

Supporting Foreign Language Learning With a Browser Extension

The Zeeguu Reader

Emma Lerche Teglbrænder
emte@itu.dk

Frida Helene Beck-Larsen
frib@itu.dk

September 1, 2022

Supervisor: Mircea Lungu

MSc Software Design

IT UNIVERSITY OF COPENHAGEN

1. Abstract

In this thesis, we investigate how language learners experience a browser extension that allows them to translate and practice vocabulary in foreign language texts of their choice. We do this by developing a browser extension named The Zeeguu Reader and evaluating it with users.

The extension allows users to practice a foreign language directly in the browser by transforming news websites, blogs, encyclopedias, etc., into learning environments with improved legibility (e.g. removing distractions from the text), one-click translations, and vocabulary exercises.

The participants use the extension for two weeks and answer a survey about their experience. We find that (1) they experience the extension as useful and convenient, while also contributing to their language learning in ways that alternatives do not, (2) they have positive perceptions of self-selecting authentic reading material, and it can increase their motivation and engagement, and (3) ad removal and layout with improved legibility is experienced positively.

Building an extension for foreign language learning also comes with challenges. We find that (4) the biggest challenge is extracting text content and detecting whether texts are suitable for reading since websites are not always built using best practices. These challenges lead to scalability and maintainability problems.

Contents

1. Abstract	2
2. Introduction	5
3. Related Work	7
3.1. The Zeeguu Ecosystem	7
3.2. Language Learning Browser Extensions	8
3.3. Personalized Language Learning and Self-Selected Reading	9
4. Methodology	11
4.1. The Process	11
4.2. Technical Decisions	13
4.2.1. Support for Multiple Browsers	13
4.2.2. React JS	13
4.2.3. Readability Library	13
4.3. Usability Testing	13
4.4. Experiment Set-Up	14
4.4.1. Initial Survey	14
4.4.2. Issue Detection Survey	15
4.4.3. Final Survey	15
4.4.4. Participant Activity Data	16
5. Solution	17
5.1. Usability Test Findings	17
5.2. User Interface Design	18
5.2.1. Installation	18
5.2.2. The Popup	19
5.2.3. The Reader View	20
5.3. Implementation	22
5.3.1. Architecture and Communication With Zeeguu.org	23
5.3.2. Utilizing Mozilla’s Readability Library	25
5.3.3. Extracting Web Page Content	26
5.3.3.1. General and Individual	27
5.3.4. Reused Components From Zeeguu.org	30
5.3.5. Build and Deployment	30
6. Experiment	33
6.1. New Users	33
6.1.1. Initial Survey Data	33
6.1.2. Findings From the Experiment	35
6.1.2.1. Overall Usage	35
6.1.2.2. Observations	37
6.2. Pre-existing Users	46
6.2.1. Findings From Experiment	46
6.2.1.1. Overall Usage	46
6.2.1.2. Observations	47
6.3. Post-Experiment Activity	47
6.4. Summary	48

7. Discussion	49
7.1. Reader View	49
7.1.1. Removal of Adverts	50
7.1.2. Legibility	50
7.1.3. Images in the Reader View	51
7.2. Self-Selected Reading and Interest	52
7.2.1. Authentic Material and Text Difficulty	54
7.3. Maintainability and Scalability	55
7.3.1. Detection of Readability	55
7.3.2. Individualized Website Cleaning	55
7.3.3. Cross-Browser Extension	57
8. Reflections, Limitations, and Future Work	59
8.1. More Reading Material at the Expense of Access	59
8.2. Increasing Personalization	59
8.3. Translation Accuracy	59
8.4. DOM Extraction and Article Detection	59
8.5. Evaluating with More Pre-Existing Users	60
8.6. Evaluating Learning Objectives	60
9. Conclusion	61
References	62
Appendix A. Usability Tests	69
A.1. Usability Test Set-Up	69
A.2. Usability Test Summary	70
A.3. Usability Test Transcripts	72
Appendix B. Surveys	88
B.1. Initial Survey Questions	88
B.2. Issue Detection Survey Questions	88
B.3. Final Survey Questions	88
B.4. Changes to Survey Questions	91
B.5. Survey Questions for Pre-Existing Users	91
Appendix C. Survey Answers	92
C.1. Initial Survey Answers	92
C.2. Issue Detection Survey Answers	93
C.3. Final Survey Pilot Answers	93
C.4. Final Survey Answers	93
Appendix D. Queries	93
D.1. RSS Feed Data	95
Appendix E. Test Websites	95
E.1. Top 5 News Sites in Denmark, France, Germany, and UK	95
E.2. Top 10 Norwegian News Sites	95
E.3. Most Visited Danish Websites, and Whether They Needed Individualized Cleaning	96
E.4. Most Recent Read Articles On Zeeguu.org and Whether Images are Fetched	96

2. Introduction

In Europe approximately 92% of students are learning a foreign language in school [13] and almost 60% of upper secondary students learn two or more languages [26]. Thus, there are bound to be many language learners with different learning preferences and interests.

In recent years many digital platforms for language learning have emerged. The online and app-based platform Duolingo had over 500 million worldwide users in 2021 [23], while a similar platform, Babbel, celebrated 10 million premium subscribers in 2020 [5]. These growing numbers indicate that language learners are interested in using new technology to learn foreign languages digitally out of the classroom.

The field of computer-assisted language learning (CALL) has evolved from “(...) examining questions about whether CALL is effective for language learning to how the affordances of technology might best be exploited to provide learners with optimal language learning opportunities” [64, p. 17]. One way we can exploit the affordances of technology is the opportunity to use authentic material.

Before the internet, gaining access to up-to-date authentic texts in foreign languages was problematic. However, “authentic materials can increase the quality of foreign language teaching” [63, p. 334]. With the rise of the internet, most of the world’s population was given access to immediate resources like native news websites, which in turn could be used as learning material [56, p. 33-34]. This kind of access is “(...) crucial to develop learners’ confidence in facing the unpredictable challenges of authenticity and their awareness of the cultures where this language is the main means of communication and expression” [56, p. 33]. Using news websites and other authentic foreign material as a learning resource also has the advantage that learners can choose the material they find interesting.

Research in educational psychology has demonstrated that the interest of readers contributes to increased comprehension and learning [33, p. 195], [44, p. 172], [45, p. 120] and has a significant effect on text recall [25, p. 228]. However, “utilizing individual interest in educational settings may be a very time-consuming and effortful task” [33, p. 203]. Therefore, tools for making it more accessible to include individual interests are needed.

One such tool is the foreign language reader and exercise prototype Zeeguu. The platform crawls the web to provide personalized article recommendations based on interest and difficulty with the aim of helping the learner create a “personalized language textbook” [52]. Users can then choose articles to read from the recommendations and do exercises based on the words they translate.

However, Zeeguu faces one big challenge; crawling the web limits the number and variety of articles, because the website must have an RSS feed, and it must be added manually. Zeeguu must also make sure that all websites allow for the content to be redistributed into zeeguu.org to avoid copyright infringement. This limits the number of resources that can be made available.

To solve these challenges, we develop a browser extension, building on top of the functionality from zeeguu.org. In general, browser extensions are used to improve or alter the browser experience by adding new features to pages, providing shortcuts, or personalizing the visit in other ways [31].

The new Zeeguu extension makes news websites, blogs, encyclopedias, etc., readable directly in the browser on the original website while at the same time leveraging all the same learning features as zeeguu.org. Thus, it will allow users to personalize their reading experience and increase the number of possible articles they can read without compromising their access to translations and vocabulary exercises.

Therefore, we define the focus of this thesis to be the following research question:

How do language learners experience a browser extension that allows them to translate and practice vocabulary in foreign language texts of their choice?

To answer the question, we build and deploy a browser extension, test the usability of the extension, and then further evaluate it by conducting a two-week experiment with language learners. After the two weeks, the participants answer a survey about their experiences, and we then discuss the technical challenges and opportunities that are associated with our approach in relation to the results from our study.

3. Related Work

Much has been written about foreign language learning; thus, the literature is extensive and diverse. Because the scope of our project is limited to foreign language learning with an extension, we will investigate literature in this area.

We will start by briefly looking into previous studies of the Zeeguu ecosystem to outline how the current system works and what has previously been investigated academically.

We will then examine the literature on foreign language learning browser extensions while also investigating what is currently available for download. This provides us with knowledge on how others have developed similar extensions and what is not available.

We finally explore the literature on personalized language learning and self-selection of material, in order for us to use previous research results when developing our extension and constructing our surveys.

3.1. The Zeeguu Ecosystem

The foreign language learning prototype Zeeguu is a “personalized language textbook that uses the web as its content source” [52, p. 11]. The system provides text recommendations based on interest and difficulty (computed by Zeeguu), which is then used for reading comprehension and vocabulary practice [52, p. 3]. The system crawls the internet for articles, and while reading texts of the learner’s choosing, translations are provided by the click of a word [35, p. 3-4]. The system uses multiple existing industrial-grade translation APIs to increase the likelihood that at least one is correct [52, p. 3]. The translated words are used to generate personalized vocabulary exercises by utilizing the original context in which the words were encountered¹.

In a 2018 study, Lungu et al. concluded that students of zeeguu.org take advantage of and appreciate the opportunity to read personalized material. But they also conclude that learners want better ways to find personally relevant content [52, p. 11]. In regard to the learning outcome the study found that the participants “vocabulary is enriched with new words and the knowledge of other words is strengthened” [52, p. 11]. Since the study, zeeguu.org has been further developed, and e.g., topic filtration of articles has been implemented.

In conclusion, zeeguu.org helped participants strengthen their knowledge of words, but they also saw that the system could be improved by providing an easier way to find and practice with personalized material.

¹<https://www.zeeguu.org/>. To create an account use invite code: zeeguu-extension

3.2. Language Learning Browser Extensions

In terms of academic literature on browser extensions for language learning, there are a few studies that have implemented browser extensions to support foreign language learning.

Trusty and Truong have built an extension that works on native language web pages. They try to achieve accidental learning and microlearning by selecting words on native language web pages and translating them to the foreign language directly on the page [77, p. 3179]. They focus on translating mainly nouns, and they provide exercises when hovering over a word, for instance, by asking the user to provide the meaning of the foreign word [77, p. 3181].

Similarly, Elbatanony et al. built a Chrome extension that provides selected translations directly on first language webpages. Users can then see the original word and add it to their dictionary [24, p. 4]. Corbin et al. have developed an extension that opens new language learning material every time users open a new tab or window [16]. The material can be completed within 30 seconds or less, thus also falling into the category of microlearning [24, p. 345].

What these three have in common is that they try to make the users practice a language by nudging them in their day-to-day use of the web. Likewise, none of these extensions focus on foreign language reading but mainly on vocabulary training.

One study partly focused on reading and vocabulary practice with the browser extension Readlang. Readlang provides translations directly on the website and then allows the user to review the translated words. They study the effect of “narrow reading”, which is defined as “reading thematically-related texts as well as the books or texts written by the same author” [41, p. 14]. Their participants were instructed to read from a fixed set of 12 articles on the same topic from the same website, and they found that it helped with vocabulary learning within the topic they were studying. Thus, they did not investigate the effect of letting learners choose their own reading material within their interests.

Chen, Zheng et. al. has developed and evaluated a Chrome browser extension “that allows readers to learn a second language vocabulary while reading news online” [14, p. 34]. Their study focused on translating English to Chinese, and the extension is triggered when the user visits a set of predefined English news websites. They conclude that “one respondent noted that they would like to use it on arbitrary websites”, but that the translation technology they have built using word sense disambiguation is difficult to develop, even in the narrow domain of news sites [14, p. 41]. The extension is not currently available in the Chrome web store.

There are other language learning extensions focused on reading, but they have not been investigated in academic research. The extensions Vocab Tracker [78] and Remembermy [67] work directly on websites by making it possible to translate words and hear pronunciations, while the webpage layout is kept unchanged. They both have either exercises or flashcard functionality to review or practice the translated words.

Remembermy requires more clicks to get the translations, whereas Vocab Tracker and Readlang (Readlang, 2022) provide translations with just one click.

The extension LingQ [51] instead only lets the user import texts from various websites through the extension to LingQ's own website. To read and practice with the text, users have to go to the website, and here they are provided with one-click translations. Their system also cleans up the article, removing all formatting, except for the headline, keeping only the text and sometimes a tiny 64x64 px image. The user can also customize the font size, line spacing, and page width to fit personal needs.

The extension Toucan works directly on websites, but only on pages in the user's native language. It chooses words to translate into the language the user is learning [39], thus being similar to the approach implemented in Trusty and Truong's extension, where a selection of words on native language websites are translated to enforce vocabulary learning [77].

Both Vocabtracker, Readlang, and Remembermy are only available in Chrome browsers, whereas LingQ and Toucan are available for both Firefox, Safari, and Chrome users.

In conclusion, nobody has developed and evaluated a browser extension for foreign language reading that removes clutter, improves legibility, includes language exercises, and works directly on the website. Secondly, the existing extensions that work directly in the browser tab are only available to Chrome users. However, they all transform websites into learning material; thus, personalizing the web, which goes beyond the intentions of the host. So, extensions are a way of personalizing foreign language learning.

3.3. Personalized Language Learning and Self-Selected Reading

Personalized learning has been a topic of research for some time [71, p. 1]. The term has become a buzzword in educational settings [79, p. 235]. Still, it has a wide range of definitions [7], [12, p. 3-4]. The U.S. Department of Education defines learning personalization as "paced to learning needs, tailored to learning preferences, and tailored to the specific interests of different learners" [62], and Bernacki has summarized the different definitions and found that the majority of personalization definitions have focused on identifying and accommodating students' "interests" and "needs" [7].

So, despite no clear definition of personalization "Recent educational changes in methods, curriculum design, and pedagogical approaches stress the importance and effectiveness of personalized learning as opposed to traditional cohort-based learning" [37, p. 107].

One literature review by Ismail et al. focuses on personalized learning as something where systems are personalized based on user information that is either given to the system or derived from analyzing user data [37]. They divide personalization into four categories based on whether it stems from self-description or system predictions [37, p. 198-199]. Personalization through self-description is based on questionnaires, surveys, or forms. In contrast, personal-

ization through segmentation is based on demographics and cultural characteristics, while other categories are based on predictions from systems that collect data about the user. According to the study, most of the investigated literature implements cognitive-based personalization, which is more advanced than self-described personalization, as it tries to personalize based on cognitive characteristics that are not explicitly expressed by the learner [37].

They conclude that Computer-aided language learning (CALL) has become more sophisticated toward personalized language learning, where learners are now given more choice and voice in their learning [37, p. 202].

However, in the study by Ismail et al. personalization does not equal self-selection of material. Instead, personalization is gained when a user provides information to a system, and the system then provides material based on the information it has received.

Another form of personalization is based on topic interests. The effect of topic interests on reading comprehension has not shown consistent results, and the results also vary depending on the learner's language level. L1 reading research has previously found that topic interest is important in reading comprehension, but L2 reading has not shown these results [44, p.162]. However, Lee investigated reading comprehension with sixty Korean L2 readers and found that "interest had subtle but systematic effects on the comprehension" [44, p. 173]. Another study using intermediate and advanced classes found that "Topic interest has been shown in this study to be an important motivational condition of text comprehension." [68, p. 334].

Other studies focus more on personalization concerning self-selection of topics for learning and the learner's experiences with this. One study regarding self-selected topics of learners found "(...) learners' perceived topic interest and familiarity to be more significant than topic importance and difficulty" [70, p. 20]. So, the study found that "learners' interests play an important role in English language classes" which is also found in multiple other studies regarding learners' interests [70, p. 19], [10], [82], [76]. Shakourzadeh and Izadpanah even found that "Learners would feel more motivated to learn the language and take part willingly in interesting activities" when they used self-selected topics of their interest [70].

So overall, studies agree that personalization in the form of interest can be an important factor in learning, and this can lead to greater motivation and comprehension [33], [4], [2], [79]. But some studies have also found that self-choosing materials do not increase motivation or engagement for all types of students, as too many options could also have a negative impact [27], [11].

4. Methodology

To answer our research question, we build and deploy a browser extension. To evaluate how the learners experience this extension, we conducted a two-week experiment, which ended with a survey.

None of the existing in-production extensions met the expectations we had for a language learning browser extension. Thus, we decided to build and deploy our own extension for the language learning platform zeeguu.org. We did this because we wanted to be able to:

1. Perform the experiment with an extension, where legibility and the general reading experience was in focus.
2. Make changes and updates to the extension based on usability testing and pilot testing with new users.
3. Gain insight into participant activity by having access to the database.
4. Explore how pre-existing users of a website-based language learning platform experience a new extension-based language learning platform.

The data emerging from the study consist of the following: usability test data, initial survey data, final survey data, and activity log data. Besides this, the study also resulted in the extension, The Zeeguu Reader, which is now available for download in the Chrome Web Store and Mozilla's Add-ons library.

In the following sections we describe our process when conducting this study, as well as motivate the decisions we made in regard to the technical aspects of building the extension. Furthermore, we motivate the methodological choices made in regard to usability testing, the experiment, and the associated surveys.

4.1. The Process

When conducting this study we have gone through three different phases: the initial phase, the pilot phase and the final phase. Our process has been iterative, meaning that continuous user involvement has resulted in further development of the extension. The three phases are described in Figure 1.

Figure 2 shows a Gantt chart of how the different activities in this study were placed in relation to each other. We began development in week 5, in February 2022, and we committed the last changes in week 35, in August 2022. Besides fixing small bugs, no changes were deployed to the extension while participants were active in the two-week experiment.

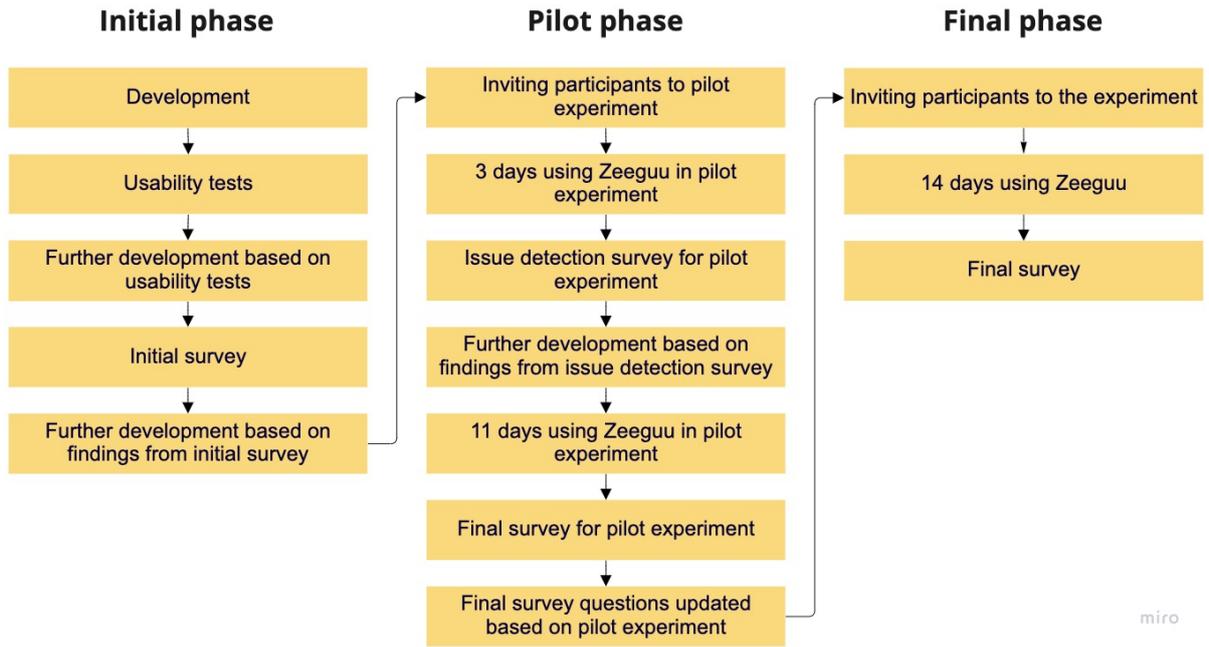


Figure 1: Flowchart displaying the phases in our study

The initial survey and the final survey were open for 9-10 weeks, thus they overlapped each other. As people signed up in the initial survey, they were invited for the two-week experiment, which was followed by the final survey.

Activity	Week of 2022															
	5-13	14	15	16	17	18	19	20	21	22	23	24	25	26-29	30	31-35
Development	[Yellow bar]															
Usability tests		[Yellow bar]														
Initial survey			[Yellow bar]													
Pilot experiment begin				[Yellow bar]												
Issue detection survey				[Yellow bar]												
Experiment begin					[Yellow bar]											
Final survey for pilot						[Yellow bar]										
Final survey							[Yellow bar]									

Figure 2: Gantt chart of our weekly activities

4.2. Technical Decisions

4.2.1. Support for Multiple Browsers

When developing a browser extension, one must decide which browsers it should be available to. We have decided to make a cross-browser extension for Chrome and Firefox, which means it also works in Edge, Opera, and Brave browsers because they support the use of Chrome-developed extensions.

According to statcounter.com in October 2021 the Desktop Browser Market Share Worldwide was divided into 67,17% for Chrome, 9,63% for Safari, 9,14% for Edge, 7,89% for Firefox, 2,89% for Opera and 3,29% for others [75]. This means that by being available on Chrome, Firefox, Edge, and Opera, we can target around 87.1% of desktop users worldwide.

It is also possible to add support for Apple's Safari, but to develop an extension for Apple requires joining the Apple Developer Program, which has a yearly fee of 99 USD per membership [65]. Additionally, the extension has to be developed on Apple's developing platform Xcode. Because of this, we decided not to implement a Safari extension.

4.2.2. React JS

The project is developed in React with JavaScript. We use React because zeeguu.org is implemented with this technology. The repository for zeeguu.org is called zeeguu-react. Using the same language and framework allows us to reuse large parts of zeeguu.org. We can simply import components into the extension and therefore improve maintainability because changes to zeeguu-react components will automatically be updated inside the extension.

4.2.3. Readability Library

We used Mozilla's stand-alone library Readability [58] to prepare the text content for the reading interface. This library is actively used for Firefox's Reader View; thus, it is a library that Mozilla updates regularly. This limits the chances of the library suddenly being deprecated. Mozilla's Readability is based on Arc90's open-source Readability library, which also applies to Safari's Reader View. Safari's implementation is, unlike Firefox's, not publicly available [29].

4.3. Usability Testing

Before starting the experiment, we performed eight usability tests of the first version of The Zeeguu Reader with users who had never interacted with the Zeeguu platform before. This helped us catch initial problems, which could be fixed before the experiment began. We did the usability tests in two iterations. First, we tested with four users and made small adjustments. We did so because the first tests revealed problems, and we wanted to see if small changes would fix them. This is also the approach suggested by Jakob Nielsen [61]. Afterward, we tested with the last four users.

The usability tests were conducted with a facilitator and an observer. When performing the usability tests, we encouraged the testers to think aloud. The test setup can be found in Appendix A.1: Usability Test Set-Up. We tested with six females and two males aged 22-31.

After the usability test, we followed up with six questions about their experiences.

We divided the findings from the usability tests into the six categories proposed by Søren Lauesen [43, p. 413-442]. After identifying and categorizing the errors, we choose what to change based on cost, e.g., the number of work hours it will take to correct the problem, and the benefit, e.g., the importance of the problem to the user [43, p. 442].

4.4. Experiment Set-Up

We recruited two different kinds of participants for our experiment: new users who had never used Zeeguu before and users who were already actively using the website; we call them pre-existing users.

We recruited new users for our experiment by sharing a post with a signup form in Facebook groups for people learning languages and on multiple language learning groups on Reddit. Thus, we aimed to recruit people who are interested and motivated to learn a foreign language.

The initial survey was only used to recruit new users; thus, the pre-existing users of zeeguu.org we recruited have not answered it.

We recruited pre-existing users of zeeguu.org to collect knowledge on how they experience The Zeeguu Reader compared to zeeguu.org. We recruited a class of Dutch students, who study French. We contacted other existing users but only recruited one more, who also answered the survey.

We call the users who used the extension our participants and we call the users who answered the final survey our respondents.

We used descriptive statistical methods to analyze the quantitative survey data. In order to keep more nuances and complexity [54, p. 21], we analyzed the free text survey responses and summarized our findings in observations.

The users we invited to participate in the experiment were instructed to use the extension for a two-week period. We gave brief instructions on how to install the extension, and after the two-week period, we sent them a survey, which will be described in further detail below.

4.4.1. Initial Survey

The initial survey included general background questions about the participants, such as their name, age, and occupation. We also asked them about their native language, the language they

were learning, their language level, how they usually learn, and how long they spend learning each week. The questions can be found in Appendix B.1: Initial Survey Questions.

We added eight new native languages (Albanian, French, Polish, Spanish, Ukrainian, Russian, Portuguese, and Vietnamese) and three new languages the users are trying to learn (Ukrainian, Hungarian, and Norwegian), to accommodate the respondents' wishes.

4.4.2. Issue Detection Survey

To minimize problems during the two-week experiment, we started with a pilot. Here we invited 25 users to try the extension. Nine participants made an account. After using it for three days they were sent a survey. Five people answered the small questionnaire about the installation process and whether they encountered problems (See Appendix B.2 for questions and Appendix C.2 for answers). They continued to use the extension for the next 11 days, after which they received the final survey.

4.4.3. Final Survey

Because we had two different kinds of participants, we also had two different final surveys. One for new users of the Zeeguu platform and one for pre-existing users of the Zeeguu platform.

The final survey for new users contains six parts; the first part consists of background questions to identify the participant. The second part consists of general questions about their experience with the extension. The next three parts are composed of questions about Finding Articles, Reading, and Exercises. We have placed the general questions first because they are more appropriately placed before the specific questions [50, p. 256].

