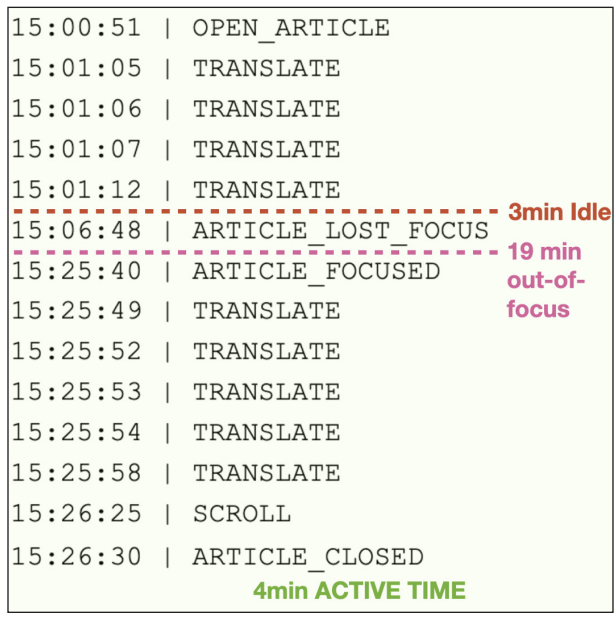
approximations based on user interactions. Based on the collected interaction data we evaluate two alternative ways.

- **Apparent time**: the full duration between opening and closing a text, used in past works not specific to language learning (Hwang, Tsai, Tsai, & Tseng, 2008; Sheard, Ceddia, Hurst, & Tuovinen, 2003).

- **Active time**: the result of subtracting from apparent time (1) out of focus time – time spent on other windows or tabs and, (2) idle time – time when a user does not interact with the system for more than a minimum threshold[5].

Figure 1 shows all the events that are recorded from the interaction of the user with ID 3083 (from the DB dump mentioned above) with the article with ID 1973640.

Figure 1. An example of a very large difference between Apparent Time (26min) and Active Time (4min)



5. Threshold of 2min in Zeeguu is used for the analysis in this paper

3. Reading time versus other metrics

We investigate whether any metric derived from the events that we collect can be used to predict reading time. This would be useful for estimating reading time when there is no previous user data about a given article.

We use Spearman's correlation to measure the strength and direction of the association between two variables computed from user interactions. These reading time measures are averaged across all events/readers for a given article. We consider the following metrics for each article:

- word_count, the number of words in a given article;
- liked, the number of times the article was liked by readers;
- translated, the number of words that were clicked to be translated in this article;
- spoken, the number of words that were clicked to be pronounced in this article;
- opened and closed, the number of times a given article was opened or closed;
- difficulty, the (between 1 and 100) normalized Flesch Kincaid reading score (higher=easier to read), as well as perceived difficulty (easy/ok/hard, the user feedback about their perceived difficulty after reading an article);
- time_to_first_translation, the time before the first translation occurs, i.e. the duration between opening an article and the first translation.

Table 1 presents the correlation coefficients ρ between these variables across all articles of all languages. Note that the correlation values are generally moderate, since the data is sparse, i.e. many fields are empty because it is optional for the user to give feedback or like an article. These results show that text difficulty does not correlate with reading times, presumably because readers automatically choose articles they feel confident they can finish in reasonable time. However, the number of words a user clicks to get a translation or pronunciation correlates with text difficulty suggesting that the used difficulty metric is a good predictor of actual perceived difficulty.

The strongest correlation in Table 1 is seen between word count and reading time (both the reading time metrics in this case). However, ACTIVE_TIME correlates most strongly with word_count and text difficulty suggesting that it is a better measure for estimating the actual time the users are involved with the text.

The results also show that users are more likely to like an article if they read it faster and need less translations. This hints at the importance of ensuring that the articles are of the right difficulty for the users.

Table 1. Correlation analysis

var1	var2	Spearman's ρ
word_count	difficulty	-0.05*
word_count	translated	0.20*
word_count	APPARENT_TIME	0.42*
word_count	ACTIVE_TIME	0.42*
word_count	time_to_first_translation	0.26*
difficulty	liked	0.05*
difficulty	translated	-0.18*
difficulty	spoken	0.17*
difficulty	opened	-0.19*
difficulty	closed	-0.18*
difficulty	APPARENT_TIME	-0.06*
difficulty	ACTIVE_TIME	-0.07*
difficulty	time_to_first_translation	-0.04*
liked	translated	-0.11*
liked	APPARENT_TIME	-0.03*
liked	ACTIVE_TIME	0.06*
liked	time_to_first_translation	-0.10*
translated	spoken	0.27*
translated	APPARENT_TIME	0.39*
translated	ACTIVE_TIME	0.42*
translated	time_to_first_translation	0.81*
spoken	APPARENT_TIME	0.30*
spoken	ACTIVE_TIME	0.27*
spoken	time_to_first_translation	0.44*
APPARENT_TIME	ACTIVE_TIME	0.84*
time_to_first_translation	APPARENT_TIME	0.59*
time_to_first_translation	ACTIVE_TIME	0.59*

Significant results ($p<0.01$) marked with *

4. Conclusions and future work

Our results suggest that active reading time is a better metric than apparent reading time and that it is possible to approximate reading time from both text-based and user-derived metrics. In the future, we plan to investigate personalizing the reading time estimation by applying prediction models based on a single user's past text interaction. We also plan to investigate presenting to a user reading times that other users required for a given article, in the same way in which sports tracking applications report comparable users achievements (West, 2015).

References

Hollenstein, N., Barrett, M., & Björnsdóttir, M. (2022). The Copenhagen Corpus of eye tracking recordings from natural reading of Danish texts. *Proceedings of LREC*.

Hwang, G. J., Tsai, P. S., Tsai, C. C., & Tseng, J. C. (2008). A novel approach for assisting teachers in analyzing student web-searching behaviors. *Computers & Education, 51*(2), 926-938. https://doi.org/10.1016/j.compedu.2007.09.011

Lungu, M., & van Engen, J. (2019). An application for L2 free reading on the web and its usage in the classroom. *Paper presented at the EUROCALL Conference 2019*.

Lungu, M. F., van den Brand, L., Chirtoaca, D., & Avagyan, M. (2018). As we may study: Towards the web as a personalized language textbook. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*.

Sheard, J., Ceddia, J., Hurst, J., & Tuovinen, J. (2003). Inferring student learning behaviour from website interactions: A usage analysis. *Education and Information technologies, 8*, 245*266. https://doi.org/10.1023/A:1026360026073

Wallot, S., O'Brien, B. A., Haussmann, A., Kloos, H., & Lyby, M. S. (2014). The role of reading time complexity and reading speed in text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 1745-1765. https://doi.org/10.1037/xlm0000030

West, L. R. (2015). Strava: challenge yourself to greater heights in physical activity/cycling and running. *British journal of sports medicine, 49*, 1024. https://doi.org/10.1136/bjsports-2015-094899

Published by Research-publishing.net, a not-for-profit association
Contact: info@research-publishing.net

Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022
Edited by Birna Arnbjörnsdóttir, Branislav Bédi, Linda Bradley, Kolbrún Friðriksdóttir, Hólmfríður Garðarsdóttir, Sylvie Thouësny, and Matthew James Whelpton

Publication date: 2022/12/12

Typeset by Research-publishing.net
Cover photo by © 2022 Kristinn Ingvarsson (photo is taken inside Veröld – House of Vigdís)
Cover layout by © 2022 Raphaël Savina (raphael@savina.net)

ISBN13: 978-2-38372-015-7 (PDF, colour)

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2022.