We have aimed to keep the questions as short as possible to reduce the cognitive load on the participants [50, p. 250]. We have used 5-point scales for most quantitative questions because having an uneven number of options on scales increases validity, as it gives a middle, neutral option [50, p. 265]. For the scale and multiple-choice questions, we have included an open text field for users to elaborate on their answers. The quantitative questions using scales can more easily be summarized for analysis, whereas the questions with open responses can be more significant. However, they can also be more time-consuming to analyze [18]. The survey questions can be found in Appendix B.3: Final Survey Questions.

The survey for the pre-existing users is a shorter version of the same survey: we excluded questions about reading and the exercises because they already know these parts from zeeguu.org. Instead, we ask them questions comparing the extension to the website zeeguu.org. The survey can be found in Appendix B.5: Survey Questions for Pre-Existing Users.

The 9 participants who created an account after being invited to the pilot experiment, were also sent a pilot final survey (Appendix C.3: Final Survey Pilot Answers). Based on their answers

four changes were made to the survey before sending it to everyone else. The changes are described in Appendix B.4: Changes to Survey Questions.

4.4.4. Participant Activity Data

The activity data is collected by logging user actions into the database. The logged activity includes e.g. opened articles, words translated, reading duration, and exercise duration.

We decided to include the activity data from the participants, who did not answer the final survey. We did this because we can still use their answers from the initial survey, how they used the extension during the experiment, and if they continued to use it afterwards.

We have fetched the data from the database with SQL and processed in a Jupyter Notebook and analyzed it in Google Sheets and Excel (See Appendix D: Queries). We used descriptive statistical methods to analyze the activity data.

After the two-week experiment, we check the database to see if some participants have continued to use the extension.

5. Solution

To use The Zeeguu Reader², users have to be on a web page with a “readable” article. The extension evaluates whether the article can be read by the reader or not, by assessing if the language is supported and if the page contains a main text area. If the chosen article is readable, when the users click on the extension icon in the toolbar, they see a “Read article” button. This button opens the The Zeeguu Reader window in the same tab, overwriting the current websites DOM. We will call this window the reader view. Here a cleaned version of the article is displayed, meaning that only the main text content remains, and everything else, like ads, navigation, buttons, and other elements are removed.

In the reader view, users can interact with the text by translating and hearing the pronunciation of the words by clicking them. Each word must be clicked individually. When clicking on an adjacent word the translation is extended. In this way, an entire sentence can also be translated. Afterward, they can review their translations and do vocabulary exercises with the words.

5.1. Usability Test Findings

The usability test revealed 20 issues. The categorization of the issues and a summary of the tests can be found in Appendix A.2: Usability Test Summary. The usability tests were conducted in week 14 of 2022. Transcriptions of the eight usability tests can be found in Appendix A.3: Usability Test Transcripts. Based on the two iterations of usability tests, we made changes to the interface.

The changes we made were regarding the process of finding articles, unclear wordings on buttons and in texts, and buttons that did not function as users expected. We encourage the reader to read the appendix after they have read the chapter “User Interface Design”.

We implemented the solutions to the errors before we started the two-week experiment. After each usability test, we interviewed the users, and we found that they were overall happy with the extension and their experiences. They liked the minimal design and simplicity. One user said it was “Visually appealing, and it is clear what to do” (Usability test 1). Another said that “It’s just really nice and simple” (Usability test 6) and that it was “easy to navigate”(Usability test 3). Overall, they did not find any important functionalities missing, and despite some minor issues, they had a positive experience.

²<https://chrome.google.com/webstore/detail/the-zeeguu-reader/ckncjmaednfephhbpeookmknhmjjodcd>.
To create an account use invite code: zeeguu-extension

5.2. User Interface Design

5.2.1. Installation

When a user creates an account and does not have the extension installed, they see a message encouraging them to install it (Figure 3). The first time a pre-existing user visits zeeguu.org after the extension is published, and does not have the extension installed, they will be met by a popup, which tells them to install the new extension (Figure 4).

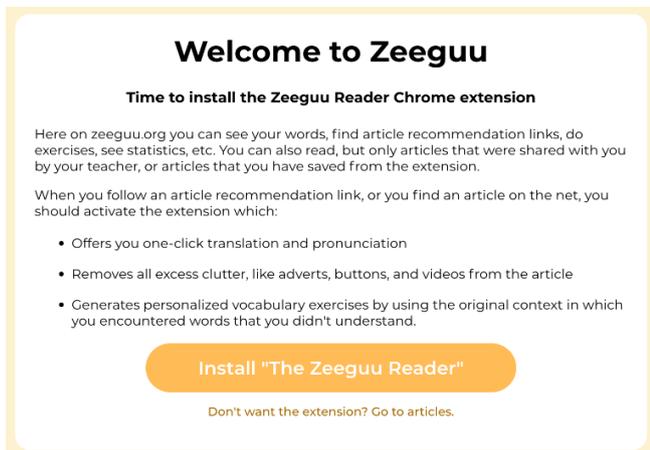


Figure 3: Installation for new users



Figure 4: Installation for pre-existing users

If a zeeguu.org user does not install the extension, a reminder will be displayed every time they go into the article recommendations (Figure 5). This is necessary because previously, users could read articles directly on zeeguu.org, whereas now the original article opens in a new tab.

Texts



Figure 5: The top of the recommendations page showing a reminder

When the user installs the extension, they are led to an installation page on zeeguu.org (Figure 6). This page displays a GIF animation and links to a video that explains how to pin the extension to the toolbar. If the user is not logged in, they will be met with the possibility to sign in or create an account. If the user is logged in when installing the extension, they will be given the option of going to Zeeguu's article recommendations.

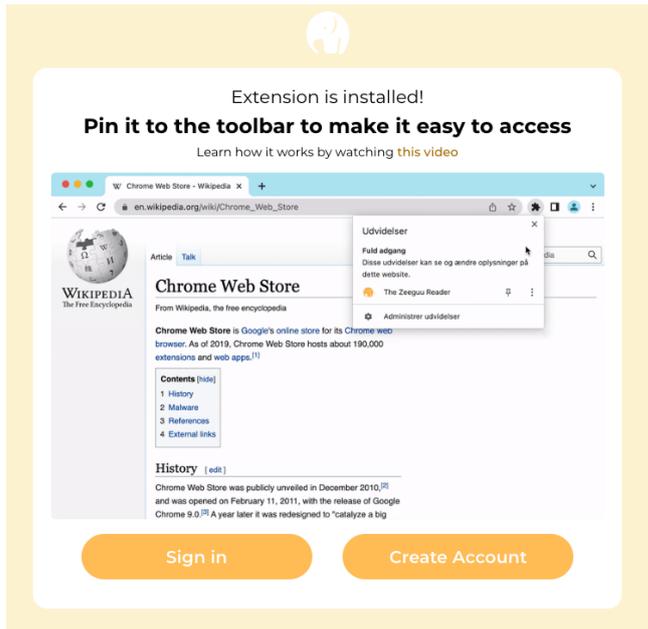


Figure 6: Installation page on zeeguu.org

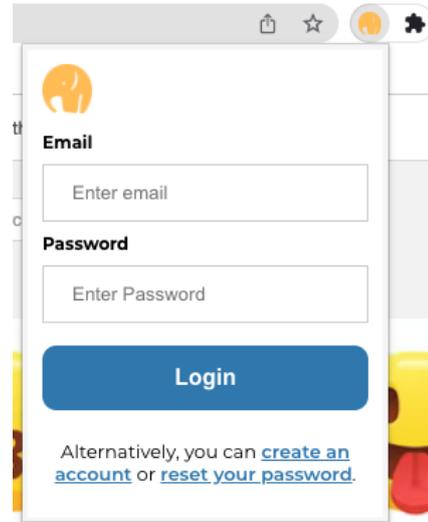


Figure 7: Login component of the extension

5.2.2. The Popup

The goal when designing the popup was to make it as simple and minimalist as possible. If the user is not logged in on zeeguu.org or in the extension beforehand, the login form will be displayed when the user opens the popup (Figure 7). If they do not have an account, users can go to zeeguu.org to create one through the link at the bottom. This way we did not implement duplicate behavior but reused the account creation from zeeguu.org.

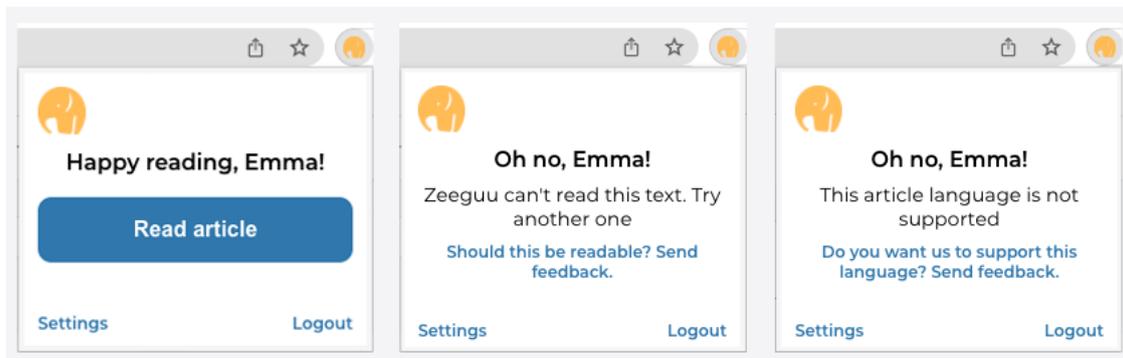


Figure 8: Possible interfaces in extension

When users are logged in and open the extension, they can be met by three different interfaces (Figure 8). When the popup is opened, it will check whether the article is readable and whether the language is supported. While this happens, the popup will display a loading circle animation. If the loading takes less than 100 seconds, we will not display the animation, as this will appear

as an annoying glitch. But if the loading takes longer than 100 milliseconds, we will show the animation for at least 900 milliseconds. Suppose the article is readable and the language is supported. In that case, users will see one big button “Read article”. If the article is not readable, they have the option of notifying us if they think this is an error. The same feedback option is available if the language is not supported.

5.2.3. The Reader View

We have focused on legibility in the design of the reading interface because the extension’s purpose is foreign language reading. Legibility is defined as “the effort to distinguish individual characters from the background and each other and includes visual aspects such as text-background luminance contrast, letter spacing, and letter case” [55, p. 2]. Thus, legibility is concerned with the visual presentation of text, whereas readability is concerned with the writing style [22, p. 3], even though readability is often used interchangeably with legibility.

We tried to improve the legibility in several ways. Firstly, by using the Readability library by Mozilla we improve the perceived readability [49, p. 1].

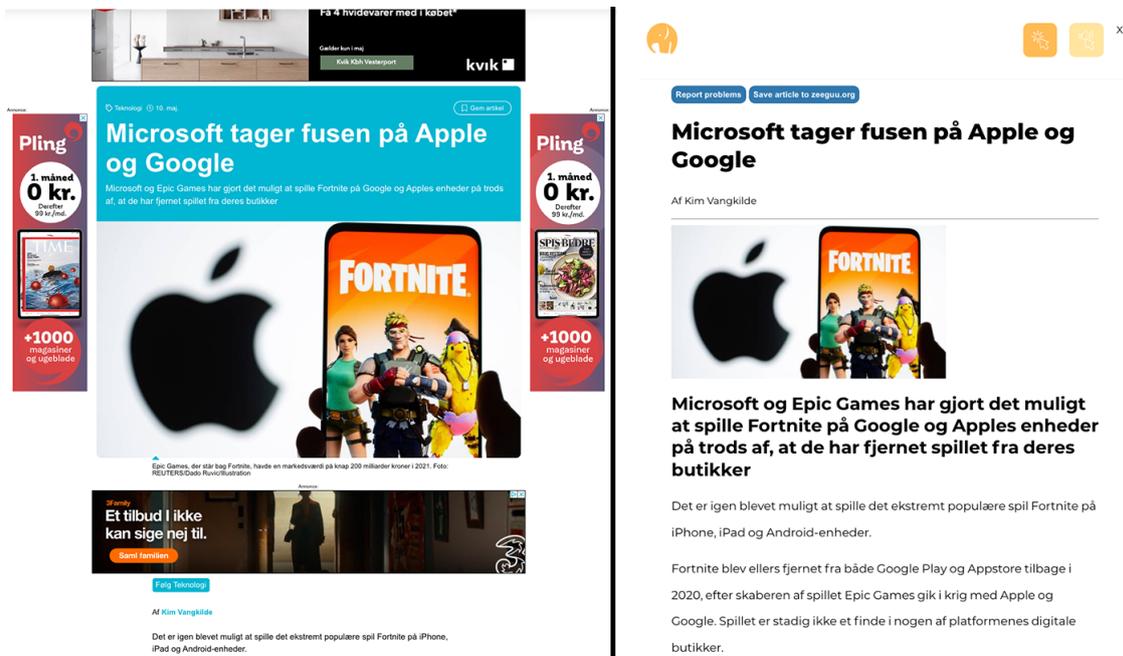


Figure 9: On the left is the original article on the website. On the right is the same article in the reader view

Secondly, we have improved the legibility of the text by altering the text presentation. We need a line-height above 2 to fit the translations above the original words (Figure 10). This has the added benefit that the larger the line height, the longer the lines can be without affecting legibility. Thus, we have set the width of the text in the extension such that the line length does not exceed approximately 80 characters per line, depending on the window size. This aligns

with recommendations from several sources [9, p. 139]. An example of an original webpage and the reader view version is displayed in Figure 9.



Figure 10: The reader view with translations

We also decided to change the color schemes otherwise used by zeeguu.org. We have preserved the yellow color but also added a complementary blue color. We did this because the yellow color used on zeeguu.org is not accessible as the background color for buttons, as it does not comply with WCAG 2.1 [15]. So, to improve readability for all users, we made sure that everything now follows this standard in terms of color.

We have continued the use of the font Montserrat from Zeeguu to keep the design similar. It is a sans serif font which have been proved to be better for screens than serif font types [34]. Compared to the articles on zeeguu.org, the extension improves legibility because, in most cases, sub-headers and lists are preserved.

After the article, the user is presented with the option to practice the words they have translated from the text (Figure 11). Clicking this buttons leads to the exercises (Figure 12).

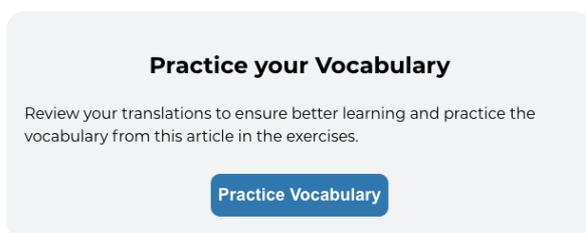


Figure 11: The bottom of the reader view

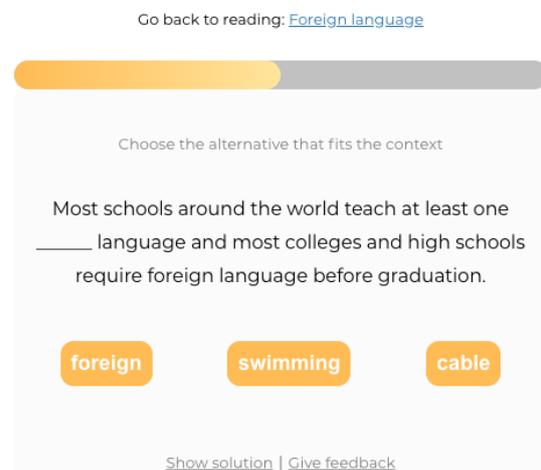


Figure 12: Example of exercise in the reader view

5.3. Implementation

The first part of the study was to develop the first version of the extension for Chrome and Firefox, so that we could later conduct our experiment. The extension is implemented with help from Mozillas' Readability, but also reusing components from zeeguu.org.

The source code for the extension is available on GitHub:

<https://github.com/zeeguu-ecosystem/ZeeguuExtension>

In our React project we have two primary components: `Popup.js`³ and `Main.js`⁴. In figure 13 we walk through the main flow of the two components from a technical point of view. Thus, this does not include all functions or possible flows, instead this is an overview of what happens if the chosen article is readable with the extension. The dark blue square represents user actions. The light blue diamond represents what happens in the user interface. The yellow squares represent functions from our React project, and the gray square represents objects.

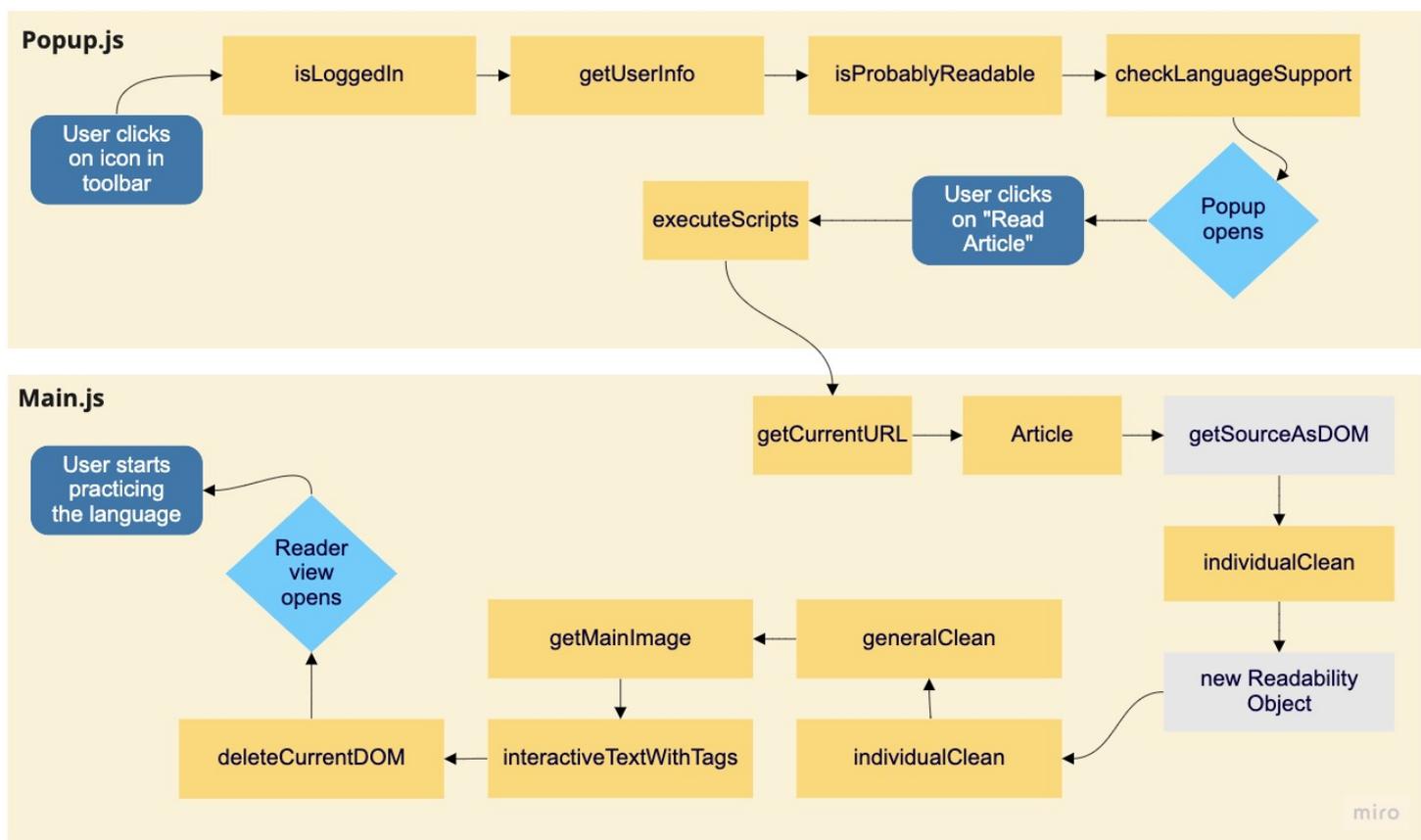


Figure 13: Overview of the main flow when users open the extension.

³<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/main.js>

⁴<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/popup/Popup.js>

As the figure displays, we run a number of checks when the extension icon in the toolbar is clicked. When the user clicks “Read article” the DOM of the current website is transformed. Most importantly, we start out by cleaning the current website’s DOM, transforming the Document object to a Readability object, and then cleaning the Readability object, because Readability does not always sufficiently extract the correct text content. Afterwards the text content is transformed into interactive text and the current DOM is deleted, and lastly, the extension opens the reader view with a cleaned and interactive version of the text content, where all functionality from zeeguu.org is implemented.

In the following sections we go through how we implemented the different parts of the extension, mainly focusing on architecture, communication, Readability and extraction of text content.

5.3.1. Architecture and Communication With Zeeguu.org

The architecture of the extension is summarized in Figure 14. The Zeeguu Reader is a 3-layer extension that runs mainly in the browser. Below we will go through the different aspects of the extension. Each communication path has an annotation letter that will be used in the description of the model.

Main.js is injected into the current webpage’s DOM when the “Read Article” button is clicked (See “a” in Figure 14). The component clears the current website’s DOM and creates a new DIV element, in which all our React components are rendered. Main.js also communicates with the database to generate interactive text with translations, exercises, and when logging the participant activity, etc (See “b” in Figure 14).

The file index.html is defined as the default popup. This is what users see, then they initially click the extension in the toolbar. Index.html renders the component Popup.js (See “c” in Figure 14). Popup.js communicates with Main.js through local storage to share user information such as the name and the native language (See “d” in Figure 14).

In order to unify the login between zeeguu.org and The Zeeguu Reader, we changed the login implementation such that it required more complex communication between the extension and zeeguu.org. We did this by making the login depend on a session id stored in a cookie⁵. This allows the extension to check whether there are cookies on zeeguu.org by using `chrome.cookies.get` [20]. If a user is logged into zeeguu.org, the extension will detect this⁶, and they will also be logged into the extension and vice versa. If they log out from either the extension or the website, they will be logged out of both platforms. This creates a smooth experience for the users.

⁵<https://github.com/zeeguu-ecosystem/zeeguu-react/blob/master/src/utils/cookies/userInfo.js>

⁶<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/popup/cookies.js>

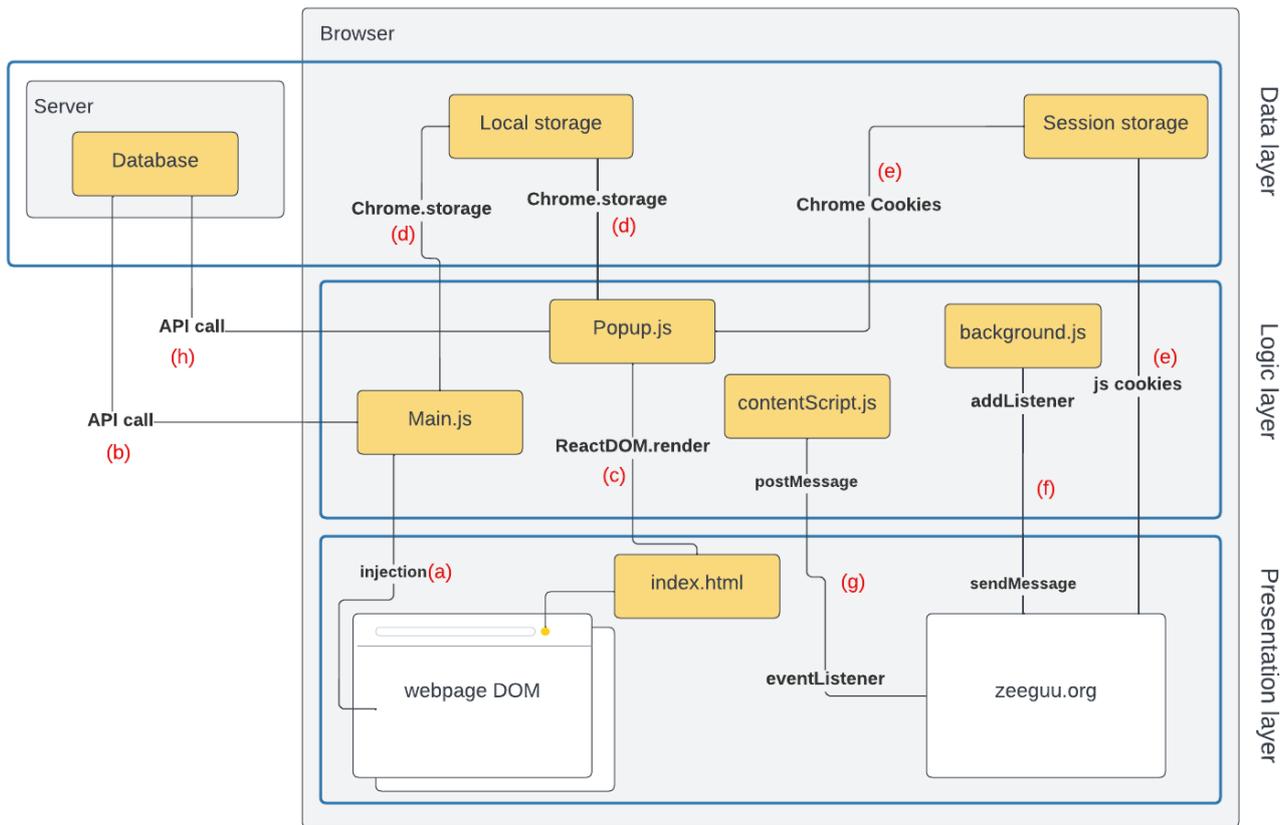


Figure 14: The architecture of The Zeeguu Reader.

Popup.js communicates with zeeguu.org about the logged-in user with cookies through session storage (See “e” in Figure 14) and with the database through the Zeeguu API (See “h” in Figure 14). For instance, if the user decides to change their native language, this will also be visible in the extension immediately.

Lastly, the extension also communicates with zeeguu.org through background.js (See “f” in Figure 14) and contentScript.js (See “g” in Figure 14), which talks to the Chrome and Firefox browsers respectively. We do this, so that we can display a message on zeeguu.org if the extension is not installed. In both instances, we send messages between the extension and zeeguu.org. A listener listens for a message, and if it is received, then the website will know if the extension is installed or not.

When using a Chrome browser, background.js detects if the extension is installed using the function “onMessageExternal”. This function uses the “externally_connectable” parameter in the manifest to send a message to Zeeguu.org⁷. Inside the extension, we set up a listener, which

⁷<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/public/background.js>

waits for a message from zeeguu.org. If the listener receives a message, we send a response that lets zeeguu.org know that the extension is installed. The function `onMessageExternal` is not compatible with Firefox [57]. Thus, when using Firefox we instead use the `“window.postMessage”` function inside of `contentScript.js`⁸. Here we send a message to Zeeguu.org. If this message is received by Zeeguu.org, the website will likewise know that a Firefox extension is installed.

The communication between zeeguu.org and the extension is implemented inside of `extension-Communication.js`⁹ in `zeeguu-react`.

5.3.2. Utilizing Mozilla’s Readability Library

To transform the articles in the browser we use the Readability library from Mozilla [58].

Readability is used to remove adverts, navigation, buttons, etc., and to extract the text from the website. It assesses which text to extract, normally the main text content from a website.

The library scores “a set of possible text-centric HTML elements in the DOM” and chooses the highest scored element as a top candidate based on the number of characters [59] and the total number of commas in the text [53]. So, it looks for the element on the page with the highest density of texts and also uses text patterns to identify, for instance, the author and source [29, p. 2]. When we use Readability, everything besides the main text content is removed.

Readability also helps us determine whether a website should be readable or not, meaning whether it can be opened in a reader view. This is based on if the text is, amongst other things, long enough. If the text has a top candidate, it is deemed readable [59], [53]. This is a difficult assessment, and according to the documentation their method will produce both false negatives and false positives [58]. If a website is determined not to be readable, then the user cannot read it with the extension.

The method used to determine if the text is readable, `isProbablyReadable`, has a parameter “options”. This options object accepts a number of properties: `minContentLength`, `minScore`, and `visibilityChecker`. These properties will affect the outcome of using `isProbablyReadable`, because we can determine what it takes for an article to be readable [58]. If we are too strict, then we will limit the number of articles that can be read in The Zeeguu Reader. If we are not strict enough, The Zeeguu Reader will open text content, which should not be readable.

As Mozilla forecasted, we did experience that `isProbablyReadable`¹⁰ did not always produce the correct result. Therefore, we implemented our own check, `checkReadability.js`¹¹. We use this

⁸<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/public/contentScript.js>

⁹<https://github.com/zeeguu-ecosystem/zeeguu-react/blob/development/src/utils/misc/extensionCommunication.js>

¹⁰<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/f52197cb50f8362d56a6102b8720d6ab3b8bbdb9/src/popup/Popup.js#L64>

¹¹<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/popup/checkReadability.js>

only if we know that a specific webpage is wrongly assessed by Mozilla’s Readability library. For instance, live articles on dr.dk are not rendering correctly. Thus, our checkReadability.js function will determine that live articles on dr.dk should not be readable. In some cases, we also label articles behind paywalls as not readable because they do not render well.

5.3.3. Extracting Web Page Content

Besides using Readability to extract the main text content, we often also have to manually extract or remove additional information from websites; thus, the websites need additional cleaning. We need to do this, because Readability removes information that it does not consider a part of the main text content, but this is not always correctly identified. Likewise, sometimes it accidentally preserves text which is not part of the main text content.

In Figure 15, we show an example of a Danish website needing cleaning. BT.dk adds two elements to the article: One asking the user to subscribe and another asking them to listen to the article. It also displays a tiny image, which is not present in the original article, but instead derives from a link to another article. So, in an article from BT.dk extra text is inserted into the reader view, because readability extracts it as part of the article’s main text content. This would likely confuse a user trying to learn danish because the text has no relevance to the actual article.

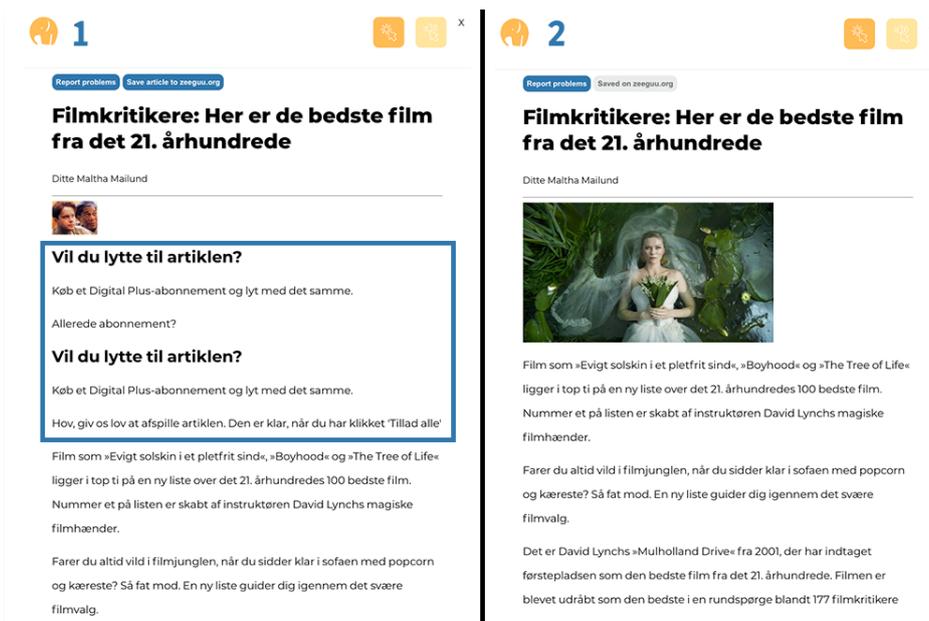


Figure 15: Section 1 displays the article in the reader view before individual cleaning (blue box showing the content wrongly extracted). Section 2 displays the article after individual cleaning.

We tested the article detection, content extraction, and layout on a selection of websites:

- The websites where most zeeguu.org users read articles recently (Appendix E.4: Most Recent Read Articles On Zeeguu.org and Whether Images are Fetched)

- Top 5 news websites in four countries (Appendix E.1: Top 5 News Sites in Denmark, France, Germany, and UK)

We prioritized cleaning the sites that looked the worst. We also tested a few of the websites which are part of the recommendations on zeeguu.org. During our experiment, the participants could also report problems through the reader. This led to the cleaning of multiple additional websites.

5.3.3.1. General and Individual

We use two types of cleaning. General cleaning, which is applied to all websites, and individualized cleaning, which is written for a specific website.

These cleaning methods are organized in `generalClean.js`¹² and `pageSpecificClean.js`¹³. For all websites, we remove links, figures, and SVG images and extract the main article image. Since this is often not sufficient, and the extracted text still has “noisy” elements, we have also implemented individualized cleaning functions for specific websites. Some of these are:

- removing text that is not relevant to the article (e.g. invitations to subscribe, suggestions of extra readings, etc.)
- removing contact info and comment sections

The extension cleans the pages by manipulating the DOM by using, for instance, `querySelector`, `createElement`, and `prepend`. We take the entire HTML content from a website or the cleaned content from Readability. When we clean websites individually, we mostly use attribute-based locators [3], which means that we depend on specific classes and ID’s of elements in the DOM. So, we look for specific tags, class names, id, or queries if we want to remove elements altogether or append new elements. In some cases, we also use content-based locators [3] to determine if a string contains a specific sentence; if it does, we remove it.

To make the process of cleaning websites individually more efficient, we have written a handful of helper-functions. These can be used to perform the more common cleaning tasks. Like using `removeAllElementsIfExistent`¹⁴, which as arguments takes the element we want to remove and the object where it is located. This can then be used on all websites, where we want to remove a specific class or id etc.

We clean websites in two different places; before and after we use Readability. Readability removes classes and ID’s from the website content. Thus, when we target and remove a specific class from a website, we must do it before using Readability.

¹²<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/Cleaning/generelClean.js>

¹³<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/Cleaning/pageSpecificClean.js>

¹⁴<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/Cleaning/util.js>

Lastly, we return a cleaned-up version of the document. Before we send this to the reader view and the database, we use DOMPurify to sanitize the content for security reasons [17].

Some of the cleaning errors arise because website owners do not always use best practices when building their websites. Because of this Readability and our cleaning functions encounter problems when website owners for instance:

1. Include the first part of the article text inside of a <figcaption> tag. Thus, when we during general cleaning remove all figures, the first part of the text is also removed.
2. Use tags to display headlines visually. The reader view does not know to format this as a headline.
3. Uses a <H2> tag for the first part of an article, even though it is visibly a paragraph text, as seen in the example in Figure 16. In this case the first seven lines of text are wrongly displayed as an H2 headline inside of the extension. With such a long text, this is not optimal.
4. Make use of an <article> tag around every comment written on an article. As seen in Figure 17, the Danish website Ingeniøren has a comment section on all their articles. Readability sees this section as the main text content, because the website owners have used an article tag here. The actual article text content is placed inside of a <div>. So, the reader view will only display the comments.

Éric Sadin : "Nous sommes dans une société où chacun est livré à lui-même"

Propos recueillis par Kévin Boucaud-Victoire
Publié le 28/10/2021 à 16:45

JE M'ABONNE POUR 1€

Comment reprendre le contrôle de nos vies ? C'est la question que pose Éric Sadin, dans son dernier ouvrage, « Faire sécession ». Pour le philosophe, connu pour son analyse fine de la numérisation du monde, il n'y a plus rien à attendre des élections, ni de l'insurrection alors que l'État providence se réduit comme peau de chagrin face au néolibéralisme. Selon lui, la seule solution est d'organiser une sécession collective. Rencontre... et explications.

Marianne : Qu'est-ce qui ne va pas, selon vous, dans la société actuelle ?

Éric Sadin : Nous vivons un moment fait de désillusions successives, d'impressions de saturation et d'inutilité de soi. Il devient, dans ces conditions, difficile de croire à l'ordre commun, aux institutions et aux structures politiques traditionnelles. En même temps, lors de chaque élection présidentielle, malgré toutes nos déconvenues passées, nous

1

Report problems Save article to zaeqaz.org

Comment reprendre le contrôle de nos vies ? C'est la question que pose Éric Sadin, dans son dernier ouvrage, « Faire sécession ». Pour le philosophe, connu pour son analyse fine de la numérisation du monde, il n'y a plus rien à attendre des élections, ni de l'insurrection alors que l'État providence se réduit comme peau de chagrin face au néolibéralisme. Selon lui, la seule solution est d'organiser une sécession collective. Rencontre... et explications.

2

Marianne : Qu'est-ce qui ne va pas, selon vous, dans la société actuelle ?



3

Report problems Save article to zaeqaz.org

Comment reprendre le contrôle de nos vies ? C'est la question que pose Éric Sadin, dans son dernier ouvrage, « Faire sécession ». Pour le philosophe, connu pour son analyse fine de la numérisation du monde, il n'y a plus rien à attendre des élections, ni de l'insurrection alors que l'État providence se réduit comme peau de chagrin face au néolibéralisme. Selon lui, la seule solution est d'organiser une sécession collective. Rencontre... et explications.

Marianne : Qu'est-ce qui ne va pas, selon vous, dans la société actuelle ?

Figure 16: Section 1 displays the original article. Section 2 displays how the extension formatted the article by default when using Readability. Section 3 displays how the extension formatted it after we cleaned the website.

De har rejst til Jupiter, Saturn og Neptun: Nasa lukker gradvist ned for Voyager-sonder



En fire år lang mission begynder for næsten 45 år siden, hvor to sonder skulle tage billeder af Saturn og Jupiter.

Af Nadja Josephine Arpe 22. jun 2022 kl. 11:37



Efter 44 et halvt års rejse er Nasa begyndt nedlukningsprocessen af sine to Voyager-sonder for at afslutte sonderens i forvejen forlængede rejse- og levetid, fremgår det af en artikel fra [Scientific American](#).

Tilbage i 1977 begyndte rejsen for Nasas Voyager 1 og 2, der skulle ud på den største ekspedition foretaget af ubemandede rumsonder.



#2 John Johansen 22. juni 2022 - 21:51

Pale Blue Dot

"Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilization, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every "superstar," every "supreme leader," every saint and sinner in the history of our species lived there - on a mote of dust suspended in a sunbeam.

The Earth is a very small stage in a vast cosmic arena. Think of the endless cruelties visited by the inhabitants of one corner of this pixel on the scarcely distinguishable inhabitants of some other corner, how frequent their misunderstandings, how eager they are to kill one another, how fervent their hatreds. Think of the rivers of blood spilled by all those generals and emperors so that, in glory and triumph, they could become the momentary masters of a fraction of a dot.



De har rejst til Jupiter, Saturn og Neptun: Nasa lukker gradvist ned for Voyager-sonder

Nadja Josephine Arpe



"Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilization, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every "superstar," every "supreme leader," every saint and sinner in the history

Figure 17: Section 1 displays the original article. Section 2 displays the comment on the article. Section 3 displays how Readability wrongly placed the comment inside the reader view of the extension

Besides our individualized cleaning functions, we also do some general cleaning and extractions when we display the article. With the function `interactiveTextsWithTags.js`¹⁵, we determine the HTML tags of the article content, and we decide what content should be clickable and therefore translatable. We save the tags and their text content in an array.

The function loops through all tags from the Readability text content and makes different alterations based on the tag. Thus, it also cleans up the document, e.g., by removing p-tags inside of header-tags and saving `` elements as children of `` elements. Inside `ReadArticle.js`¹⁶ we loop through the list of element objects and only display the elements we have chosen. In the current implementation, we display headers, paragraphs, and lists. Thus, we filter out all other HTML tags.

Another form of general cleaning is that we delete all timeouts and intervals on all sites¹⁷, because we found that website owners sometimes add content to their websites after running a timeout or interval.

¹⁵<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/Modal/interactiveTextsWithTags.js>

¹⁶<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/Modal/ReadArticle.js>

¹⁷<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/f52197cb50f8362d56a6102b8720d6ab3b8bbdb9/src/popup/functions.js#L68-L80>

5.3.4. Reused Components From Zeeguu.org

We reuse components from the codebase of zeeguu-react in our implementation. This codebase is available on GitHub at: <https://github.com/zeeguu-ecosystem/zeeguu-react>

For instance, we use components that are related to the translation and pronunciation of words, buttons, and exercises. This means that we keep the functionality and design from zeeguu.org, and we avoid writing redundant code.

To reuse the codebase, we added the zeeguu-react project inside our project in the src folder. However, we have made alterations by overwriting some design features, like colors, line heights, and buttons. We have also made changes inside the zeeguu-react codebase, for instance, improving the design of the Congratulations component¹⁸, which the user sees when they have finished the exercises. We also made changes based on our findings in the usability tests.

The extension cleans the user's article of choice while keeping the formatting and an image, whereas on zeeguu.org all formatting and all images are removed. Because we reuse components from zeeguu-react we do not need to make any changes to get the extension in-sync with zeeguu.org. All words clicked, number of exercises done, starred words, etc., will be registered automatically, so that users can see this information on zeeguu.org. We also log whether actions were performed inside of the extension or the website. These changes were made inside of zeeguu-react.

5.3.5. Build and Deployment

We encountered three main challenges during the build and deployment of the extension.

1. The first challenge was that when building a React project, all Javascript code is minimized into chunks of JavaScript and added into build/static/js. These are automatically generated and the names are not static. But for us to be able to inject JavaScript and CSS into a website when "Read article" is clicked by the user, we have to use the executeScript function¹⁹. To use the executeScript functionality, we have to define specifically which file we want to execute.

Because of this, we need to keep the Main.js file, which we want to execute, inside of the build folder without minimizing and chunking the code; thus it needs to remain static.

To make the file static, we change the Webpack configurations. In general, Webpack bundles JavaScript modules and generates static assets that can be used to serve the content. So, it takes multiple files and bundles them into a few files, which can be used to run the application [81]. In the standard configuration, React uses Webpack under the hood when building the application.

¹⁸<https://github.com/zeeguu-ecosystem/zeeguu-react/blob/development/src/exercises/Congratulations.js>

¹⁹<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/f52197cb50f8362d56a6102b8720d6ab3b8bbdb9/src/popup/PopupContent.js#L14>

But Webpack can also be used when we do not want to minimize the code. We use our Webpack configuration to define that “./src/JSInjection/main.js” should be put into the build folder without being bundled together with the other files. To do this, we create two different builds inside package.json: build:app and build:bg²⁰. The first one builds the project with React’s standard Webpack build configurations, but we also add “INLINE_RUNTIME_CHUNK=false”. We do this because inline scripts violate extensions’ content security policy. Running this line as part of our build tells React not to add an inline script to index.html. In the second build, we use our Webpack configurations when building. We choose not to eject the whole project because only one file, Main.js, is causing an issue. This way, we do not need to change the configurations of the whole application.

Before every build, Rimraf [36] is used to safely and consistently delete the previous build folder.

We use the Babel Loader[6] to transpile JavaScript files with Webpack. By default Webpack only reads Javascript and JSON files [80], and since we are using React, Webpack has to transpile JSX. We have used additional loaders to transpile CSS, styled-components, and SVG’s. They are defined in webpack.config.js²¹. By doing this, we can successfully inject code into a website.

2. The second challenge we encountered was related to project dependencies. To reuse components from zeeguu.org we placed the zeeguu-react repository inside of the extension’s repository. However, the React imports inside of the extension did not know which React project dependency to use. To solve this, we have to define, inside of webpack.config.js, that when running React in this project, we should resolve to use React from ./node_modules/react, instead of resolving to the zeeguu-react version of React.

3. The third challenge was that Chrome supports Manifest V3 and will soon stop supporting Manifest V2. On the other hand, Firefox only supports Manifest V2. This means that we have two manifest files, manifest.chrome.json, and manifest.firefox.json, but not a manifest.json file. Since the manifest file is required to be named “Manifest.json” to be published to Chrome and Firefox, we need to create a workaround.

We do this by creating manifest.json every time we use the publishing script²². When running pub-ci we publish the extension to Chrome, and when running pub-ci-firefox we publish to Firefox. In this process the scripts copy the information from the correct manifest into a new file manifest.json. In the same script we also pull the newest version of zeeguu-react, build the project, and zip the build folder. In this way, we only need to run one script when we need to publish a new version of the extension.

Additionally, we decided to automate the process further. We did this by setting up two CI/CD

²⁰<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/package.json>

²¹<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/webpack.config.js>

²²<https://github.com/zeeguu-ecosystem/ZeeguuExtension/blob/f52197cb50f8362d56a6102b8720d6ab3b8bbdb9/package.json#L46>

workflows in GitHub²³, which uses GH Actions to automatically publish new versions of the extension when something is pushed to the main branch. We created a workflow, `.github/workflow`, that runs a script when code is pushed to main, and these scripts are given our secret tokens to our Chrome and Firefox accounts. The scripts are found in `firefox.yml` and `main.yml` and they still use previously described `pub-ci` and `pub-ci-firefox`. In this way, the workflow is simply to push to the main branch, and both extensions get updated in the extension stores. The workflows and scripts can also be seen under “Actions” on GitHub, where we have two workflows, “Publish to Chrome store” and “Publish to Firefox Add-ons”.

²³<https://github.com/zeeguu-ecosystem/ZeeguuExtension/tree/main/.github/workflows>

6. Experiment

In total, 280 people signed up through the initial survey from April to June 2022. 17 of these people chose a native language we could not support, and 18 chose foreign languages we could not support. Despite this, two users still decided to participate in the experiment using other languages.

92 people made an account on zeeguu.org with the invite code. 50 people ended up using the extension, meaning that they opened at least one article and translated a couple of words. We had active participants in the experiment from April to June 2022.

6.1. New Users

6.1.1. Initial Survey Data

The data in this section is based on the answers that the 50 participants gave in the initial survey (Appendix C.1: Initial Survey Answers).

Out of the participants 50% were female, 44% were male, and the remainder answered: “Prefer not to say” or “Other”. Most participants were between 18-35 years old, but older age groups were also represented (Figure 18).

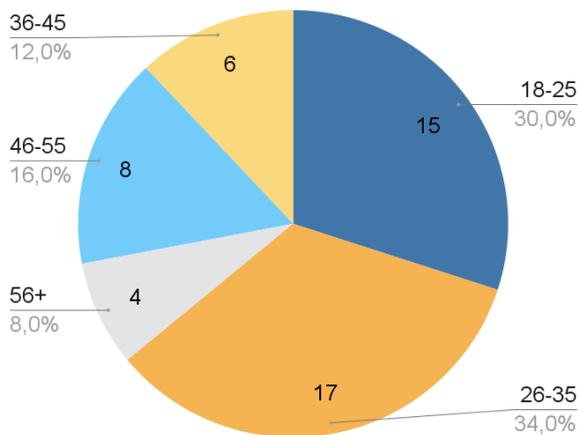


Figure 18: Age distribution of the participants

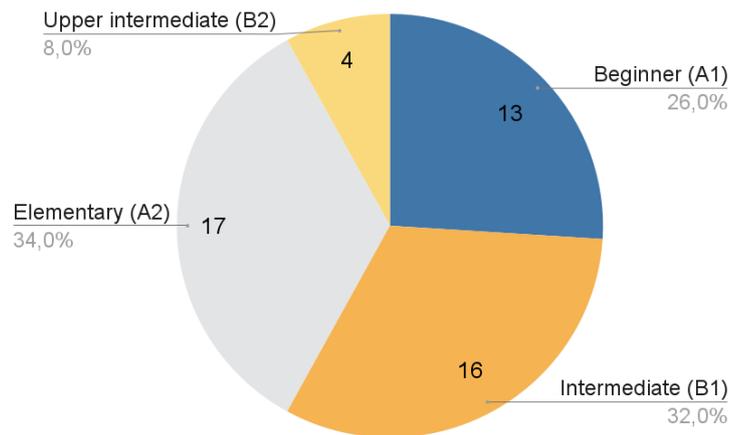


Figure 19: Level of the language the participants are learning

Out of the 50 participants, 60% were beginners (A1-A2), and the rest were intermediate learners (B1-B2) (Figure 19).

Figure 20 and 21 display the native and learned languages of the participants. The majority of participants have English as their native language. Norwegian, Dutch, and Danish were the languages that most participants were learning.

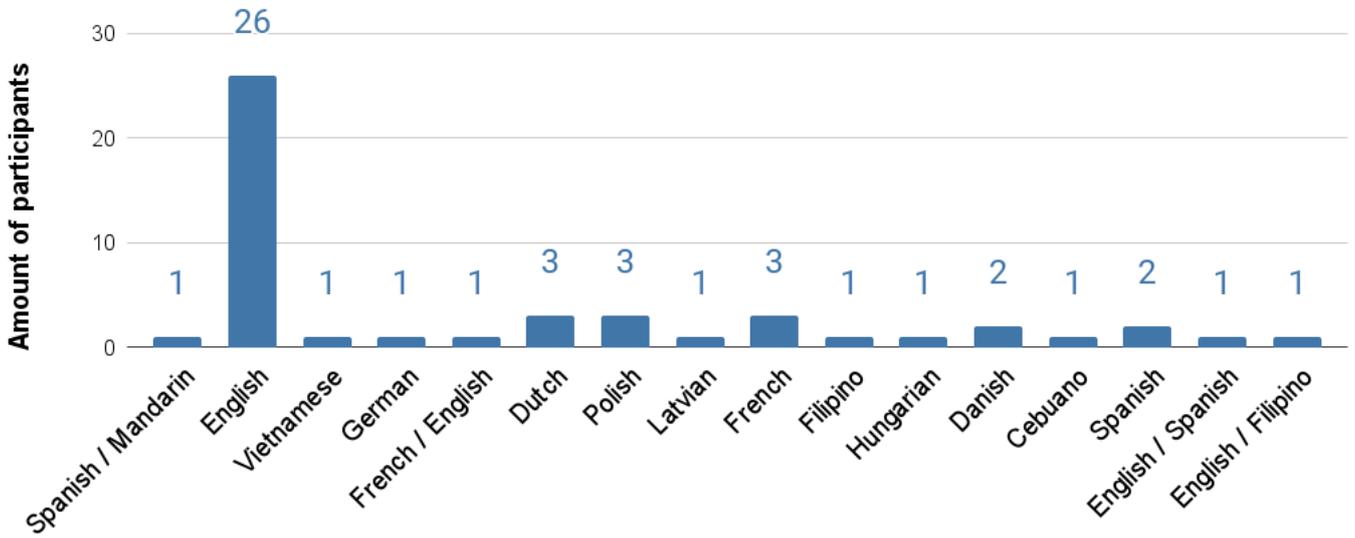


Figure 20: Native languages of the participants. Based on 49 participants.

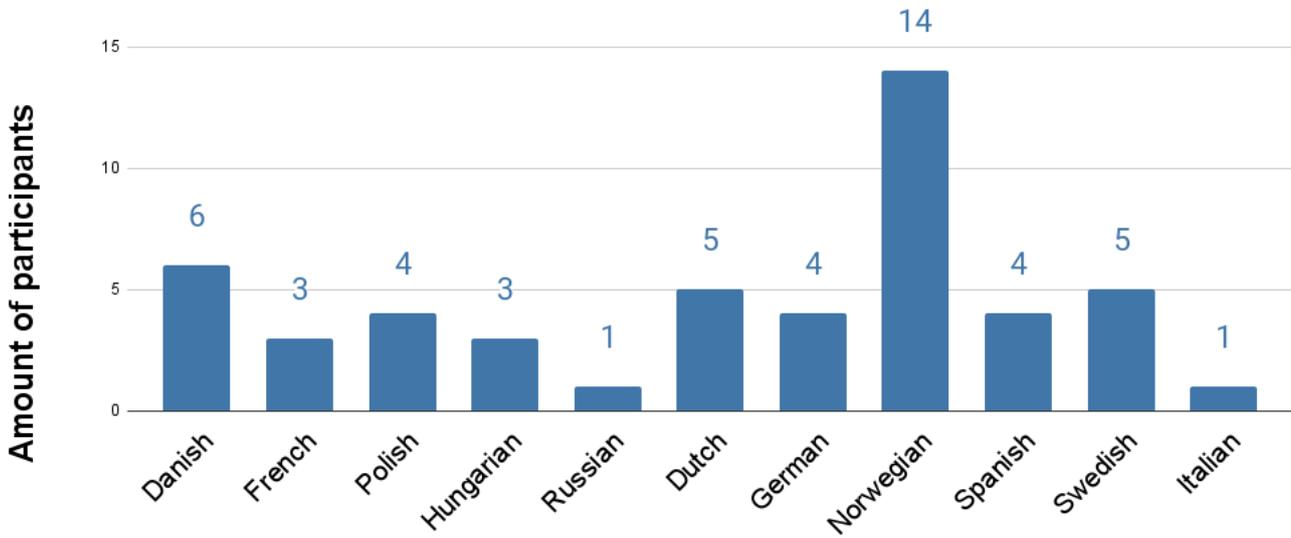


Figure 21: Learned language of the participants

The participants generally reported that they already spend a significant amount of time learning languages every week, where 56% spend between 1-5 hours a week (Figure 22).

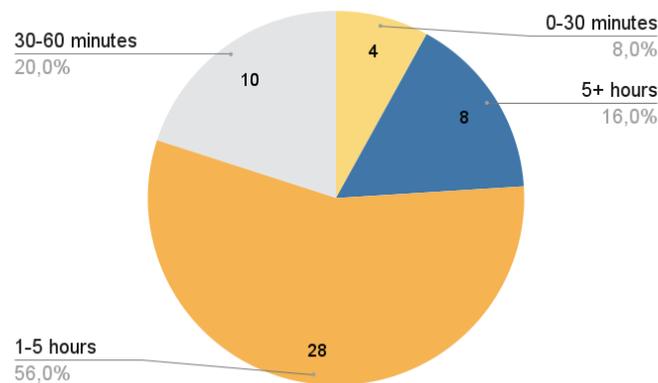


Figure 22: How long participants spend actively studying their language every week

We also asked the participants how they otherwise study languages and found that many participants had experience with online platforms, language courses, and in general, communicating with people in the foreign language. However, these were also the three predefined possibilities in the survey. The other answers were written as “other” in the survey, and we organized their answers into groups (Figure 23).

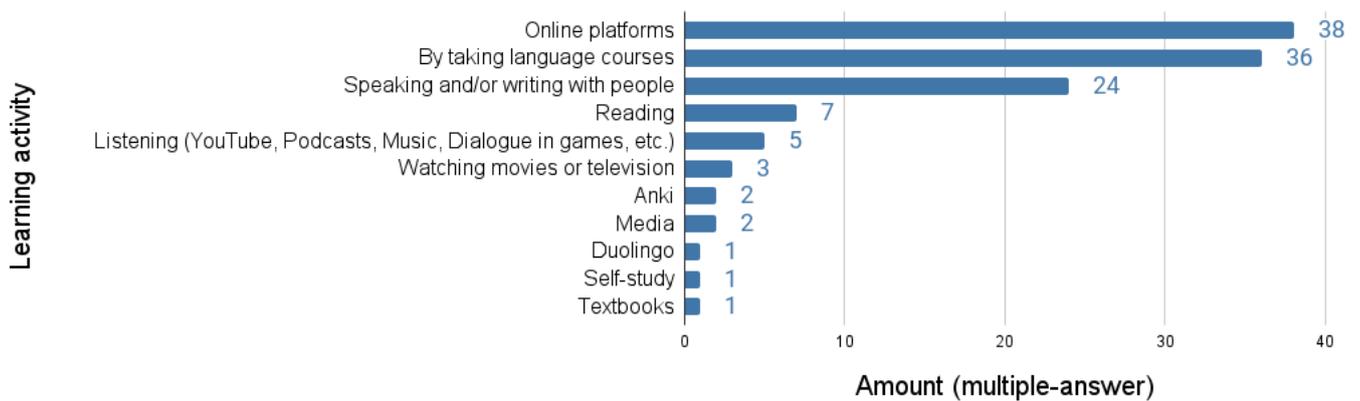


Figure 23: How the participants usually learn their language of choice

6.1.2. Findings From the Experiment

6.1.2.1. Overall Usage

Initially, we will present data on how much the 50 participants used the extension during the two weeks. We instructed them to use the extension as much as they liked, and we did not remind them to use it during the two weeks.

The following results are based on the logged participant activity. The data can be fetched from the database with the queries from Appendix D: Queries. On average, the participants opened 4.5 articles during the two weeks, the median being 3 (Figure 24).

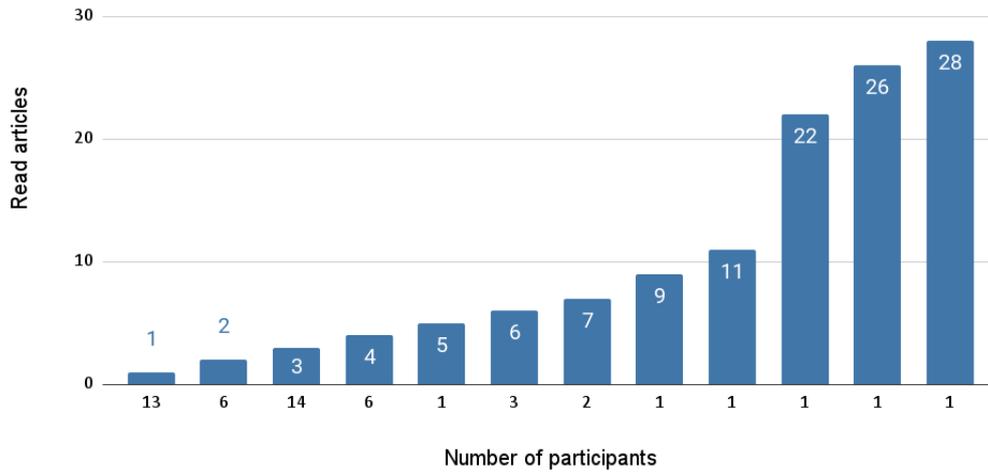


Figure 24: Distribution of how many articles were opened by the participants based on participant activity data

The participants used The Zeegu Reader for 2.7 days on average during the two weeks (2 is the median). The data is based on days with logged participant activity. One participant used it 13 days out of the 14 days in the experiment. 38% of the participants used it for only one day (Figure 25).

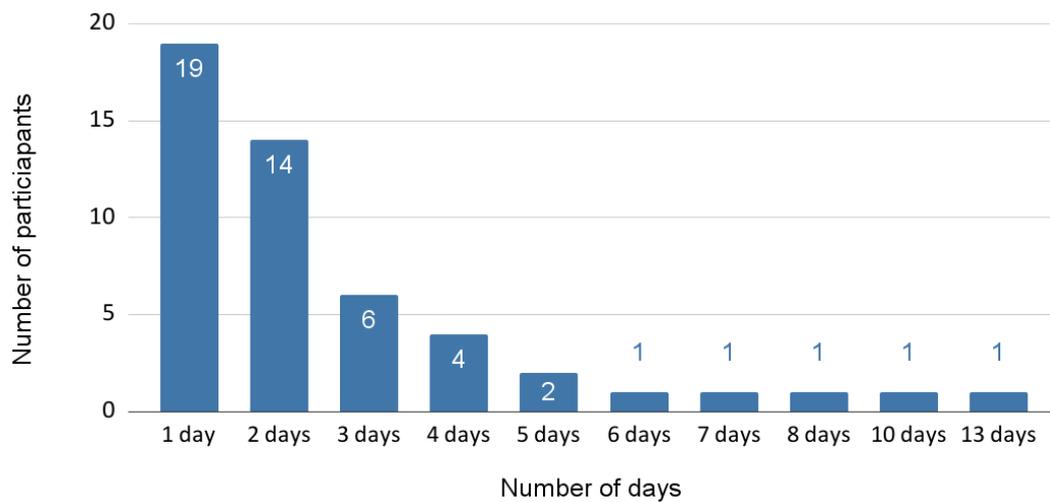


Figure 25: Distribution of how many days the participants used the extension

On average, the participants read for 37.7 minutes during the two weeks, and the median reading time was 18 minutes (Figure 26).

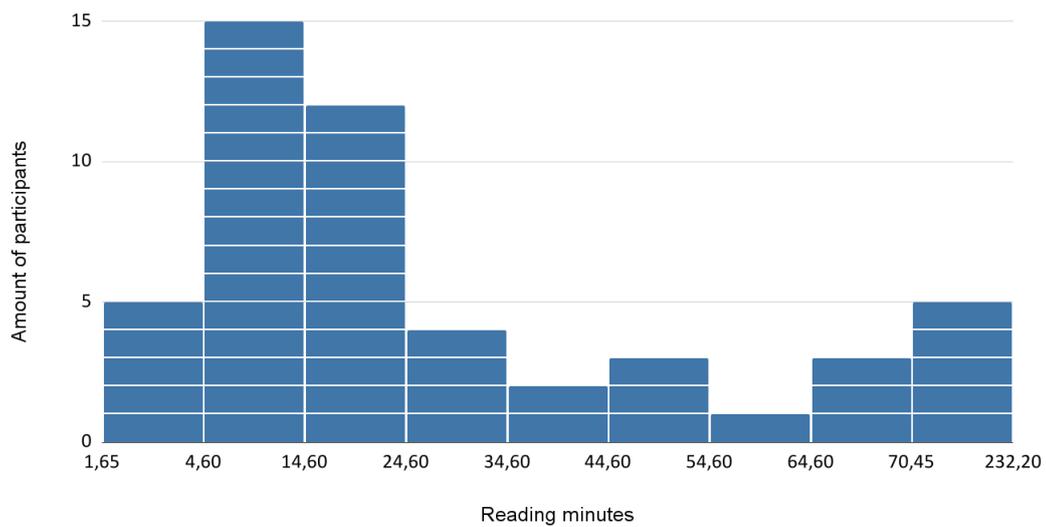


Figure 26: Histogram summarizing the participants's reading time in minutes

27 participants out of 50 did exercises. They did so for 7.3 minutes on average. The median was 3.7 minutes. Due to a bug in the database, data about exercises were lost during 4 weeks of our experiment. Because of this, these numbers are most likely higher, and this should be considered a lower-bound result.

6.1.2.2. Observations

32 out of 50 participants answered the final survey (for all answers see Appendix C.4: Final Survey Answers). The observations stem from the survey data, triangulated by data from the activity log. They are summarized in Table 1. Most quantitative questions were asked on a scale of 1-5 where 1 is negative, and 5 is positive. See Appendix C: Survey Answers for the answers to all the surveys.

Observation 1	Overall, respondents had an very positive experience.
Observation 2	The extension was easy to use for most respondents, but difficult to get started with for a minority.
Observation 3	Respondents experienced that the extension helped with their language learning in several ways and it contributed to their language learning in ways other learning resources do not.
Observation 4	Respondents liked that they could find their own articles and found mainly advantages to this.
Observation 5	The design and formatting of The Zeeguu Reader were slightly preferred over the original website, and they found it helpful for their learning.
Observation 6	The layout was not ideal for all articles, and some respondents were bothered by it.
Observation 7	Respondents who did exercises found them useful.
Observation 8	Respondents saw ways translations and pronunciations could be improved.
Observation 9	Respondents found the extension useful and generally said that it was convenient for practicing a foreign language.

Table 1: Summary of the observations found in this section

In the following, we will go through the 9 observations.

Observation 1: Overall, respondents had an very positive experience

Based on the survey data from the 32 respondents, they rated their general experience with The Zeeguu Reader 3.91 on average with a median of 4, where 5 is “very positive” and 1 “very negative” (Figure 27).

They elaborated that “it is an awesome tool for learning”, “i really like the idea of the extension”, “it’s a good learning tool”, “i loved the concept” and one saying it is “quite an innovative idea”. We also found that the majority of respondents would recommend the extension to others. They answered with 4.2 on average to the question "Would you recommend The Zeeguu Reader to others?", where 1 is "No, not at all" and 5 is "Yes, very much so". Thus, most respondents have had a positive experience.

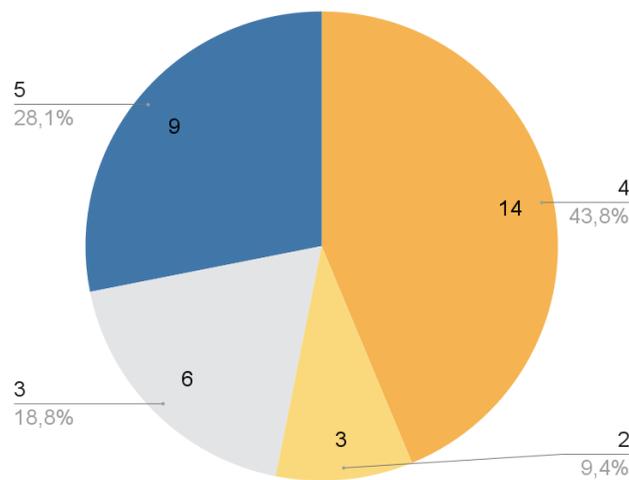


Figure 27: Answers on a scale of 1-5 to the question: “How was your general experience with The Zeeguu Reader?”. 1 is "Very negative" and 5 is "Very positive".

Observation 2: The extension was easy to use for most respondents, but difficult to get started with for a minority

The respondents reported 4.13 on average and a median of 4 to the question “How did you find the extension in terms of easiness?” where 5 is “Very easy” and 1 “Not very easy” (Figure 28). 14 respondents elaborated that it was easy to use, and three described it as intuitive. Moreover respondents described it as “easy to figure out”, “logical to use”, “just 2-clicks”, “user friendly”, “darned easy to use” and “doesn’t get much easier than that”.

However, for five, learning how to use it was somewhat complicated at first. One said that it was “Quite tricky at the beginning, but you will finally get the hang of using it once you have already started practicing by reading the article” and another said: “I was a bit lost on the very first time I used the extension, but it took less than 20 minutes to learn”. Thus, the description in the Chrome web store, and the GIF animation showing how to pin the extension, was not helpful

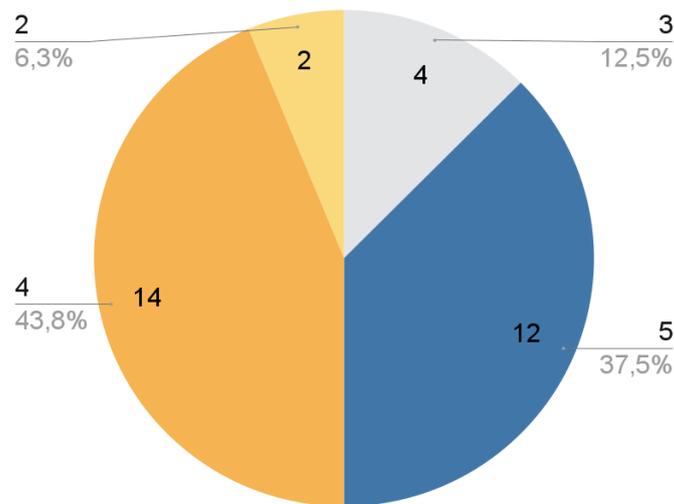


Figure 28: Answers on a scale of 1-5 to the question: "How did you find using the extension in terms of easiness?". 1 is "Not very easy" and 5 is "Very easy".

enough for some respondents. However, the actual user interface of the reader was intuitive and easy.

Using an extension was also foreign to 18 out of 32 of our respondents. The four people who wrote in the survey that they struggled to get started, also reported that they had never used an extension before. To combat the problem, we later added an instruction video to zeeguu.org, where we explain how the extension works in detail²⁴.

Observation 3: Respondents experienced that the extension helped with their language learning in several ways and it contributed to their language learning in ways other learning resources do not

The respondents answered with an average score of 3.97 and a median score of 4 to the question "Do you feel like the extension contributed to your foreign language learning?".

Two respondents highlighted that The Zeeguu Reader helped them learn more naturally. One respondent said that it "helps you practice more and learn new stuff in a very natural way" and another said that it "induces learning naturally".

A respondent also noted that it helped with learning everyday terminology: "It helps with new vocabulary and new grammar constructions, also gives some 'real' expressions or colloquial phrases within the articles that most learning courses don't have". Another found that "It made a lot more native content feel approachable".

²⁴<https://vimeo.com/715531198>

The respondents said that it was a great feature “being able to select multiple words to get context” and that the extension “Helped in reading large passages”. One said “I learned a lot of new words in a short space of time”, so it also contributed to their vocabulary based on their own assessment.

62.5% said that The Zeeguu Reader contributes to their language learning in a way that other language learning resources do not, while 25% did not find that it contributed differently and 8.3% did not know.

They highlighted that “It offers a way to practice that is different from what learning apps commonly provide, and seeing how the words are used in real life examples such as news articles helps in the retention”, “Other learning resources often put stress on learning vocabulary in an artificial way”, “The reader allowed for learning through longer texts compared to single sentences or short paragraphs that other learning services provide”, “being able to read larger pieces of text mostly uninterrupted by having to look elsewhere for a translation should help one acquire a language better”, and that “being able to read new stuff and articles related to something real and about topics that I am interested in”. They also said that “Most resources don’t make you read real articles”, and “It helped give me confidence to tackle a larger piece of text whereas other tools help with short sentences”.

Observation 4: Respondents liked that they could find their own articles and found mainly advantages to this

The average answer to the question “Do you like that you get to choose any article to read from the web yourself?” was 4.75, and the median was 5. Thus, the respondents liked that they could find articles to read themselves (Figure 29).

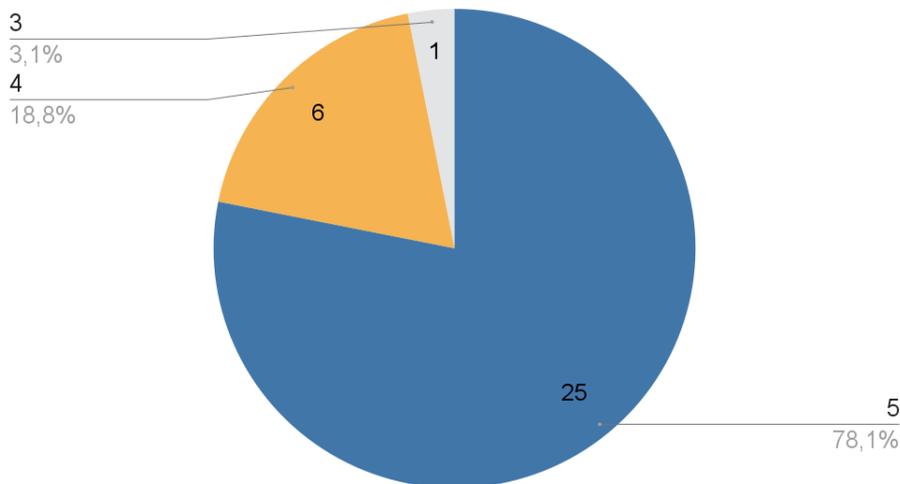


Figure 29: Answer on a scale from 1-5 to the question: “Do you like that you get to choose any articles to read from the web yourself?”. 1 is "Difficult" and 5 is "Easy".

In general, they also found it easy to find their own articles. The median was 4, and the average

was 3.75 to the question “How was it finding foreign-language articles to read with the extension?”, where 1 was difficult, and 5 was easy. From the qualitative answers, it is apparent that this could be because many used the recommendations on zeeguu.org. Some said it was easy because “you suggest articles in the page”, “List provided in the extension” and “The pre-set preferences about the type of articles you are interested in is a great idea”. On the other hand, some also mentioned that “It is easy to find articles from internationally known websites” and “I just went to foreign news websites and read articles there”.

To learn where the participants found articles, we checked whether they found articles through the recommendations on zeeguu.org or by themselves. We define that an article is found through the recommendations, if it has an rss-feed id in the database (Appendix D.1: RSS Feed Data). Thus, there could be cases where participants have found articles themselves that were already fetched from an rss-feed, and therefore the number of articles found without the recommendations could be higher. So, this is a lower-bounds result. 21 participants only used other websites. 17 participants used both other websites and recommendations. 12 participants only used the recommendations (Figure 30).

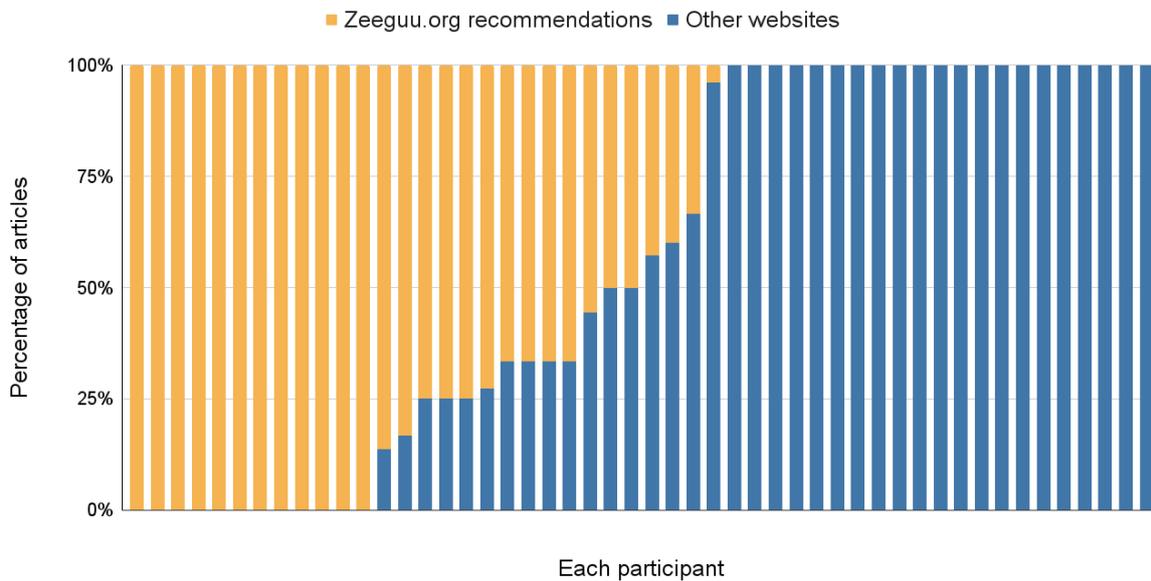


Figure 30: Where participants found the articles they read. Based on participant activity data, grouped by participant

From the survey, we could see that 22 respondents said they found articles through the recommendations, 17 found articles on websites they knew beforehand, and 14 found articles through a search engine (Figure 31).

To the free text question “Did you find articles to read that was of your interest? If yes, how did you find them?” 22 people answered that they did find articles of their interest, while six answered that they failed in finding articles of their interest, and four did not answer.

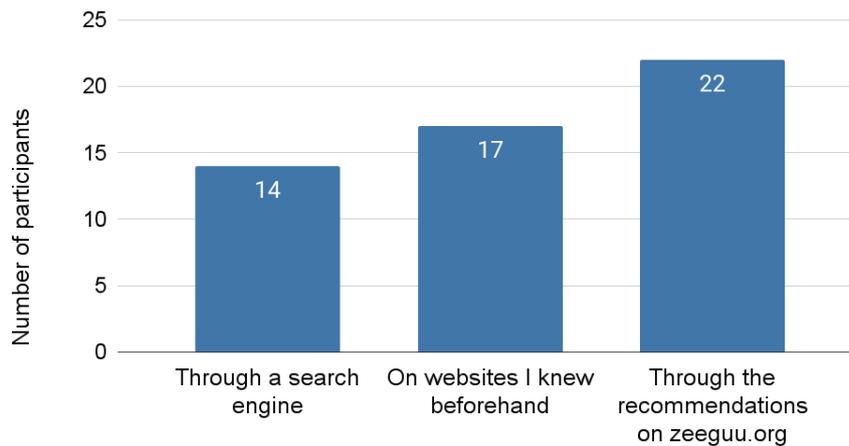


Figure 31: Answers to the question: "How did you find articles to read with The Zeeguu Reader?" They could choose more than one answer.

To the question “Was it a problem finding articles of the appropriate difficulty for you?” eight respondents said that this was not a problem at all, but another 10 respondents said that it was indeed difficult. The respondents that found it difficult elaborated that “Yes, but I am only a beginner so that is expected” and “Yes. News articles are usually intended for native speakers so there aren’t many that exist for beginner levels that aren’t too childish or boring”, while another respondent expressed that it was “A little bit” difficult. However, multiple respondents misunderstood the question. They thought it was regarding the meaning of “levels” inside of zeeguu.org and said that “I didn’t really know what levels mean” and “It doesn’t appear that the way in which articles are suggested for your level is quite right”.

Respondents also found several advantages to finding their own articles, for instance: “It helps me find something that suits my interest and will help make learning a language more interesting and engaging”, “Encourages me to read more”, “having fun just by reading exciting pieces”, “If you come across an article at random you can always read it, so there’s more freedom”, and “It motivates me more to read”.

To the question “What were the disadvantages of finding your own articles?” nine respondents answered that they found no disadvantages while six respondents found disadvantages like “having to find articles that worked”, “not easy to find articles of my interest”, and that the articles they found “could have been at any level of difficulty”.

We asked two questions regarding the easiness and difficulty of the article’s language, but looking at the qualitative elaboration from the respondents, we could see that they were misunderstood. We saw that people generally answered “3” regardless if they experienced great difficulty or found no problems. Because of this, we have not considered these two questions; instead, we looked at the qualitative answers.

Observation 5: The design and formatting of The Zeeguu Reader were slightly preferred over the original website, and they found it helpful for their learning

50% of the respondents preferred The Zeeguu Reader, 25% had no preference, and 25% preferred reading on the original website (Figure 32). Because this question was not part of the pilot survey, it has only been answered by 28 respondents.

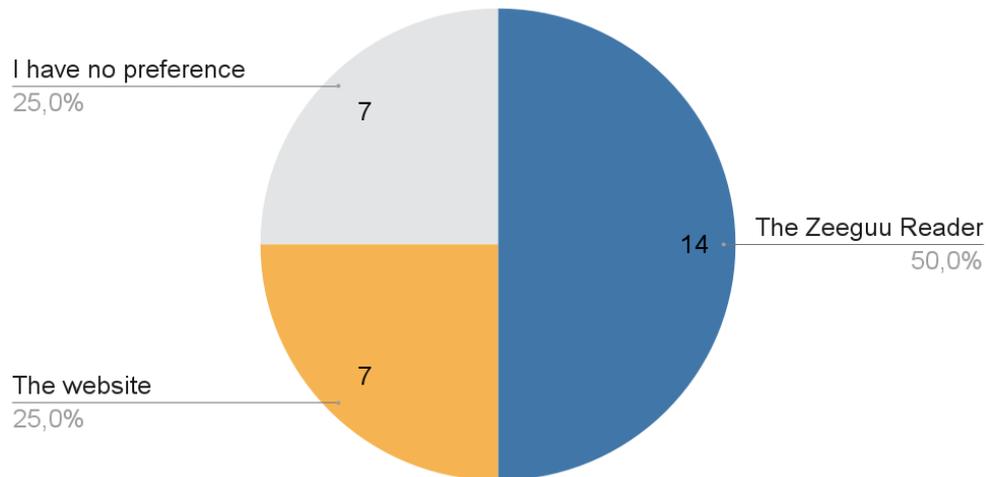


Figure 32: Answers to the question: "Do you prefer reading articles with the design and formatting of the The Zeeguu Reader or with the design and formatting of the original website". Respondents could choose one answer.

They found that the advantages of The Zeeguu Reader were "Simpler design, less clutter on the page", "much more simple therefore easier to concentrate on", "removed all the advertisements, hyperlinks and focused solely on the text I wanted to read", "no adverts to confuse the screen", "Fewer distractions allowed me to concentrate on the text", "Ads distract you from the reading" and that "The design and formatting keeps me aware that I am in a learning environment".

On the other hand, one found that it "would be nice just to open article I want to read in the website and read it without doing some more actions to get article in Zeeguu". Three respondents were also bothered that videos and images were often removed from the articles, saying that "Images/videos are important to an article", "it also eliminates the images that sometimes are helpful when reading very long texts" and it "removes too much of the pages formatting like videos and images". One even said that "(...) seeing ads in other languages can also help with learning".

The respondents who did not have a preference said that "My focus as a reader was only on the article itself, and it did not matter to me how it was presented" and that it "isn't really important what format".

On the other hand, we also saw that the question "Did you find it helpful for your reading

experience that the extension only displays the article text and maybe an image - and not, for example, adverts, buttons, and links?” yielded positive feedback. The median response was 5, and the average was 4.34 (Figure 33).

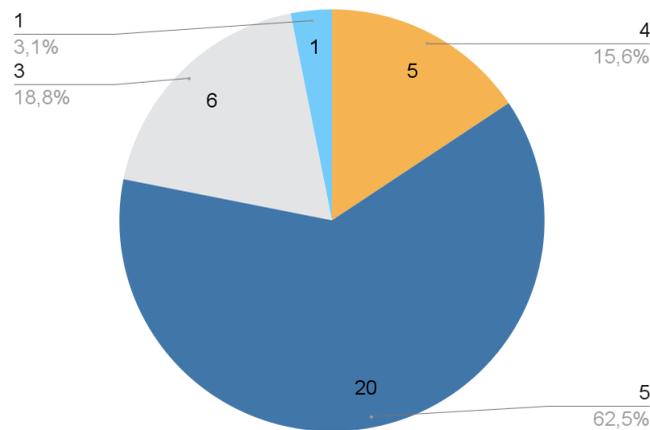


Figure 33: Answers on a scale of 1-5 to the questions: “Did you find it helpful for your reading experience that the extension only displays the article text and maybe and image - and not for example adverts, buttons and links”. 1 is "Not helpful" and 5 is "Very helpful".

13 respondents elaborated that it was helpful because it removed distractions. One noted that it was an advantage that they “(...) can only focus on the article itself” when the clutter is removed, and others wrote that it “is very helpful in the sense that it lets you focus on just reading the article without any distractions, or without the possibility of accidentally clicking the wrong buttons”, “Reduce distractions and increase my focus“, “allowed me to concentrate on the text” while another said “I struggle a lot with focusing so this was wonderful”.

Observation 6: The layout was not ideal for all articles, and some respondents were bothered by it

Six respondents noticed the absence of images or experienced formatting issues. Respondents generally found it helpful that clutter on a page was removed, but those who did not find it helpful said that “It removed too much”. Even though The Zeeguu Reader adds more of the original formatting than zeeguu.org, several respondents thought that some formatting was still missing: “(...) there is a problem with the paragraph headings”, “My biggest gripe was with how it tended to mangle page layout”, “this one [The Zeeguu Reader] often formats the articles poorly”, and “It’s nice there are no ads, but it does miss some of the style”.

Thus, they experienced the unintentional issues with formatting in The Zeeguu Reader, which arise because Readability’s text extraction from websites is not always sufficient enough.

Observations 7: Respondents who did exercises found them useful

66.7% of the respondents did exercises and they found them very useful. To the question “Did you find it useful that you could do exercises with the translated words after reading an article?”

the median answer was 5, and the average was 4.55 (1 being "No, not useful" and 5 being "Yes, very useful").

They found that the exercises were "(...) very helpful in checking what I've learned from practicing by reading the article and in retaining that knowledge", "one of my favorite features", and that it "made me understand the words I had difficulty with".

Observation 8: Respondents saw ways translations and pronunciations could be improved

Some improvements were requested regarding the translation and pronunciation of words.

Five out of 32 respondents expressed dissatisfaction because they had to click each individual word in a sentence to translate in its entirety. One said that it would be "more convenient if the user can translate by highlighting the phrase instead of clicking per word", another said that it was a disadvantage "Only being able to click one word at a time", while another expressed that it was "Quite tedious if you click the words", so it "would be better to get the translation of several words at once". Two respondents also experienced problems in regard to translations, because only one word in the sentence is translated, even though it is connected with another word: "[it] struggles a bit with phrases as selecting only one word in a phrase can cause it to translate wrong" and "an effective translation only appears when you click the whole of a clause rather than the individual words".

On the other hand, one person mentioned this as their favorite part of the extension: "Being able to click on multiple words at once to highlight a phrase or a sentence to get a better translation is my favorite feature".

Four also found that translations were not always correct and five found that pronunciations could be improved: "occasionally gave me some mistranslations", "there are some linguistic peculiarities which online translators all seem to struggle with", "can improve when it comes to translating and pronouncing numbers", "A few words were pronounced wrongly".

On the other hand, another respondent wrote that what they liked about the extension was "The ease of use, and how relevant the translations are. I also noticed how it can handle expressions and set of words very well", and "The ability to select from multiple possible translations".

Observation 9 Respondents found the extension useful and generally said that it was convenient for practicing a foreign language

To the question "How convenient did you find the extension for practicing a foreign language?" the average answer was 4.1, and the median was 4. They elaborated that "The quick translation and reading aloud features made it convenient to quickly learn new words or phrases", "The add on was always there when I wanted to read (...)", "The accessibility was the key", "I like the convenience and accessibility of it", and "I liked the convenience it provided where I no longer have to open multiple tabs for translating".

One also said that "It reduces the hassle that one usually has to go through when encountering

unknown words (...) no need to look it up somewhere else. Also, the pronunciation help is very handy. But what I liked the most was that it collected the new words & expressions for me and prepared the exercises to learn and practice them. Truly great idea”. Thus, the in-text translations and pronunciation were valued by multiple respondents.

6.2. Pre-existing Users

10 pre-existing users used the extension, and four answered the survey (Table 2). Thus, we do not have as much data from the existing users; therefore, our data is not as reliable.

ID*	Learned language	Native Language	CEFR-level (self selected)
51	Dutch	Romanian	B1
52	Dutch	German	B2
53	Dutch	English	B2
54	Danish	English	No level chosen

Table 2: Pre-existing zeeguu.org participants. *ID is referring to the id's in the provided excel sheets with data²⁵

Three out of four participants were part of a Dutch class, where the teacher instructed them to read one or two texts each week. Due to cancellations of their classes, they used the extension for three weeks. The participant learning Danish, was an Academic Language Consultant at a university.

6.2.1. Findings From Experiment

6.2.1.1. Overall Usage

The four pre-existing users that answered the survey read 90 minutes across five days on average (Figure 34).

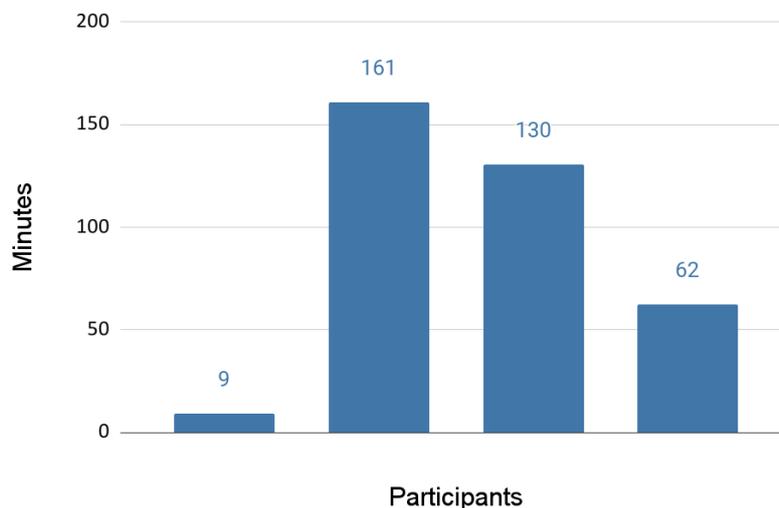


Figure 34: The reading time in minutes of the pre-existing users during the experiment

6.2.1.2. Observations

Observation 10: The pre-existing users liked that they could find articles on the web

Overall, their general experience with The Zeeguu Reader was, according to their own reporting, better than with zeeguu.org. To the question “How was the general experience using the extension compared to using zeeguu.org” where 5 is “better experience than zeeguu.org” and 1 is “worse experience than zeeguu.org” three responded with 4, and 1 responded with 5.

Two respondents were indifferent to the design and format change of the articles from zeeguu.org to the extension, one liked the extension better, and one liked zeeguu.org better.

They also all liked that they could find their own articles to read. They highlighted that the advantages to finding their own articles were: “I could read topics that I am interested in”, “more variety” and “I can translate anything no matter the topic (...)”. Only one out of four mentioned a disadvantage: “I was overwhelmed and did not know which one to choose”.

6.3. Post-Experiment Activity

After the experiment, by August 24th, 2022, 14 participants have continued to read with the extension (Figure 35). These participants all finished the experiment between May and mid-June. They were all active in the month following the experiment, and three used the extension beyond that. Eight participants used it only for one day after the experiment. Three participants used it for more than 1,5 hours. The three participants used it for:

- 1 hour and 48 minutes distributed onto 7 days
- 2 hours and 13 minutes distributed onto 4 days
- 2 hours and 48 minutes distributed onto 7 days

Thus, 14 participants out of 50 have read with the extension without the experiment as an incentive.

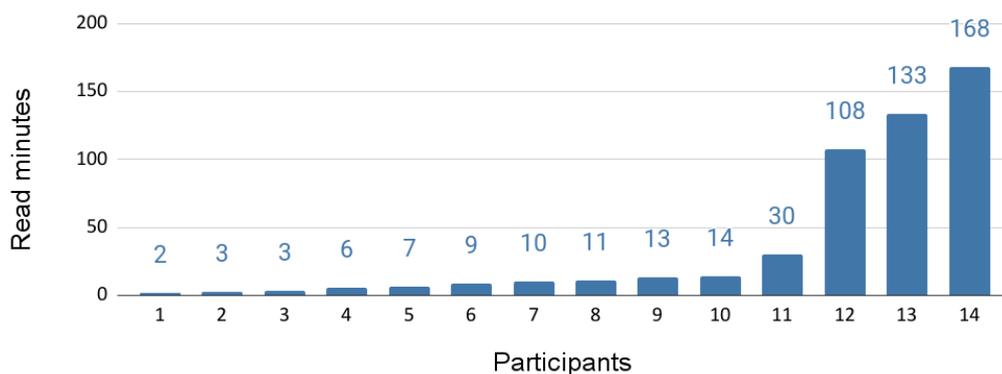


Figure 35: Reading time per participant from experiment finish to August 24th, 2022

6.4. Summary

Overall, the respondents had a positive experience using The Zeeguu Reader. They found it convenient for language learning and easy to use. They especially liked that they could find their own articles to read. The majority found it very convenient and useful while it also contributed to their foreign language learning in ways that other alternatives do not. Some of them appreciated practicing their newly encountered vocabulary with exercises immediately after reading a text. A slight majority preferred reading with the reader's layout compared to the original website. They also liked that ads were removed, but some were bothered by the formatting in regard to unintentional errors and design choices, while some also disliked that images were removed. A minority had problems getting started, and multiple criticized the quality of the translations and pronunciations.

Lastly, we found that 14 people continued to use it after the two-week experiment ended with at least 3 using it significantly.

7. Discussion

The number of participants, 50, and respondents, 32, in the survey only provide us with the experiences of a minor group, who are specifically interested in language learning. We also presented Zeeguu as a system where you learn by reading. Because of this, we could also have a bias, because we have mainly tested the extension with people who prefer reading as opposed to people who prefer using videos, textbooks, or other language learning methods. To be able to draw generalized conclusions, we will discuss the findings in relation to other studies and academic literature that touch on the themes in our observations. Thus, we will discuss the layout of the reader view, self-selection of reading material, and the maintainability and scalability of the extension.

7.1. Reader View

In this section we will discuss our findings related to three aspects; images, ad removal, and legibility.

Even though a slight majority of our respondents preferred reading with The Zeeguu Reader (Observation 5), some preferred reading the article on the original website. When we asked the usability test users about the design, they were mostly very positive, especially about the removal of ads. Most respondents also agreed that removing ads was a good aspect of the extension because they were less distracted, and it looked nicer. However, the page layout was sometimes not ideal (Observation 6).

One way to remove these issues altogether would be to activate the extension on the web page itself instead of opening it in a reader view. This was suggested by one of the usability testers. He mentioned that the extension could be improved if “Instead of having it be activated and isolate the article, then if I could just browse directly in it. So, if it was already activated I could just press a word directly in the article” (Usability test 8). However, he also rightfully mentioned that this could conflict with the website’s navigation and the ability to click on links.

By staying on the webpage we would have to exclude, for instance, links from being translatable. It could potentially confuse users that all words are clickable, but some words are translated when clicked, and others send users to a new URL when clicked. It seems the only alternative would be to remove all links from a webpage. However, removing the possibility to navigate a website does not seem like a possible solution, as this would be a very intrusive act, though, this is what the previously mentioned extension Vocab tracker does.

Even though seven out of 28 respondents preferred reading the article on the original website, most of our respondents liked the reader view or had no preference. Thus, we will now discuss the opportunities and challenges met when using a reader view.

7.1.1. Removal of Adverts

One important aspect of opening articles in a reader view is the removal of ads. Though some studies have found that people can ignore ads [74, p. 175], Simola et. al. also concluded that “online ads are not ignored during reading, irrespective of whether the task was engaging (reading for comprehension) or free browsing” [74, p. 188]. However, they also found that the users were more affected when free browsing than when they read for comprehension [74, p. 188]. They also found that ads located to the right of the text, and abrupt onset of ads in some cases would capture the attention of the reader [74, p. 188]. The Zeeguu Reader works on many different websites, and therefore also on websites that are monetized. Thus, ads will sometimes appear in these ways and positions.

Another study found that adverts used as in-text images can distract readers [8, p. 458-459]. In our case, we did see examples of adverts for news sites’ own products, like subscriptions, inside of article texts because the Readability library does not detect these adverts. These kinds of adverts, which are not related to the article text, will distract the user, however, it will not impair the comprehension of what they have read [8, p. 459].

So, some people can ignore ads, and their reading comprehension might not be affected. However, we have some indicators that ads can affect the reader negatively, and nonetheless we have not found any negative aspects of removing ads, only possible positive aspects.

7.1.2. Legibility

Another aspect of opening the text with a different layout than the original website is the fact that we can alter the text presentation.

A study found that using a font size of 18px improved the reading experience for dyslexic people [66]. In general, studies on dyslexic people agree that “the application of dyslexic-accessible practices also benefits non-dyslexic readers” [66, p. 7]. They concluded that using font size 18px improved comprehension, readability, and fixation times.

Reader-friendly font sizes are not always used on websites; for instance, bbc.co.uk uses 16 px, bold.dk uses 12 px, and wikipedia.com uses 14 px. By displaying the article in our reader view, we can ensure that the text is always the appropriate font size.

Sans serif fonts have been shown to provide better readability and reading performance than serif font types [34, p. 172], but not all news sites comply with this. For instance, both lefigaro.fr and lemonde.fr, two of the biggest french news sites (Appendix E.1: Top 5 News Sites in Denmark, France, Germany, and UK), use serif fonts. So, opening their articles inside The Zeeguu Reader could improve the legibility.

Likewise, the line length impacts the legibility of texts, but not all news sites keep the line length within the recommended 60-75 characters per line [9, p. 139]. For instance, the German state-

owned international broadcaster Deutsche Welle, dw.com, uses up to 99 characters per line, while also using a small font size of 15, without additional spacing. A website text like this could also benefit from the formatting in The Zeeguu Reader.

Another study specifically looked into Firefox’s Reader View, and they concluded that the design alterations “(...) result in significant improvements in reading speed, perceived readability and aesthetics for people with and without dyslexia” [49, p. 11]. They found that reading speed improved by 5% in the reader view and that people with and without dyslexia benefitted the same amount from the reader view. They also asked their test users to rate the readability and found that the readability of the reader view was significantly higher. The test users found the reader view “clean” and “pleasant” compared to the website [49, p. 8], similar to our study.

So, by opening a reader view, where we determine the font size, line height, characters per line, and remove unnecessary clutter, we ensure that the reading experience is always consistent because not all website owners are aware of best practices concerning legibility. Thus, we would argue that even though some people did not prefer The Zeeguu Reader, the reader view should be improved rather than adding the functionality directly to the original website.

7.1.3. Images in the Reader View

We saw in observation 5 that some respondents were unhappy that images were removed. In other studies, it has also been found that pictures inside of text have a positive effect on reading comprehension [32], also specifically on foreign language reading with low-proficiency English learners [83]. Thus, images are an important part of the extension.

One part of the issue is that, in many cases, Readability did not retrieve the main image from the articles successfully. Because of this, we created individualized functions to fetch the main image from specific websites. This was cumbersome, as we did it for each website. When we ran the experiment, this was the approach deployed in the extension. After discovering that some respondents missed images, we made a bigger adjustment to the code.

First, we did a small experiment to test our adjustment. We looked at the article sources of the 5000 most recently read articles on zeeguu.org, and we checked if the main image was fetched when using the Readability library. We found that for 14 of 18 websites Readability did not fetch the main image (See Appendix E.4: Most Recent Read Articles On Zeeguu.org and Whether Images are Fetched). Given that a majority of our participants were trying to learn Norwegian, we also checked the top 10 Norwegian news sites (See Appendix E.2: Top 10 Norwegian News Sites). Here Readability did not fetch the image in five out of 10 websites. Thus, less than 50% of news articles had an image when opening in The Zeeguu Reader.

Because of this, we stopped using Readability to fetch images. Instead, we created our own function `getMainImage`²⁶. This function checks if the article has an image meta tag. In many

²⁶github.com/zeeguu-ecosystem/ZeeguuExtension/blob/main/src/JSInjection/Cleaning/generelClean.js

cases, news articles have this meta tag, because news sites often share articles on their social media accounts. If no meta tag is found, we check all images on the web page and filter based on specific limitations, like size and file type.

This approach is implemented in the current version of the extension, but was not deployed with the participants in this study. Readability only fetched 4 out of 18 images on the most popular news sites on Zeeguu, but the new function successfully fetched all 18 (Appendix E.4: Most Recent Read Articles On Zeeguu.org and Whether Images are Fetched). In Norway's top 10 news sites, where Readability only fetched images for 50%, we successfully fetched images for 100% (Appendix E.2: Top 10 Norwegian News Sites). So out of the 28 websites we have randomly checked using `getMainImage`, all had the main image displayed in the reader view.

But some respondents also wanted all images and videos included within the article text. However, we decided only to include one image in each article because we experienced that articles with many images became unmanageable and bulky, as the images took up a lot of space in the reader view. Instead, we wanted to focus on the reading experience without having too many images.

We have not evaluated if the changes affect the respondents' issues with missing images. However, images are now added to a significantly larger number of articles.

7.2. Self-Selected Reading and Interest

In the experiment, we found that many of our respondents were happy to find their own articles because it allowed them to read texts based on their own interests. The respondents said that it encouraged them to read more and that they had fun while reading. It made it more interesting, engaging, and motivating because they could find exciting pieces themselves (Observation 4).

Multiple studies support this observation. One study found that "individual interest is an important if not a critical factor of academic motivation and learning" [33, p. 202]. Asgari et al found that "interest can influence students' motivation for learning, involvement with the learning tasks, and educational achievement" [4, p. 69]. Similarly, Ainly et al conclude that "topic interest was related to affective response, affect was then related to persistence with the text, and persistence was related to learning" [2, p. 558].

Interest-based motivation has a positive effect on the outcome of learning [42, p. 32] and it elicits students' interest in learning [79, p. 142], while also enhancing immediate performance and long-term learning [79, p. 143]. So, not only are learners more motivated to learn, but the learning outcome is also better. Additionally, studies about children who are learning to read have found that self-selected reading creates a positive attitude toward reading and a greater proficiency [38, p. 197].

On the other hand, others have found that topic-interest was not as impactful as situational interest, and that "topic-interest may be more important to catch attention, whereas situational

interest may sustain attention in a manner that increases online engagement while reading” [27, p. 111]. Topic interest (or personal interest) means an interest that is stable and content-specific. Situational interest means an interest that is short-lived, context-dependent, and based on curiosity and emotions; thus, it arises spontaneously [27, p. 95].

Our experiment backs up other studies, which reveals that for many students, being able to choose their own reading material based on their own interest improves motivation, engagement and learning. However, many participants still used the recommended articles on zeeguu.org. The recommendations make it possible to choose reading material based on interests, but with a more limited selection of articles. One respondent expressed that it was “not easy to find articles of my interest”, two respondents did not find any advantages to finding their own articles, while multiple respondents expressed that they liked the recommendations on zeeguu.org.

A study found that self-choosing material is not optimal for every student and that teachers “(...) discover that it does not increase motivation in all cases or for all students (...) sometimes, choice may lead to less engagement, lower quality work, and even negative affect” [27, p. 111]. So, letting learners find their own material and having many options is not the best solution for some. Even though a study on students who are learning, also learning to write, found that the students who can choose themselves were “perceived to be more motivated and encouraged to write” [11, p. 390], some students favored not choosing their own material [11, p. 391]. The study goes on to suggest that the teacher should provide two options “either their own favorite topic or the one suggested” [11, p. 391].

One of the respondent in our study expressed that “I was overwhelmed and did not know which one to choose” in concern to finding articles to read on the web. Likewise, multiple people in our usability tests had difficulties finding articles without having the possibility to look at recommended articles. Thus providing suggestions on zeeguu.org seems to accommodate these issues to some extent.

By having these recommended texts, which can be based on personal interest, we still obtain personalization based on self-description [37, p. 198-199]. We ask the learner to choose their interests and to write their level in the language, and based on this information, they are presented with relevant material.

It could be argued that our experiment has a bias because the people we included in the study are all very interested in learning a language. This may not always be the case if the platform is to be used by teachers in schools. In these cases, having recommended articles will also make it easier for those who don’t know any foreign language websites and those not interested in spending time looking for material to read.

So, both the results of our study and previous studies support having both the option to choose from recommended articles and to choose articles freely from the web will embrace more language learners because some prefer finding their own material, while others benefit from having a limited selection of material.

7.2.1. Authentic Material and Text Difficulty

Letting learners select their own reading material with The Zeeguu Reader also allows them to read authentic material. Authentic material in this context means material “produced by native speakers for native speakers” [30, p. 98]. Authentic material can help “attract students’ attention and increase their motivation towards learning a foreign language.” [63, p. 334] because it allows the students to “comprehend the learned foreign language better and to use it as in real life.” [63, p. 334]. In our experiment multiple respondents mentioned, on their own initiative, that they liked that they could read real-life texts with “real expressions” and “colloquial” language, as this is normally not present in learning courses, and that the extension contributed to making native content feel approachable and helped their general retention (Observation 3).

But this approach also poses a challenge. Authentic material can cause negative feelings if the material is not of the correct difficulty [63, p. 334]. Therefore the foreign language teacher has an important role in not using random authentic material [63, p. 334], and instead, it should be carefully selected [30, p. 108]. Even though most of our respondents experienced no issues finding material of the correct difficulty, some encountered problems, which could be due to the lack of difficulty determinations when using the extension.

With Zeeguu’s current approach, there is not necessarily a teacher to provide texts, and users can choose mostly anything; thus, it could be argued that Zeeguu should help the user find texts of the appropriate level. Currently, the Zeeguu platform doesn’t help the user determine their language level, and even though zeeguu.org helps the user with information on text difficulty the extension does not. As a solution, the extension could determine the difficulty when opening the popup, so this could be displayed to the user directly, but it can also be a challenge to determine the difficulty of a text based on grammatical criteria [30, p. 108].

However, our approach also has an advantage for teachers. In earlier studies, it has been described that personalizing reading material for students is challenging for teachers. Learning based on individual interest has been seen as a “(...) very time consuming and effortful task” [33, p. 203] because the teacher must individualize the learning for each student, which makes it “(...) very difficult for teachers to accomplish personalization” [79, p. 167-168]. But The Zeeguu Reader can give teachers an advantage because the students themselves can personalize their reading.

Thus, self-selection of authentic reading material from the web has the possibility of motivating and engaging learners. Still, it should also be further investigated how to help users find appropriate text material of the correct difficulty. Given that currently, only zeeguu.org provides difficulty determinations, it could be advised that users who have trouble finding articles of the correct difficulty should mainly use the recommendations on the Zeeguu website.

Before users can benefit from authentic material in their studies, we must ensure that the foreign language websites work inside The Zeeguu Reader. This leaves us with the challenge of maintaining and scaling website cleaning.

7.3. Maintainability and Scalability

We experienced multiple issues regarding the maintainability and scalability of The Zeeguu Reader, both during development and the experiment.

We had to implement our own readability check and cleaning because the Readability library did not always correctly determine if a text was readable, and because it also did not always extract the text sufficiently, we had to implement our own readability check and cleaning. This gave us a maintainability problem because we had to clean each individual website, which as a result, posed a scalability problem. After all, we can not clean all text-based websites available in all the languages we support.

7.3.1. Detection of Readability

The first maintainability issue we encountered was the detection of readable articles. We use the detection method that Readability provides, but “Reader Mode has always been an unconventional and complex problem statement due to the lack of any published standards or guidelines.” [53, p. 1]. We experienced the lack of standards in Readability’s insufficient detection of readable articles. For instance, two respondents saw it as a disadvantage having to find "articles that would work with the extension". This led us to create our own `isReadable` function in addition to the detection that Readability provides.

Nayak, Thadichi, and Harshita investigated how to improve text detection and found that their Machine Learning Random Forest approach had a precision of 99% and a recall of 94%, whereas Mozilla’s Readability only had a precision of 83% and a recall of 90%. Precision is how many selected articles were actually articles, and recall is how many articles which should have been selected were actually selected [53, p. 5].

7.3.2. Individualized Website Cleaning

The second maintainability issue was the cleaning process. A big part of the cleaning process is handled by Mozilla’s Readability open-source library.

However, the library does not always work perfectly and we are left either with unwanted content in the extracted text or content that is removed that should have remained.

As a small experiment, we looked at the ten most-visited Danish news sites in April 2022 [28]. (See table overview of the ten sites in Appendix E.3: Most Visited Danish Websites, and Whether They Needed Individualized Cleaning). Among these ten sites, we deem that:

- Two of them were cleaned completely by Readability and our general cleaning functions.
- Three sites would need to be cleaned because of problematic errors like articles not being readable, scripts interfering with the reader, or because of substantial unwanted text. For instance, `Berlingske.dk` articles start with five lines of text regarding their subscription.

- Five sites could be improved by individualized cleaning. For instance, in some cases Politiken.dk references one or two other article titles at the bottom of the page. They would be nice to remove, but not critical for the reading experience.

Thus, most websites in this small experiment needed additional cleaning.

To clean a specific website, we must analyze the HTML of the website to determine which elements we want to remove, alter, or add when opening The Zeeguu Reader. This is not a scalable approach because we cannot analyze the code of every website on the internet. Instead, we chose to look at the most popular news sites in the languages that a lot of people wanted to learn. This also means we have only tested cleaning on a tiny percentage of supported websites.

Cleaning based on attribute locators leaves us with a maintainability issue. In general, “Web locators are fragile upon Web Content upgrades that can make extensions no longer pinpoint the right DOM element” [3, p. 6]. If we clean a specific website, and the host renames their classes or restructures their website, then our cleaning function will break. However, we will not learn this unless we are notified by users or actively go through websites daily to check.

Choosing the correct locators is also an issue in regard to test automation of web applications. Leotta et. al. define three kinds of locators: visual, DOM-based, and coordinate-based [46, p. 322-323]. Through an empirical study, they found that DOM-based web locators were more robust. On the other hand, visual locators based on algorithms were easier to repair because the visual appearance of an application does not often change as drastically as the underlined structure [46, p. 339].

The maintainability problems with locators have been investigated on multiple occasions [60], [3]. Aldalur and Diaz suggest the possibility of using regenerative locators, which improves the robustness of web locators [3]. The regenerative locator uses contingency data, using both ID, Name, TagName, etc., so when one locator fails, it can be restored by using the other saved information [3, p. 3]. However, they also conclude that even though this increments the unbroken lifespan of the sample browser extension by 70%, this approach is more relevant for single-site extensions [3, p. 6].

Nguyen, To, and Diep have looked into XPath, which is the most common locator. It uses absolute or relative paths to navigate through web elements [60, p. 1]. They looked into improving XPath, and found that multiple people have tried to make this locator more robust [60, p. 2]. Nevertheless, if the target attribute changes, XPath is still vulnerable no matter the improved solution. They then suggest using neighbor-based XPaths [60, p. 2]. But this approach is also vulnerable if the neighbor of a target element changes.

Thus, there is not yet a solution on how to make locators robust enough so that we can rely on them working at all times. If the host changes their website, it will likely affect the locators we target during cleaning. Likewise, XPath locators are generally considered fragile. Therefore, their creation requires effort and experience [47, p. 450].

Jung et. al. have developed an algorithm that extracts elements based on visual features [40]. They evaluated it on the top websites of seven countries and they found that it was better than several other extraction methods, including Readability. However, they do not provide an available implementation of the algorithm.

Another study investigated the use of multi-locators to improve the robustness of locators and found that “Overall, multi-locator (worst order) is able to outperform ROBULA+, the algorithm that produces the most robust locators, globally reducing the number of broken locator from 78 to 68” [48, p. 8]. Thus, our method could likely be improved by using, for instance, multi-locators.

It is also part of the maintainability problem that we currently have many functions to clean specific websites, but we do not know when these functions will become outdated. Thus, we could end up with a large amount of unused code because the specific website changed its code. In our short process, we did experience that code became unnecessary because of updates to a website.

For instance, we previously mentioned how the website Ingeniøren.dk had a comment section on their articles, which was seen as the main text content by Readability, because the comments were placed inside of an article tag. The problem was discovered and fixed in a commit in March 2022²⁷, and in June, three months later, we discovered that the issue had returned. The website had been changed, so Readability no longer saw all comments as the main text content. But then another issue arose. Readability saw the comment as the main text if the comment was longer than the article. In one instance the article text contained 353 words, but the article had received one comment, which was 368 words. Because of this, Readability extracts the comment, and it was displayed in The Zeeguu Reader. Thus, a new fix must be written to remove this issue.

There is generally no agreement in the literature on the best way to locate elements. Though finding a robust locator does not seem to be the solution to the maintainability problem and the scalability problem that arises from it. It is impossible to clean all websites, even if locators are robust. It seems more likely that a solution should be implemented to improve the Readability library or to find or develop a better DOM extractor tool, so more articles work without individualized cleaning.

7.3.3. Cross-Browser Extension

To accommodate more users we made the extension available in multiple browsers: Chrome, Firefox, Edge, and Opera. However it is a challenge to maintain extensions across browsers.

Others have previously stated that one of the main problems for extensions that implement web annotations is maintainability regarding browser upgrades and, in particular, host upgrades [21,

²⁷<https://github.com/zeeguu-ecosystem/ZeeguuExtension/commit/09b2967573cf48da613518332af6f28f1f86c1dd>

p. 16-17]. Dealing with multiple browsers you also have to deal with different browser upgrades.

We experienced this challenge because Chrome and Firefox currently use different manifest versions. This means that when Firefox moves to Manifest V3, like Chrome, the Firefox manifest has to be updated. The move to Manifest V3 does not only mean that the manifest must be changed, but it also affects functionality and code in other files. Thus, maintaining cross-browser extensions will pose maintainability challenges.

8. Reflections, Limitations, and Future Work

8.1. More Reading Material at the Expense of Access

The solution we have implemented greatly expands the number of texts users have access to. But we are also limiting who can use the Zeeguu platform, because the extension can only be used in desktop browsers that support Firefox or Chrome extensions. If the user is using Safari, an iPad, or a smartphone, they can no longer use Zeeguu. Since iOS 15 Safari extensions also work on iPads and iPhones [19], building a Safari extension would increase the target audience. So, to reach the same number of users as before, we would need to create an Android app and a Safari extension.

Another problem we faced was that some of the people who signed up in our initial survey did not know that extensions only work on desktop browsers; thus, they thought they could participate using their phones or iPads. Therefore, the technological knowledge of users could be a barrier for potential users.

8.2. Increasing Personalization

Multiple respondents wished for more images and even videos in the reader view. We could include all images for some articles without overpowering the text content. Still, for other types of texts, like blog posts, which sometimes contain many images, this can easily look messy.

A solution could be to increase personalization and let the user decide if they want images and videos displayed in the articles. Firefox's Reader View currently allows the option of deciding the font size, type, color, etc. This could also be implemented in the extension. Thus, we could make the user decide themselves if they want to focus on only the text or if they want to include all original images and videos in the article while also giving them the option to choose font sizes and colors.

8.3. Translation Accuracy

We learned from the survey that four respondents were not satisfied with the translations and described them as “(...) somewhat acceptable (sometimes it's just plain wrong).” and that “It occasionally gave me some mistranslations”.

Another challenge regarding language is that we ourselves are not able to assess the quality of the translations in most of the languages we offer because we do not understand the languages. That is one of the dangers of adding new languages to the system without being able to assess if the translation tools are reliable.

8.4. DOM Extraction and Article Detection

Improving text content detection and extraction would improve the extension's maintainability, but it is also the most challenging future work. That could, for instance, require us to implement

algorithms for article extraction [40] or improve article detection by using machine learning [53]. It would benefit The Zeeguu Reader greatly if we could generalize the detection of e.g., subscription invitations snippets, “Read more with an account” elements, or references to other articles. A future study could be to review and evaluate the relevant libraries and algorithms, and test them with The Zeeguu Reader.

Another issue regarding article detection is paywalls. These are common on news websites, but they are implemented in different ways. So, there is not one solution to detecting them. In some cases, the code of the website will contain a “paywall” class if the article is behind a paywall. In other cases the paywall is defined by a script running on the website, which can not be detected. If a user has a subscription to a website, they should be able to read the premium articles with the extension. Currently, articles behind paywalls do not work consistently. Some will work in the extension and some will not. As of now, zeeguu.org will also fetch articles in the recommendations which are behind a paywall, which could be frustrating to users, because, without a subscription, they cannot read them.

8.5. Evaluating with More Pre-Existing Users

Since we only collected four responses to the survey for the pre-existing users, we relied mostly on the data from the new users. Evaluating The Zeeguu Reader with more pre-existing users would give us more reliable data concerning whether The Zeeguu Reader consistently improves the user experience compared to using zeeguu.org.

8.6. Evaluating Learning Objectives

In this study we have focused on the personal experiences from the respondents’ point of view. We have not researched the effect on learning with the use of self-selected authentic material on language learning with a browser extension. In general, not much literature exists on this subject, where teachers are not involved in choosing the reading material. Thus, the effect of this approach on actual foreign language learning should be researched further. This could, for instance, be done by testing the language skills of users before and after using the extension consistently over a longer period, or by comparing the progress of two parallel high school classes: one using the extension and another one using a traditional textbook.

9. Conclusion

In this thesis, we have investigated how language learners experience a browser extension that allows them to translate and practice vocabulary in foreign language texts of their choice by developing a browser extension and evaluating it with users. Analyzing the data gathered in the experiment and discussing the findings, we found five key points.

(1) Firstly, a majority of the respondents experienced that the extension contributed to their language learning in a way that other learning resources they have used did not. The respondents also generally said it was useful and convenient for practicing a foreign language.

(2) Personalization in the form of self-selection of material was experienced positively by respondents. They reported that finding their own articles was an advantage because they could learn while reading texts of their interest, which some said increased motivation, engagement, and overall enjoyment. However, it can be less optimal for some people because it can be difficult or overwhelming to find material; thus, providing suggestions alongside self-selection is the most optimal approach. Reading authentic material can also motivate learners and provide them with real-life uses of the language. However, The Zeeguu Reader currently does not offer language-level detection, which made it difficult for some respondents to find texts of the correct difficulty. Thus, reading authentic texts can also cause decreased motivation for some learners.

(3) Respondents generally found that ad removal contributed to them being less distracted and more focused on the text during reading. Although there is no agreement in the literature on whether ads are distracting, we found no negative aspects of removing them in the literature or our experiment. By opening texts in The Zeeguu Reader, we ensure consistent legibility that can possibly improve reading speed and comprehension compared to the various layouts of websites.

(4) Our observations also led us to discuss the challenge of extracting readable text content from websites, which leads to formatting issues in the reader view. To combat the wrong extractions by the Readability library, we added both general and individualized website cleaning. This creates maintainability problems because hosts change their websites, and browsers are updated. These problems also lead to scalability issues because it is impossible to clean all relevant websites individually.

In conclusion, the language learning browser extension, The Zeeguu Reader, has great potential, because learners experience self-selecting authentic material as a positive feature, as it motivates and encourages them to read while having the freedom to choose articles of their own interest. However, a solution to extracting text content from websites and detecting whether texts are suitable for reading must be found as these problems challenge the maintainability and scalability of the extension.

References

- [1] ACPM. Classement des sites février 2022. <https://www.acpm.fr/Les-chiffres/Frequentation-Sites-et-Applications/Classement-des-Sites>. (Accessed August 31 2022).
- [2] Mary Ainley, Suzanne Hidi, and Dagmar Berndorff. Interest, learning, and the psychological processes that mediate their relationship. *Journal of educational psychology*, 94(3):545, 2002.
- [3] Inigo Aldalur and Oscar Diaz. Addressing web locator fragility: a case for browser extensions. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, pages 45–50, 2017.
- [4] Majid Asgari, Saeed Ketabi, and Zahra Amirian. Interest-based language teaching: Enhancing students' interest and achievement in 12 reading. *Iranian Journal of Language Teaching Research*, 7(1):61–75, 2019.
- [5] Babbel. Language learning app babbel celebrates 10 million subscriptions sold. <https://www.babbel.com/press/en-us/releases/2020-09-21-Babbel-10million-US.html>. (Accessed August 30 2022).
- [6] Babel. babel-loader. <https://github.com/babel/babel-loader/>. (Accessed August 30 2022).
- [7] Matthew L Bernacki, Meghan J Greene, and Nikki G Lobczowski. A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)? *Educational Psychology Review*, 33(4):1675–1715, 2021.
- [8] David Beymer, Peter Z Orton, and Daniel M Russell. An eye tracking study of how pictures influence online reading. In *IFIP Conference on Human-Computer Interaction*, pages 456–460. Springer, 2007.
- [9] Nanavati Bias and Randolph G Bias. Optimal line length in reading - a literature review. *Visible language*, 39(2):120–145, 2005.
- [10] Alireza Bonyadi. The effect of topic selection on efl students' writing performance. *Sage Open*, 4(3):2158244014547176, 2014.
- [11] Alireza Bonyadi and Sheida Zeinalpur. Perceptions of students towards self-selected and teacher-assigned topics in efl writing. *Procedia-Social and Behavioral Sciences*, 98:385–391, 2014.
- [12] Monica Bulger. Personalized learning: The conversations we're not having. *Data and Society*, 22(1):1–29, 2016.
- [13] Pew Research Center. Most european students are learning a foreign language in school while americans lag. <https://www.pewresearch.org/fact-tank/2018/08/06/most->

european-students-are-learning-a-foreign-language-in-school-while-americans-lag/. (Accessed August 29 2022).

- [14] Tao Chen, Naijia Zheng, Yue Zhao, Muthu Kumar Chandrasekaran, and Min-Yen Kan. Interactive second language learning from news websites. In *Proceedings of the 2nd Workshop on Natural Language Processing Techniques for Educational Applications*, pages 34–42, 2015.
- [15] World Wide Web Consortium. Web content accessibility guidelines (wcag) 2.1. <https://www.w3.org/TR/WCAG21/#contrast-minimum>. (Accessed August 31 2022).
- [16] Connor Corbin, Deniz Cetinkaya, and Huseyin Dogan. Design and development of a web extension to help facilitate the learning of a foreign language. In *International Conference on Human-Computer Interaction*, pages 345–364. Springer, 2020.
- [17] Cure53. Dompurify. <https://github.com/cure53/DOMPurify>. (Accessed August 30 2022).
- [18] Maria De Marsico and Stefano Levialdi. Evaluating web sites: exploiting user’s expectations. *International journal of human-computer studies*, 60(3):381–416, 2004.
- [19] Apple Developer. Safari web extensions. https://developer.apple.com/documentation/safariservices/safari_web_extensions. (Accessed August 30 2022).
- [20] Chrome Developers. chrome.cookies. <https://developer.chrome.com/docs/extensions/reference/cookies/>. (Accessed August 30 2022).
- [21] Oscar Díaz and Cristóbal Arellano. The augmented web: rationales, opportunities, and challenges on browser-side transcoding. *ACM Transactions on the Web (TWEB)*, 9(2):1–30, 2015.
- [22] William H DuBay. The principles of readability. *Online Submission*, 2004.
- [23] Duolingo. 2021 duolingo language report. <https://blog.duolingo.com/2021-duolingo-language-report/>. (Accessed August 29 2022).
- [24] Ahmed ElBatanony, AbdelRahman Abouneqm, and Giancarlo Succi. A browser extension to facilitate language acquisition. In *2021 7th International Conference on Education and Training Technologies*, pages 1–6, 2021.
- [25] Gülcan Erçetin. Effects of topic interest and prior knowledge on text recall and annotation use in reading a hypermedia text in the l2. *ReCALL*, 22(2):228–246, 2010.
- [26] Eurostat. Foreign language learning statistics. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics. (Accessed August 29 2022).
- [27] Terri Flowerday, Gregory Schraw, and Joseph Stevens. The role of choice and interest in reader engagement. *The Journal of Experimental Education*, 72(2):93–114, 2004.

- [28] GemiusAudience. Toplisten: Media channel. <https://e-public.gemius.com/dk/>. (Accessed August 31 2022).
- [29] Mohammad Ghasemisharif, Peter Snyder, Andrius Aucinas, and Benjamin Livshits. Speedreader: Reader mode made fast and private. In *The World Wide Web Conference*, pages 526–537, 2019.
- [30] Alex Gilmore. Authentic materials and authenticity in foreign language learning. *Language teaching*, 40(2):97–118, 2007.
- [31] Google. Make webpages more useful with extensions. https://support.google.com/chrome_webstore/answer/1047776?hl=en&topic=1212379#zippy=%2Cmake-webpages-more-useful-with-extensions. (Accessed August 30 2022).
- [32] Shaoru Guo, Ru Li, Hongye Tan, Xiaoli Li, Yong Guan, Hongyan Zhao, and Yueping Zhang. A frame-based sentence representation for machine reading comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 891–896, 2020.
- [33] Suzanne Hidi. Interest, reading, and learning: Theoretical and practical considerations. *Educational psychology review*, 13(3):191–209, 2001.
- [34] Nafiseh Hojjati and Balakrishnan Muniandy. The effects of font type and spacing of text for online readability and performance. *Contemporary Educational Technology*, 5(2):161–174, 2014.
- [35] Nanna Inie and Mircea F Lungu. Aiki-turning online procrastination into microlearning. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.
- [36] Isaacs. rimraf. <https://github.com/isaacs/rimraf>. (Accessed August 30 2022).
- [37] Heba M Ismail, Saad Harous, and Boumediene Belkhouche. Review of personalized language learning systems. In *2016 12th International Conference on Innovations in Information Technology (IIT)*, pages 1–6. IEEE, 2016.
- [38] Denise Johnson and Anne Blair. The importance and use of student self-selected literature to reading engagement in an elementary reading curriculum. *Reading Horizons: A Journal of Literacy and Language Arts*, 43(3):3, 2003.
- [39] Jointoucan.com. Toucan. <https://chrome.google.com/webstore/detail/lingq-importer/eaanhanppiiifopiabnfmhjbikjmeeale>. (Accessed August 30 2022).
- [40] Geunseong Jung, Sungjae Han, Hansung Kim, and Jaehyuk Cha. Don't read, just look: Main content extraction from web pages using visually apparent features. *arXiv preprint arXiv:2110.14164*, 2021.
- [41] Ferit Kiliçkaya. Improving formative assessment in language classrooms using " gradecam go!". *Teaching English with Technology*, 17(4):78–92, 2017.

- [42] Andreas Krapp. Interest, motivation and learning: An educational-psychological perspective. *European journal of psychology of education*, 14(1):23–40, 1999.
- [43] Soren Lauesen. *User interface design: a software engineering perspective*. Pearson Education, 2005.
- [44] Sang-Ki Lee. Topic congruence and topic interest: How do they affect second language reading comprehension? 2009.
- [45] Sunjung Lee and Diana Pulido. The impact of topic interest, l2 proficiency, and gender on efl incidental vocabulary acquisition through reading. *Language Teaching Research*, 21(1):118–135, 2017.
- [46] Maurizio Leotta, Diego Clerissi, Filippo Ricca, and Paolo Tonella. Visual vs. dom-based web locators: An empirical study. In *International Conference on Web Engineering*, pages 322–340. Springer, 2014.
- [47] Maurizio Leotta, Andrea Stocco, Filippo Ricca, and Paolo Tonella. Reducing web test cases aging by means of robust xpath locators. In *2014 IEEE International Symposium on Software Reliability Engineering Workshops*, pages 449–454. IEEE, 2014.
- [48] Maurizio Leotta, Andrea Stocco, Filippo Ricca, and Paolo Tonella. Using multi-locators to increase the robustness of web test cases. In *2015 IEEE 8th International Conference on Software Testing, Verification and Validation (ICST)*, pages 1–10. IEEE, 2015.
- [49] Qisheng Li, Meredith Ringel Morris, Adam Fourney, Kevin Larson, and Katharina Reinecke. The impact of web browser reader views on reading speed and user experience. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.
- [50] Petra Lietz. Research into questionnaire design: A summary of the literature. *International journal of market research*, 52(2):249–272, 2010.
- [51] Lingq.com. Lingq importer. <https://chrome.google.com/webstore/detail/toucan-language-learning/lokjgaehpcnlmkebpjmjiofccpklbmoci>. (Accessed August 30 2022).
- [52] Mircea F Lungu, Luc van den Brand, Dan Chirtoaca, and Martin Avagyan. 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*, pages 1–12, 2018.
- [53] Shobhit Mathur, Pritam Nikam, Harshita Patidar, Rohan Bapusaheb Gaikwad, and Preeti Narayan Nayak. Machine-learning directed article detection on the web using dom and text-based features. In *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–5. IEEE, 2021.
- [54] Pauline M McGuirk and Phillip O’Neill. Using questionnaires in qualitative human geography. 2016.

- [55] Aliaksei Miniukovich, Michele Scaltritti, Simone Sulpizio, and Antonella De Angeli. Guideline-based evaluation of web readability. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.
- [56] Iain Mitchell. The potential of the internet as a language-learning tool. *Foreign language learning with digital technology*, pages 7–31, 2009.
- [57] Mozilla. Browser compatibility. https://developer.mozilla.org/en-US/docs/Mozilla/Add-ons/WebExtensions/manifest.json#browser_compatibility. (Accessed August 30 2022).
- [58] Mozilla. mozilla/readability. <https://github.com/mozilla/readability>. (Accessed August 30 2022).
- [59] Preeti Nayak, Biljith Thadichi, and Harshita Patidar. Article detection on the web using dom and text based features. In *2019 IEEE 16th India Council International Conference (INDICON)*, pages 1–4. IEEE, 2019.
- [60] Vu Nguyen, Thanh To, and Gia-Han Diep. Generating and selecting resilient and maintainable locators for web automated testing. *Software Testing, Verification and Reliability*, 31(3):e1760, 2021.
- [61] Jakob Nielsen. Why you only need to test with 5 users. <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>. (Accessed August 30 2022).
- [62] United States Department of Education. Transforming american education: Learning powered by technology. <https://www.ed.gov/sites/default/files/netp2010.pdf>, 2010.
- [63] Aytunga Oguz and H Ozge Bahar. The importance of using authentic materials in prospective foreign language teacher training. *Pakistan Journal of social sciences*, 5(4):328–336, 2008.
- [64] Luke Plonsky and Nicole Ziegler. The call-sla interface: Insights from a second-order synthesis. *Language Learning & Technology*, 20(2):17–37, 2016.
- [65] Apple Developer Program. What you need to enroll. <https://developer.apple.com/programs/enroll/>. (Accessed August 30 2022).
- [66] Luz Rello, Martin Pielot, Mari-Carmen Marcos, and Roberto Carlini. Size matters (spacing not): 18 points for a dyslexic-friendly wikipedia. In *Proceedings of the 10th International Cross-Disciplinary Conference on web accessibility, W4A '13*, pages 1–4. ACM, 2013.
- [67] Remembermy. Remembermy - translate and memorize. <https://chrome.google.com/webstore/detail/remembermy-translate-and/dipiagiiohfljicicepggffpbnjmgjcnf>. (Accessed August 30 2022).
- [68] Ulrich Schiefele. The influence of topic interest, prior knowledge, and cognitive capabilities on text comprehension. In *Learning environments*, pages 323–338. Springer, 1990.

- [69] Semrush. Die meistbesuchten und meistaufgerufenen websites in deutschland — top 100 in 2021 und 2022. <https://de.semrush.com/blog/top-der-meistbesuchten-webseiten/#das-sind-die-top-100-der-meistaufgerufenen-websites-in-deutschland-in-2022>. (Accessed August 31 2022).
- [70] Latifeh Shakourzadeh and Siros Izadpanah. Textbook-assigned and self-selected topics of iranian male efl learners: topic interest, topic familiarity, topic importance, and topic difficulty. *Asian-Pacific Journal of Second and Foreign Language Education*, 5(1):1–23, 2020.
- [71] Atikah Shemshack and Jonathan Michael Spector. A systematic literature review of personalized learning terms. *Smart Learning Environments*, 7(1):1–20, 2020.
- [72] Similarweb. Top websites ranking for all categories in in norway. <https://www.similarweb.com/top-websites/norway/>. (Accessed August 31 2022).
- [73] Similarweb. Top websites ranking for all categories in united kingdom. <https://www.similarweb.com/top-websites/united-kingdom/>. (Accessed August 31 2022).
- [74] Jaana Simola, Jarmo Kuisma, Anssi Öörni, Liisa Uusitalo, and Jukka Hyönä. The impact of salient advertisements on reading and attention on web pages. *Journal of Experimental Psychology: Applied*, 17(2):174, 2011.
- [75] Statcounter.com. Desktop browser market share worldwide, oct 2021. <https://gs.statcounter.com/browser-market-share/desktop/worldwide#monthly-202110-202110-bar>. (Accessed August 30 2022).
- [76] Sigmund Tobias. Interest, prior knowledge, and learning. *Review of educational Research*, 64(1):37–54, 1994.
- [77] Andrew Trusty and Khai N Truong. Augmenting the web for second language vocabulary learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3179–3188, 2011.
- [78] Vocabtracker.com. Vocab tracker. <https://chrome.google.com/webstore/detail/vocab-tracker/dmnbllpfefmpaflamlfgmimhkmdmhlcda>. (Accessed August 30 2022).
- [79] Candace Walkington and Matthew L Bernacki. Motivating students by “personalizing” learning around individual interests: A consideration of theory, design, and implementation issues. In *Motivational interventions*. Emerald Group Publishing Limited, 2014.
- [80] Webpack. Loaders. <https://webpack.js.org/concepts/#loaders>. (Accessed August 30 2022).
- [81] Webpack. Webpack concepts. <https://webpack.js.org/concepts/>. (Accessed August 30 2022).
- [82] James P Wolf. Exploring and contrasting efl learners’ perceptions of textbook-assigned and self-selected discussion topics. *Language Teaching Research*, 17(1):49–66, 2013.

- [83] Pan Yi-Chun and Pan Yi-Ching. The effects of pictures on the reading comprehension of low-proficiency taiwanese english foreign language college students: An action research study. *VNU Journal of Foreign Studies*, 25(3), 2009.

A. Usability Tests

A.1. Usability Test Set-Up

Usability test tasks:

1. You are on the Zeeguu Extension page in the Chrome Web Store. Read the overview.
2. You want to practice a foreign language. To do this you want to read an article that you find interesting. Find an article in the foreign language you want to practice.
3. To practice the foreign language on the article you choose use the Zeeguu extension
 - a) You have now opened the article of your choice. In the article, you see difficult words. Find out what they mean.
 - b) You find a word that you do not know how to pronounce. Figure out how to pronounce it.
4. You realize that the article you are reading doesn't look right. Text is missing and the formatting doesn't look right. You want to notify the extension that something is wrong.
5. You have now translated multiple words. To practice these words you now want to go to exercises. Practice the words.
6. You are finished. Go back to the article on the news site.

Corner cases test tasks:

1. You have opened an English article in the Zeeguu Extension. You want to translate the words: "on", "to", and "in". You now want to practice the exercises with these specific words (native language should be danish for this task).
 - a) <https://edition.cnn.com/2022/03/29/politics/trump-putin-hunter-biden/index.html>
2. You suddenly realized you don't want to do more exercises - you would rather read the article again.
3. Go to google.com. Activate the extension. What do you see?

Post test interview questions:

- How was your experience with the Zeeguu Extension.
- Did you find that the extension was missing any functionality?
- Did you find anything confusing when using the extension?
- Do you see any ways the extension can be improved?
- What did you think of the design?
- Any last comments?

A.2. Usability Test Summary

The findings below are categorized based on the six categories proposed by Søren Lauesen (Lauesen 2005, p. 413-442).

The six categories are Bug, Task failure, Annoying, Medium problem, Minor problem, and Setup error. In this sense, a Bug is a problem in the code implementation, which should not occur. A Task Failure is when the task cannot be completed. An Annoying error is when the task is completed, but something irritating or confusing is happening during the task completion. A Medium Problem is one that takes lengthy attempts and a Minor Problem is one that takes few attempts and is quickly figured out by the test user themselves.

We did two iterations of usability tests. We did so because the first usability tests revealed problems, which could be fixed before the next iteration. In this way new problems, if any, could be discovered in regard to our changes. The first iteration consisted of four usability tests. The second iteration also consisted of four usability tests.

Not all errors caused us to make changes to the interface and interaction. Two users were unsure how the “listen on click” and “translate on click” buttons worked. With these buttons, users can enable and disable translations and pronunciations at the top of the page. However, we did not change them because they quickly figured out how they worked by themselves (Usability test 1, task 3, Appendix A).

In the usability tests, we had two kinds of task failures. Two users did not find the “Report problems” form initially. It is placed at the bottom of the page, and here users can send feedback if they experience issues with an article or the extension. Users can either click the “Report problems” button on top of the page and be sent to the box at the bottom or by scrolling to the bottom of the page. One user initially scrolled to the bottom but did not see the box. Shortly after, she found the “report problems” button at the top of the page. The other user did find the “report problems” button at the top of the page but was confused because the “practice vocabulary” box was more prominent on the screen after clicking “report problems”. Still, they eventually realized that the report problems were in a box underneath the “practice vocabulary” box. We decided not to change this because 7 out of 8 users instinctively knew that they had to scroll to the bottom.

Another task failure that two users encountered was that they did not find the exercises. But one of the users (Usability test 6) did rationalize afterward that it “makes sense in the way that, when you are done reading the article, then you are down here [at the ‘practice vocabulary’] and then you can practice all the words”. The task was announced before the user had read the entire article, so in normal circumstances, the user will see the “practice vocabulary” box as soon as the article is read. Because of this, we did not change the setup. We discussed whether we should add an exercise button at the top of the page, but we also do not want to encourage going into exercises before reading the article, as there will be no exercises ready yet.

The “Task” column presents which task the issue was experienced in. The “Amount” column presents the number of users who experienced the issue. The “Issue” column presents the issue the testers encountered. The “Category” column presents how we categorize the issue based on Lauesens six categories. And the “Solution” column presents how we handled the issue.

Task	Amount	Issue	Category	Solution
1	4	Spent a lot of time trying to find an article to read.	Medium problem	We send the users to zeeguu.org recommendations initially, so they have something to start off with.
5	1	Did not find the exercises by themselves	Task failure / set-up error	The question was asked before the user read the article, which could cause this task to fail. No changes were made.
3	2	Was not initially sure how the “translate” and “listen” buttons worked.	Minor problem	No changes were made. The users figured it out themselves quickly.
3	1	Double clicked words because loading was too slow.	Annoying	After the last iteration, an animation was added to indicate a click had been registered.
4	1	Text not removed after submitting feedback	Bug	Fixed so that text is deleted after submitting.
4	1	Did not find the “Report problems” initially.	Task failure	FNo changes were made, as almost all users knew to scroll to the bottom of the page.
Corner case 1	1	Did not understand why some words were grayed-out.	Annoying	A text was added to indicate that grayed-out words are not relevant in exercises.
3	1	Encountered exercises with long sentences	Annoying	A limit should be set so long sentences do not appear in exercises, as these caused confusion.
3	2	Did not understand the red dotted line [importance indicator] attached to some words.	Annoying	The importance indicator was removed from Zeeguu altogether, as it caused confusion.
3	2*	Clicked the disabled button and did not understand why nothing happened.	Minor problem	A text on hover was added to the button to indicate what they should do.

Table 3: First iteration summary. *Two users did not get this task.

Task	Amount	Issue	Category	Solution
5	2	Clicking on a word from the texts in the exercises caused the extension to crash.	Bug	The bug was fixed.
5	1	Did not find the exercises by themselves	Task failure / set-up error	The question was asked before the user read the article, which could cause this task to fail. No changes were made.
3	2	Was not sure what “save to zeeguu.org” meant. If it was words or the article that could be saved.	Minor problem	The text was changed to “Save article to zeeguu.org”.
1	1	Spent a lot of time trying to find an article to read.	Medium problem	Users are now initially sent to zeeguu.org.
1	1	Double clicked words because loading was too slow.	Annoying	After the last iteration, an animation was added to indicate a click had been registered.
3	1	Did not find the “Report problems” initially.	Task failure	FNo changes were made, as almost all users knew to scroll to the bottom of the page.
Corner case 1	1	Did not realize you could star words to go to exercises	Task failure	A hover text was added on the disabled button to explain that words should be starred.
4	1	Did not find the “Report problems” initially.	Task failure	No changes were made.
3	4	Clicked the disabled button and did not understand why nothing happened.	Minor problem	A text on hover was added to the button to indicate what they should do.
Corner case 1	2	Did not know that the “Should this be readable” button was used to send feedback	Minor problem	Renamed to “Should this be readable? Send feedback”.

Table 4: Second iteration summary.

A.3. Usability Test Transcripts

Usability test 1: The tasks and corner cases

Task 1: Reads the document – no comments.

Task 2: Looks at google.com for a few seconds and says, “I don’t even know where to find a German article”. Is there a paper called “Die Deutsche Welle”? She looks at the first result and clicks on it. She looks around the website confused. It’s an English site about Germany. Realizes she can’t use the website. She goes back to google. Try searching for “German newspaper” – all the first results are in English. Scrolls around the site. She goes to Wikipedia and finds a list of German newspapers. Go to one of the websites and click on an article.

Task 3: She clicks on the extension and quickly clicks on “Read article”. She hovers over the

“translate on click” button as the first thing. Then she tries to do the same over “listen on click”. “Ahh, that makes sense. I assume the first button gives a written translation. And the other one gives one with sound”. Then reads some text and finds a word she is not sure she understands. She hesitates. Go back to the “translate on click” button. Click on it again. Realizes that it was already clicked – “I maybe would have thought that the bright yellow meant it was turned on. But maybe that’s just me”. Goes back to the text and clicks on the word, which is then translated: “That makes sense”. Then goes on to activate “listen on click” and deactivate “translate on click”. Clicks on a word – and the word is said out loud, “Oh, it’s not the translation, it just says the word in German. Oh yeah, I read that in the description of the extension. So, I know how it is pronounced”. “What if I have both buttons turned on?” She proceeds to do this and clicks on a word. Communicates that this is a nice feature.

Task 4: Is in the middle of the article when the task is explained (so she can’t see the “report problems” button at the top of the page and has not registered it beforehand). Take the mouse up to the Zeeguu logo – but do not click on it. Take the mouse up to the translate and pronounce icons – but do not click on them. “I’m not sure.” Is about to explain that she is not so good with technology – but at the same time she scrolls to the bottom and goes “ahh, right here”. Proceeds to explain that it made sense that it was at the bottom, but that she had not yet realized that the extension had opened the article in a new “window” because she did not know how the news site’s design was. So, she at first was not sure where to look.

Task 5: Find the “Practice Vocabulary” right away. Saw it when she reported a problem. Clicks on the button. “Ah these are the words I did not know. That’s just like Duolingo”. Proceeds to the exercises. Read the first exercise “Ah, so I have to write the word in German?”. Click on the hint – and proceed to use “Show solution”. Realizes she can get the correct answer read out loud. “Makes a lot of sense. I did not take the time to look at which words I clicked before I did the exercises.” She is slightly confused. Clicks the correct answer and gets the word pronounced. Click “Show solution” on the next exercise.

Corner case 1: Don’t understand the question – slightly confused. Translates the words mentioned by the facilitator. She goes to “Practice Vocabulary”. She pronounces the words. Do not realize that the “To exercises” button is disabled. Clicks on it multiple times. “Oh no – nothing is happening. Is that supposed to happen?”. Is that because they are grayed out? Oh, so I must star them. It makes sense if I read the text. “Maybe ‘star’ should be bold?”.

Corner case 2: She clicks on the arrow in the top right corner of the modal window. The extension shuts down and we are back at the news site. “That’s probably how I would do. And then open it again. Maybe I could also finish the exercises and then go back from there?”

Corner case 3: Activates the extension. “The article is not readable. That makes sense. We are not on an article”.

Usability test 1: The questions

How was your experience with the Zeeguu Extension?:

“It was a nice experience. Very pretty – nice layout. Visually appealing and it is clear what to do. It made sense that you found help at the bottom – but maybe I also expected something at the top – a cogwheel or something else. But maybe that would just be for dummies – it also makes sense that it is at the bottom”.

Did you find that the extension was missing any functionality?:

“No, I don’t think so. Can I choose the translation language?”. We explain that this can be done inside of the website. She opens the popup and clicks on settings. And she sees that you can change the language here. “How about a link to settings from inside of the article? Maybe I would think that would be nice”.

Did you find anything confusing when using the extension?:

“No, I don’t think so“.

Do you see any ways the extension can be improved?:

“No, just a link to settings from inside of the article”.

What did you think of the design?:

“Very pretty. Calming and nice that there are no commercials and ads. But I don’t know if that is something you did? That’s really nice, so you are not distracted by anything. I like that it is so light - sometimes PDFs are opened and then it is dark on the sides. I like that this is so light. Cute elephant also. And I study art history”.

Any last comments?:

"Thank you for a nice experience. It was pleasant".

Usability test 2: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: “Do I just Google?” – looks a bit confused. “So, do I have to find something in German?”. The facilitator mentioned that we also have a list of news sites if she doesn’t know how to find a news site. The tester wants us to help her. She goes to faz.de. She looks around the website. Goes to “Sport” to find something interesting for her. Scrolls a lot to find something she is interested in. She then clicks on an article.

Task 3: Clicks on “Read article” straight away. Starts reading the article and quickly says “Ah, and then I can click on the word”. She proceeds to click on a word and see the translation. Click on multiple words. “Ah, that’s nice. So, you have all the translations. It’s quite nice that you can click on every word. So, every word can be translated”. The facilitator asks “What if you want to hear the words pronounced?”. The tester quickly moves the mouse up to the icons in the top right corner “Then I saw these”. She proceeds to click on “Listen on click”. “And then I guess I click on the words again.” And sees that this works. Clicks on multiple words in a row and see that it then translated the words together.

Task 4: She is almost at the top when the question is asked. She scrolls up the last bit, so the “report problems” button is visible. She says, “My thought would be to scroll to the bottom to see if there is any contact information” While she proceeds to do this – and finds the “Report problems” form. “But I also saw this” and she then scrolls to the top and shows us the “report problems” button at the top of the page. “But my first thought would be to go to the bottom”.

Task 5: Find the “Practice Vocabulary” straight away. Reads the text in the yellow box. She looks at the list of words “So, all the chosen words from the article are in the exercises”. Says “What do I do?” but then immediately sees “To exercises” and clicks on this button. Facilitator: “You can try to do some exercises if you want to, otherwise you can click ‘show solution’”. Tester: “So, here I have to translate this word into German?”. She proceeds to click “show solution” and clicks on the pronounce button to hear it pronounced. In the next exercise, she clicked on the correct word: “Ah then I can choose”. Click “show solution” for the next exercises.

Task 6: Click on “back to reading” to go back to the article. She misunderstood that she had to go back to the news site. The facilitator clarifies “What if you wanted to go all the way back to the news site?”. Then click on the cross in the top right corner – and it is back at the news site.

Corner case 1: Translates the words mentioned by the facilitator. Go to “Practice Vocabulary”. Clicks on the disabled button multiple times. “I can’t click on it” does not seem to understand why she can’t click on the button. Reads the text in the yellow box and now understands that she needs to star the words. Proceeds to exercises.

Corner case 2: She clicks the “Go back to reading” link and goes back to the article.

Corner case 3: Go to google.com and search for German. Activates the extension. “Okay, so you have to be on an article site. But would it work on Facebook for example?”. The facilitator answers “no, that would not work either”. “So it is limited texts? I think that makes sense”.

Usability test 2: The questions

How was your experience with the Zeeguu Extension?:

“I think it is easy to activate. When you activate the extension it is like in a popup window – the translations – so it cuts off other articles within the article.” Facilitator: “So, if the article links to other articles?” Tester: “Yeah” – seems slightly confused about this. “But yeah. I think it’s easy.”

Did you find that the extension was missing any functionality?:

“I don’t think so. Maybe it depends on how good you are at the language. If it is too difficult you could end up clicking every word and then get the translations. But that is the only thing.”

Did you find anything confusing when using the extension?:

“Maybe, the star, when you showed the Trump article. That you have to star the words to use them in exercises. I did not understand why I had to do that”.

Do you see any ways the extension can be improved?:

“Maybe some sort of alert or popup explaining what the program is? What does it mean when

you star words. I was a little confused about what to do”

What did you think of the design?:

“Very nice - very simple. Not too many colors or distracting movements or elements.

Any last comments?:

“No. Very simple to use. Not that complicated”.

Usability test 3: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: “So, I have to google to find a website by myself? Okay”. Googles for “News German” and she sees that all results are English news sites about Germany. Scroll through the results. Go back to the search bar. Adds “.de” to the search. Still only see English sites. Finally she finds a news site in german. Clicks on it. The website she is on has navigation in both English and German “it is slightly confusing that they have the English and German titles”. She goes into politics and looks for an article – find one and click on it.

Task 3: She open the popup and quickly clicks on “read article”. Says “it was easy to know what to do because there was a big button that said ‘read article’ so I instantly clicked on it”. “First of all, I’m just looking at what I can do. I think I am a bit confused about these ones” while hovering over the “report problems” and “Make a personal copy” buttons. “But I won’t take any further decision about them, and I’ll just leave them for now and if I have any problems then I will scroll back and try to do what you can do. But I’m a little unsure of what ‘make personal copy’ means. Can I download a PDF of the article? Or can I download the reading help I got from Zeeguu? Or what can I do with a personal copy? It’s not very concrete – I don’t know what action there is behind it.”

She proceeds to want to translate a word. She hovers over a word but then goes to the “translate on click” button. “First I was a little confused if I have to click on a word or on the “translate on click” button.. But actually, I can now see that the button is already activated. So, maybe I just have to click on a word. And if I want to listen to the word, then I can activate this on” as she hover over “listen on click”. She then proceeds to click the button. “Oh, both can be activated. Nice”. She starts by only having the translation activated. And proceeds to click on a word. “Ah, then it tells me the translation.” She proceeds to click on multiple words in a row and get the entire headline translated. “And if I’m not sure how to pronounce it I active ‘listen on click’” and she proceeds to do just that. It then pronounces the entire sentence “ah, and it takes all of it. She now has four words translated in a row, and she clicks on one word to get it pronounced. Because the word is already translated it removes the translation for all the translated words. She is surprised by this. She wanted to hear the translation (but she also had ‘translate on click’ activated). “Oh, I undid all of them. Okay. I only want to have ‘anbieter’ pronounced, but if I had chosen the other ones as well, It also pronounced them”.

Task 4: “So, the article looks wrong – or the extension makes the article look wrong?” the facilitator answers that it is if the extension makes it look wrong. She scrolls to the bottom, but

only reaches “practice vocabulary” before she starts going back up. She scrolls up and points out a word that she thinks might look wrong. Then she clicks on the word. Scrolls up and down before she goes all the way to the bottom “I will scroll down to see what else I can do.” And finds the form to report problems. She writes a text and submits the feedback. “and then I think I have to submit the feedback and I get this nice message that my feedback is sent. It says ‘thank you for the feedback so I think it is sent now. But it is confusing that it is not deleting the text in report problems, and that might get me confused over whether my feedback was sent. Or do I have to do it again? Even though I got this popup dialog. So that might be a little confusing.’”

Task 5: She is asked to do exercises, but she scrolls past the “practice your vocabulary” button and proceeds to click on more words in the article. She understands that you can do exercises, she just wants to add more words. Note that she clicks on multiple words in a row, so she translates a lot of sentences instead of individual words. She scrolls back down to “practice vocabulary” and clicks on the button. She reads the text in the yellow box “okay, I might be missing a bit more explanation as to why some words are greyed out. Why are some of the words more important than others? It’s confusing that it’s only some of them. But it makes sense that I can delete it and that I can star it. I know from the web, that if I want to favorite something you always star it. So, that is easy for me to recognize, that these are the ones that are most important for me, or if I want to save them for later. But it might be confusing that it has priority in the exercises – and that they are not just saved. Maybe priority could be visualized in another way – with something else than a star. A number or a color or something else. Because a star normally means ‘favorite’ and I don’t know which order I favorite things in”. She proceeds to star some words. But because she has clicked on multiple sentences, she stars sentences with multiple words in them. And proceeds to go to exercises. Because she has chosen an entire sentence, the exercise she sees is confusing. She is asked to find the expression in the context – but she sees the entire sentence as the expression and the context is also the entire sentence.

Before looking at the exercise she asks: “Another thing I think was a bit confusing was that the exercises divided most of the words I clicked into one ‘word’. Can I maybe change the exercises? Can I go back? There is no back button? Oh, well.” She proceeds to ask whether she can choose what words go in the exercises. If she can change them – and only learn some words from the sentence. In general, this exercise leads to confusion. She reads the exercise “Find the expression in the context below”. “And then I should do what? Write it – or pronounce it? Maybe. I don’t know actually. I will take a hint. Oh, so I have to write it. Do I also have to pronounce it? Or is it only about writing? It might be a little bit confusing that the sentence is already there. I think I cheated. I don’t know how else to do it”. Clicks on “show solution” and realizes what she did wrong. She wrote six words, but only needed to write the first four “ah, that makes sense. Because that was what I clicked on”. When reaching the last page, she says “I like that I can see what I did good and what I might do better for next time. How this is divided. What do the red dots mean? [the importance indicator] was that because it was a bad word to choose?”

Corner case 1: This task did not work intentionally because the words were not grayed out.

Corner case 2: Did understand that she had to click the “go back to article” link to go back –

although she thought she would go all the way back to the article on the news site.

Corner case 3: She understands that she didn't choose an article "I did not choose an article so I have to choose one? It makes sense. And maybe it has to be more clear that I have to find an article. 'Article not readable' - maybe it is misleading. For example "find an article by..." or something.

Usability test 3: The questions

How was your experience with the Zeeguu Extension?:

"Good, as you saw some things were quite confusing for me. But after having done it a couple of times I understand essentially how to use it. Especially the exercises were a bit confusing for me.

Did you find that the extension was missing any functionality?:

"Yeah, a back button in the exercises would be nice. Maybe I did one exercise and wanted to go back to see the solution for the first one. So back and forth between exercises.

Did you find anything confusing when using the extension?:

"A text about how the exercises work. Make it more clear what part of the sentence I had to translate. Maybe an explanation of what I have to translate and exercise."

What did you think of the design?:

"I think it is really cool – I like the colors. It's very clean and neat. I think it can do what it should do. It was really easy to navigate - besides missing the back button. I think the buttons were nice and it was easy to find out what to do next. For example, scroll down the article – I can practice and report problems. Nothing was hidden.

Any last comments?:

"No, I think it was really nice and cool – I think it could help a lot of people to learn a new language. And a nice and easy way to do it – it's not just Duolingo. It's something that you want to learn about. And you can decide on the article by yourself – so it will not be an article about German politics, but something else that you can understand better than other topics. I think it was nice that I could choose the article by myself – and I can always decide the difficulty of the words I want to translate. I could also find an article for children where they use other not so difficult words than they might use in politics.

Usability test 4: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: She quickly searches for dr.dk and finds an article, but because it is a gallery article it is not readable. She reads the text from the popup "Article is not readable". She is not sure what to do. Clearly confused. Moves the mouse back and forth from "should this be readable?". She concludes "I need to find another one" and goes back to Google. She now searches for "news Danish" instead – but only sees English articles about Denmark. She thinks about what to search for and writes "Dansk Gyldendal" because Google suggests it. She clicks on multiple links but

no articles. She ends up on Gyldendal.dk but cannot find anything. Then she searches for dr.dk again and ends up on journalisten.dk. This article is readable. It took 3 minutes to find an article.

Task 3: Clicks on words straight away – and clicks on multiple words in a line. When she clicks on a word twice the translation for the entire sentence is removed. Some lagging issues when she tries to click on all words in the headline. She tries to double-click on all the words. So, the translation is removed before she sees it. Sometimes clicks on a word 3-4 times to get the translation – because she clicks too many times in a row. The facilitator asked how she would get the words pronounced. She sees the little black triangle if she hovers over the translations. She clicks and sees the different options for the translation “Oh, you can change the translation”. She then proceeds to go up to the icons in the top right corner – first hovering over “translate on click” and then moving to “listen on click”. She clicks on this and then clicks on a word.

Task 4: Immediately starts scrolling to the bottom – and finds the “Report problems” form.

Task 5: Quickly finds the “Practice your vocabulary” and sees the list of words. Also seems to read the text in the yellow box. Get a couple of words pronounced. She has a long list of words – she scrolls up and down. “This is the list of words I translated”. Then scrolls to the bottom and clicks “To exercises”. Do a couple of exercises. The facilitator asks “If you don’t want to do more exercises – and instead go back to the article?” and she clicks on the link next to “Go back to reading” straight away.

Task 6: Clicks on the cross in the top right corner. And is back on the news site.

Corner case 1 and 2: These tasks could not be performed because it does not work from English to English.

Corner case 3: “That it is not readable I think makes sense. It’s not an article. If I for example go to Wikipedia, it should be readable” proceed to do that. And sees that it is readable. But Google is not “because it is not one article”.

Usability test 4: The questions

How was your experience with the Zeeguu Extension?:

“It was something that really made sense to me. It’s useful if you don’t understand something. Instead of copy-pasting and opening the browser looking for the translation. Having to do that for everything. So, it’s useful.”

Did you find that the extension was missing any functionality?:

“No, I don’t think so”.

Did you find anything confusing when using the extension?:

“No, I just did not know where some things were at first, but I found it”.

Do you see any ways the extension can be improved?:

She tries to highlight a long sentence to see if this works. “To translate you have to click the different words – that is a little bit annoying. So maybe mark all that you want to translate. And translate the phrase and not only one word at a time. “

What did you think of the design?:

“It’s simple and a minimal design. It’s cool. Looks good.”

Any last comments?:

“No, it’s really useful and something I would like to have in the future to learn a language”.

Usability test 5: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: Search for “French election 2022” and see a lot of English articles. Search for French news. “I heard that the election is in France soon, so I just searched for that, to see if there was any French news. I don’t think there was. I’m not quite sure what to search for in order to get some French articles. Search for “français” and take the first result Google suggests. Goes into a news site.

Task 3: She activates the extension and clicks on “Read article”. She looks through the article. Clicks on a word and sees a translation. “Yeah, so I’m just trying to click on words that I am not quite sure what means. So, my first thought was to click on a specific word. And because I see this arrow [on the translated word] I’m thinking that I can press on this. That is the reason the arrow is there. Ah okay. So, then there are different translations from different sites, and I can create my own – I guess.” Continues to click on words. Facilitator: “And what if you wanted to hear how a word was pronounced?”. She goes straight to the icons at the top. First hovering over “translate on click” and then “listen on click”. She clicks on it and afterward clicks on a word. “That’s amazing”. “So, it makes sense, that if you want to get just the translation you click this. Oh, maybe you can choose both?” she tries and clicks on a word. “Okay, you can choose both”. “So, yeah. I think I understood that. You can choose if you want the translation or the listen button – or both- and then you can click on the words you do not understand. And I actually really like that they [the translations] stay there. Because when you are learning a new language you forget, so it’s nice that it stays there. And then you can remove them again by clicking on them”.

Task 4: Quickly scrolls to the bottom. “So, my first thought was just to scroll all the way down. And if I were to report a problem, I would write it here. Maybe a problem could be that the translation is not correct. I guess. But I don’t know the article, but I think it looks fine.” She scrolls to the top and notices the “Report problems” button at the top. “Oh you can also [she clicks on the button] – ah, makes sense. The top button sends you all the way down”. “Save to zeeguu.org?”. Facilitator: “What do you think that means?”. “Hm, maybe that it saves the article? I don’t know if you can log into Zeeguu? Then I guess it was to save the article. Or maybe the words?”.

Task 5: She scrolls down to the bottom again and finds “Practice vocabulary”. “Oh, so it keeps the ones you pressed on – the words.” Proceeds to read some of the text in the yellow box. Pronounced a couple of words. Clicks on “edit this word”. “I pressed this one to see what it means. I see the word and the translation and the context it was in. I don’t know if I would think it was nice if the context also was translated. But again it is not about the specific words, so maybe that is not what the page should do. I like the idea that there is a context because sometimes the words are in different forms. And then it is nice that it is in a context because sometimes they are spelled differently. I would probably delete this word [France] because this is one of the easy ones. The star I think indicates that you can save the word somehow. Then I go to the exercises. And then you can exercise by writing them, to learn how to spell them.” She proceeds to do the exercises. “oh, so there are different kinds of exercises, that also makes it a bit more fun”.

Task 6: “I just noticed this one [go back to reading]” and she clicks that. The facilitator does not ask to close the extension altogether.

Corner case 1: Tries to click on “To exercises”. “I’m not sure how. It looks like this red thing [the importance indicator] which I think indicates something. I am not sure.”. She tries to go to “edit word” but closes it again. “I’m not sure if you can do the exercises. Maybe it is too simple of a word? Or because it does not have a context? I don’t know. So maybe it’s impossible, and that is why it is red? It is not possible”

Corner case 2: Was not asked, because she showed us before that she could go back using this link.

Corner case 3: Activates the extension. Read the text. “okay, so ... [clicks on “should this be readable] uh.. I would not have pressed that if I knew that it was sending something. It says ‘should this be readable’ and I thought that if I pressed this it would have an error message, which told me ‘this is not an article..etc’. So, I didn’t think it would be sent to you. If that makes sense.”

Usability test 5: The questions

How was your experience with the Zeeguu Extension?:

“It was really nice. And it’s really intuitive to click the different buttons and the different icons. It makes sense. It is a very nice layout. And yeah, it is built in a way, where I know that I report a problem at the bottom because that is normally how websites are built. So, it was really easy to use. Really nice.”

Did you find that the extension was missing any functionality?:

“Only maybe that the context in the exercises – in the “edit word” I don’t know if it would be nice to make the context translated. But then again, I feel like this extension is very word-based, so I don’t know if it is annoying if it ruins the whole concept if you translate the whole context. So maybe that’s just me – that’s interesting in knowing what the whole sentence means”. But I guess I would also use this for French if I had a bit of an idea of what it actually said. So maybe it does not make sense in this case.”

Did you find anything confusing when using the extension?:

“It confused me a bit why I could not practice the word ‘to’. It makes sense that it is too simple, and it had a red line. But maybe it would be nice with a question mark you could press, just to get to know why I can’t exercise this. I don’t know who would be mad that they can’t exercise ‘to’, but it is nice with these error messages. And the same with the “should this be readable”. I was surprised when I pushed that, that it was sent directly to you. I thought, since there was a question mark, that it was a ‘this is why I can’t read this’ message.”

Do you see any ways the extension can be improved?:

“Besides the things she said. “No, actually not. It’s nice. Nice layout. But maybe – what does the red and orange means [the lines – importance of words]. That could be a bit more clear maybe. Maybe a little box that explains it. But I like that the exercises are different kinds of exercises. That makes good sense.”

What did you think of the design?:

She already explained this – so this question was not asked.

Any last comments?:

“I think I would use it”.

Usability test 6: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: Googles for a news site she knows – süddeutsche Zeitung. “I’m going to find something interesting. Maybe click on ‘Ukraine’.” She looks at the top story – but says “oh, this is a live article. Does it have to be an article – or is this fine?”. The facilitator says that she can try it. It works and the live article is opened.

Task 3: She starts reading. “So, I don’t know what this word is, so I’m clicking on it. Ah, that is very smart. And I can also try this. Ah. So, when you click on all of them it goes above the word [the entire translation]. That’s actually really smart. The facilitator asked what she would do if she wanted a word to be pronounced. “Then I think I could..” and then she clicks on a word. Afterward, she goes up to the icons at the top and hovers over “listen on click”. She clicks on it. She seems to think that the word she has just clicked would be translated if she clicked on the button “listen on click”. She clicks on it multiple times. Then she unclicks “translate on click” and tries to click on a word – and sees that now it is pronounced. “Ah okay. So you have to activate it first and then click”.

Task 4: She thinks about it. Does not understand the question, “I’m having a little trouble understanding”. She scrolls to the bottom initially – and “report problems” becomes visible, but she does not see it. She scrolls back up. “I have read that I can write an email [in the extension description], but I don’t see that here. But that is written when you download the extension. So, I could write an email.”. being translating more words. Realizes you can have both translation and pronunciation active “that is also really nice”. She then finds the “report problems” button at

the top of the page, “oh, so there is a ‘report problems’ here”. She also sees ‘Save to zeeguu.org’ – “and when I want to save some words you can click on ‘save to zeeguu.org’? Opens the popup in the extension again. Clicks on ‘settings’ and ends up on zeeguu.org. Goes to “words”. But cannot find the words she just saved by clicking ‘save to zeeguu.org.’ “I guess it would be here.”. She starts going through zeeguu.org – we stop her.

Task 5: She can’t find the exercises. Scrolls to the top. Does not realize what exercises mean. “if there were words I wanted to exercise I would use these two functions [translate and pronounce] and then I would first hear how it was pronounced and maybe click on it again so I can’t see the word in Danish. And then maybe read it a few times and try to remember. But I don’t know if it makes sense to have a longer article because this is very short.” She scrolls up and down – and we see Practice vocabulary multiple times, but she does not see it. She tries to find another article but does not find the exercises. But she closes the extension by clicking the cross in the top right corner.

Task 6: This task was not asked. Beforehand she used the cross in the top right corner to close the extension window.

Corner case 1: Before this task, the facilitator reveals that there are exercises if she goes to “practice vocabulary”. She clicks on the button and sees the one word “to” that was clicked. Pronounce the words and she tries to click on “to exercises”. But as it is disabled nothing happens. She looks at the text in the yellow box, but does not seem to read it. Tries clicking on the word itself. Tries clicking on the “edit this word” icon. Then she star it and “to exercises” becomes clickable. She clicks it. Reads the yellow box and tries to click the disabled button. does not understand it. Then stars it - and she can now go to exercises.

Corner case 2: She clicks “back to reading” without issues.

Corner case 3: Reads “Should this be readable” and clicks it immediately. “okay, so if you think the text should be readable, you can send the feedback.” But did not seem to understand that before she clicked. “It’s funny because when I was in here [in the extension article] I don’t know why, but I actually did not read this at all [the gray boxes at the bottom] because I thought it was a part of the article. Not a part of zeeguu. That was the reason I did not go down there to search for the exercises. I didn’t even read it.

Usability test 6: The questions

How was your experience with the Zeeguu Extension?:

“I think it was nice. It’s a simple idea and it’s good. And I like that it is an extension, so you can read the article with the translation being there, so present. Instead of looking up all the words as you go along the way. And it’s nice that you can practice the words within that article, so you don’t have to go into the website and practice it there. So, it is all very present – you don’t have to go everywhere. And I like the design and it’s just really nice and simple. You don’t get confused about it because there are only two functions [translate and pronounce] when you read it. And it’s very intuitive.

Did you find that the extension was missing any functionality?:

“I feel kind of dumb that I did not care to look at this [the grey boxes at the bottom] and I thought it was part of the article. I think I thought that because on a lot of articles there are these kinds of ‘subscribe to our newsletter’ or something like that. So that was what I perceived when I saw it without reading it. Hard for me to say if something is missing because I obviously missed the exercises. Facilitator: “should it be maybe somewhere else?”. Then she said “I think it should be up here also [points to translate and pronounce icons] so not only at the bottom. Then you know it is also a function that is as present as these functions [translate and pronounce]. But I guess it also makes sense in the way that, when you are done reading the article, then you are down here and then you can practice all the words. So, in that sense, it makes sense.

Did you find anything confusing when using the extension?:

Maybe you have to star the words when you want to exercise them. But given that she did not find the exercises initially, it makes sense that she is confused. She thinks you have to star every word you want to exercise.

Do you see any ways the extension can be improved?:

“Not anything else than to have the exercise button more visible – up here [in the top header]. Maybe when you download the extension there could be a walkthrough. I think that is all.”

What did you think of the design?:

“It cancels out all the commercials and all the extra things that are on articles. So much clutter. I don’t need to see all this; I just want to listen to or read the text. So I think the design is very nice and simple – and not too disturbing.

Any last comments?:

Nice tool. I think I would use it – it’s very simple and I think it works nice.

Usability test 7: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: Immediately googles “Der Spiegel”, a German news site he knows. Scrolls through the news site and looks for something interesting.

Task 3, 4 and 5 combined: He activates the extension and clicks “Read article” quickly. The first thing he does is hover over “translate on click” and “listen on click”. “So, now it will translate” and he clicks a word to see that it works. “and if I use this one it will pronounce” and he clicks a word to see that it works “yeah, that makes sense”. He clicks on multiple words. Occasionally double-clicking because the translation does not appear straight away. And he tries activating both – and neither. Clicks on more words and tries to translate multiple words by highlighting them all. But it does not work. He sees that an arrow appears if he hovers over a translation – he clicks on the arrow and sees the translation options. He closes it again. He clicks on words that have been translated and sees that they are removed. “So, if I want to translate a full sentence”

and then he proceeds to click on multiple words in a row. When he has three words in a row translated he tries to click on one of the words and sees that the entire translation is removed.

“Can I also translate.. [proceeds to click on the word ‘lehnt’] because I know from German that these two words [lehnt and ab] are connected [even though there are multiple words between them]. Oh, yeah. Cool” It seems to work the way he thought. That Zeeguu could detect that the two words are connected. He proceeds to scroll a little and we see something that looks weird. Bullet points that are empty. “This looks weird”. The facilitator asks “what if you want to notify us about this?”. “I would report problems.” And he then scrolls to the top and clicks “report problems”; he seemed to have noticed this button beforehand. The extension jumps down to “Report problems” but he does not see it. He only sees “Practice Vocabulary” and he clicks on this. “I’m not quite sure I understand what this was about. So I just clicked it”. Then he went back to the article. He again clicks on “Report problems” and is sent down again. But he only sees “Practice vocabulary”.

“So, this is if there is a problem with the translation? Oh, this is the thing I read about – the more personalized” he again clicks on “practice vocabulary”. “Ah, then I can see which words I clicked on. He reads the text in the yellow box. “ah, and then I can go to the exercises”. He goes to the exercises. He clicks on the correct word, but the extension crashes. We discovered a bug regarding clicking words in exercises. He tries it multiple times. He goes through the exercises without clicking on words. He goes back to the article and scrolls to the bottom. He then finds the “report problems” form. “that confused me. If you go to “report problems” then “practice vocabulary” is what I see. So, I don’t know if it is possible to have a tab that is called “practice vocabulary” and a tab that is called “report problems”. So one leads you to one and the other to the other one.

Corner case 1: Tries to click on the disabled button multiple times. “Do I have to press the star? Oh yeah. I should have just read the last part [of the text in the yellow box]” proceeds to exercises.

Corner case 2: This task was not presented.

Corner case 3: Opens the extension popup and clicks on “This should be readable”, but he is surprised when he sees that he has given feedback. “Oh okay, I didn’t know I was giving feedback there”. He proceeds to go to Wikipedia and open the extension. “and then I can read the article. Looks pretty good”. google something - activate extension: “this should be readable” didn’t know he was giving feedback.

Usability test 7: The questions

How was your experience with the Zeeguu Extension?:

“Good. The tool itself makes it a lot easier to read this page. Especially “Der Spiegel” [The news site he visited] was full of weird subparts of the page, pictures, and different fonts. So, it makes it easier. There were a couple of times, as you probably noticed, where I was not quite sure what I was actually doing. But it was not connected to the reading part, but more connected to the exercise and feedback part. Overall, very good. And I could definitely see myself using that when reading a page and learning a language.

Did you find that the extension was missing any functionality?:

“Intuitively it would be nice to be able to... Like when you are translating a sentence, it seems like you need to press on every single word. It would be very intuitive to be able to mark the entire sentence so that it translates the entire sentence.”

Did you find anything confusing when using the extension?:

“The part about pressing ‘report problems’ because the first thing I noticed was ‘practice vocabulary’ but other than that it was just that. And then the part where you are on the google search page and it says “this should be readable” I was thinking that this was actually your communication to me – so like a ‘we should be able to read this’ to activate it or something. So, I did not know I was giving feedback.

Do you see any ways the extension can be improved?:

He mentions specifically the bullet points part of the first article, which did not look right “that looked a bit strange.”

Usability test 8: The tasks and corner cases

Task 1: Read the description without saying anything.

Task 2: Googles for the news site “Die Zeitung” and goes to the news site. “This is a German newspaper. And then there should be an interesting article to read” he scrolls to find something interesting.

Task 3: He activates the extension and asks, “do I just press this button?” referring to “read article” and he proceeds to do so. Starts to read the article. Scrolls a bit. Reads without doing anything else. Facilitator: “and if you encounter a word you don’t understand, and you want to know the meaning, how would you go about that?” He goes up to the icons at the top and hovers over “translate on click”. “So, I will click” and he sees the translation for the word. He clicks on multiple words and sees the translations. He reads more of the article. The facilitator said, “if you find a word you don’t know how to pronounce. How would you do it?” He then goes up to the sound icon and sees “listen on click” when he hovers. He clicks it and then clicks on a word. Hears the word pronounced and sees the translation.

Task 4: The first thing he does is click on the Zeeguu logo. Does not work “is there a little help icon somewhere?” he asks while he scrolls to the bottom “ah, here. So, then I would type in here what the problem would be and then submit the feedback”.

Task 5: He does not realize that there are exercises to be done. When asked to do exercises he says that “I would try to pronounce and translate again. And click the words. He clicks on multiple words. He realizes that you can have both pronunciations and translation activated. He does not see the exercises.

Corner case 1: The facilitator tells him to go to the bottom of the article and practice the words.

He tries to pronounce the words. And he then goes down to the button “to exercises” and clicks multiple times - even though it is disabled. Reads the text in the yellow box and starts the translations. “if I actually read what it says” and then he goes to the exercises. He also clicks on a word and the site crashes. Same bug as the previous usability tester.

Corner case 2: This task was not presented.

Corner case 2: “So, they can’t read it. Should I try to google something else?” The facilitator asks “does it make sense that you cannot read the text?” he replies “no. Why should it not be?” he then goes to ‘settings’ on zeeguu.org. and looks through it. Maybe it is because it is not supporting Danish. Or is it because I typed that I wanted to translate from Danish?”

Usability test 8: The questions

How was your experience with the Zeeguu Extension?:

“I think that the feature itself is pretty cool. That I can go onto a site and when I try to read it, I can translate and get it pronounced straight away. That is a very cool feature - I would like to have such an extension tool in my own toolbar. So pretty cool. Perhaps if I had chosen more words it would be a bit more clear how it would work with the practice” because he did not find the exercises in the article where he translated a lot of words.

Did you find that the extension was missing any functionality?:

“I think you pretty much guided me to what I should be looking for, perhaps that was a bit leaning. For example, the exercises – if you had not told me it was part of the functions I would not have noticed it. Report the problems – maybe it would be more intuitive if there was a question mark icon up here [top header] or anything like that. Otherwise pretty straightforward.

Did you find anything confusing when using the extension?:

“No, not really”.

Do you see any ways the extension can be improved?:

“I think a cool function could be if you had the article itself. Instead of having it be activated and isolate the article, then if I could just browse directly in it. So, if it was already activated I could just press a word directly in the article. I don’t know if that is conflicting in terms of actually being able to navigate the site itself. For example, if they have a link to somewhere else I could imagine that this could conflict with that. But that would be a nice feature. Or that it did not turn away from the actual article. But that is a minor thing.

Any last comments?:

Working otherwise pretty good - pretty simple.

B. Surveys

B.1. Initial Survey Questions

- Email
- Name
- How old are you? (18-25, 26-35, 36-45, 46-55, 56+)
- What is your gender? (Female, Male, Other, Prefer not to say)
- What is your occupation?
- What is your native language?
- What language would you want to use in our study?
- How would you characterize your level in the language you are learning? (Beginner A1, Elementary A2, Intermediate B1, Upper intermediate B2, Advanced C1, Proficient C2)
- How do you usually learn that language? (By taking language courses, Speaking and/or writing with people, Online platforms, Other)
- How much time do you spend actively studying the language every week? (on average?) (0-30 minutes, 30-60 minutes, 1-5 hours, 5+ hours)
- Have you used Zeeguu.org before? (Yes, No)

B.2. Issue Detection Survey Questions

- How did you find the installation process? If you found it difficult, please elaborate.
- Have you encountered any issues while using The Zeeguu Reader? If yes, please elaborate.
- Any other feedback you'd like to leave for us?

B.3. Final Survey Questions

- What is your name?
- What email did you sign up with?
- Besides your native language(s), how many other languages have you studied? (1-2, 3-4, +4)
- Have you tried using other extensions before using "The Zeeguu Reader"? (No, Yes, Don't know)

Your experiences using the extension

- How was your general experience with "The Zeeguu Reader"? (Scale, 1 = Very negative – 5 = very positive)
- Please elaborate and tell us more about why?
- How did you find using the extension in terms of easiness? (Scale, 1 = Not very easy – 5 = Very easy)
- Please elaborate and tell us more about why?
- How convenient did you find the extension for practicing a foreign language? (Scale, 1 = Not very convenient – 5 = Very convenient)
- Please elaborate and tell us more about why?
- How would you rate the usefulness of the extension? (Scale, 1 = Not very useful – 5 = Very useful)
- Please elaborate and tell us more about why?
- Do you feel like the extension contributed to your foreign language learning? (Scale, 1 = Not at all – 5 = Very much so)
- Please elaborate and tell us more about why?
- Does The Zeeguu Reader contribute to your foreign language learning in a way that other language learning resources you have used don't? (No the Zeeguu Reader does not contribute in another way, The Zeeguu Reader contributes the same as other learning resources, Yes The Zeeguu Reader contributes in another way, Don't know)
- If The Zeeguu Reader contributed to your foreign language learning in a way that other alternatives don't, how did it contribute differently?
- What did you like about the extension? Please provide examples. What did you not like about the extension? Please provide examples.
- Did you face any challenges while using the extension? If yes, what challenges did you face?

Finding articles

- How was it finding foreign language articles to read with the extension? (Scale, 1 = Difficult – 5 = Easy)
- Please elaborate and tell us more about why? How did you find articles to read with The Zeeguu Reader? (Through a search engine, On websites I knew beforehand, Through the recommendations on zeeguu.org)

- Did you find articles to read that were of your interest? If yes, how did you find them?
- How often was the language level in the articles too difficult? (Scale, 1 = Rarely – 5 = Very often)
- How often was the language level in the articles too easy? (Scale, 1 = Rarely – 5 = Very often)
- Was it a problem finding articles of the appropriate difficulty for you? Please elaborate
- Do you like that you get to choose any article to read from the web yourself? (Scale, 1 = No, not at all – 5 = Yes, very much)
- What were the advantages of finding your own articles?
- What were the disadvantages of finding your own articles?

Reading

- Do you prefer reading articles with the design and formatting of the The Zeeguu Reader or with the design and formatting of the original website? (The Zeeguu Reader, The website (where the article is published), I have no preference)
- Please elaborate and tell us more about why?
- Did you find it helpful for your reading experience that the extension only displays the article text and maybe an image - and not for example adverts, buttons and links? (Scale, 1 = Not helpful – 5 = Very helpful)
- Please elaborate and tell us more about why? Have you done any exercises in The Zeeguu Reader? (Yes, No, I did not find any exercises)

Exercises (Only displayed if the previous answer was yes)

- Did you find it useful that you could do exercises with the translated words after reading an article?(Scale, 1 = No, not useful – 5 = Yes, very useful)
- Please elaborate and tell us more about why?
- How do you find the workflow of reading articles, then reviewing the words and then doing vocabulary exercises? (Scale, 1 = Dislike – 5 = Like)
- Please elaborate and tell us more about why?

Final comments

- Can you see a way to improve The Zeeguu Reader? Would you recommend The Zeeguu Reader to others? (Scale, 1 = No, not at all – 5 = Yes, very much so)
- Do you have any other comments?
- Can we follow up in case we need further clarifications? (Yes, No)

B.4. Changes to Survey Questions

We initially did a pilot test of the final survey, so issues could be revealed before sending it to all participants. 25 users were invited, nine made an account and four answered the survey.

Based on this survey, the following changes were made:

1. We added one new question “Was it a problem finding articles of the appropriate difficulty for you? Please elaborate”. Because we needed more information on text difficulty.
2. We updated one question from “Do you prefer reading articles with The Zeeguu Reader or reading on the website where the articles originate from?” to “Do you prefer reading articles with the design and formatting of the The Zeeguu Reader or with the design and formatting of the original website?”. This change was made because we meant to gather information on the design and formatting, but the first question was not answered with focus on this aspect. Thus, we decided to emphasize on this.
3. We changed the configuration of the survey, so the users had to answer if they did exercises. And only if they answered yes, were they asked questions about the exercises.
4. In the pilot survey we asked the users to write how much time they spent answering the survey, so we could take this into consideration.

B.5. Survey Questions for Pre-Existing Users

- What is your name?
- What email did you sign up with on Zeeguu?
- Have you tried using other extensions before using "The Zeeguu Reader"? (No, Yes, Don't know)

The Zeeguu Reader compared to zeeguu.org

- How was the general experience using the extension compared to using zeeguu.org? (Scale, 1 = Worse experience than zeeguu.org – 5 = Better experience than zeeguu.org)
- How did you like the design of the article in The Zeeguu Reader compared to zeeguu.org (in regard to formatting, colors etc.) (Scale, 1 = Worse than zeeguu.org – 5 = Better than zeeguu.org)
- Was there any advantages to using the extension compared to zeeguu.org?
- Was there any disadvantages to using the extension compared to zeeguu.org?

Your experiences using the extension

- How was your general experience with "The Zeeguu Reader" (Scale, 1 = Very negative – 5 = Very positive)

- Please elaborate and tell us more about why? How convenient did you find the extension for practicing a foreign language? (Scale, 1 = Not very convenient – 5 = Very convenient)
- Please elaborate and tell us more about why?
- How would you rate the usefulness of the extension? (Scale, 1 = Not very useful – 5 = Very useful)
- Please elaborate and tell us more about why?
- Did you face any challenges while using the extension? If yes, what challenges did you face?

Finding articles

- How did you find articles to read with The Zeeguu Reader? (Through a search engine, On websites I knew beforehand, Through the recommendations on zeeguu.org)
- Do you like that you get to choose any article to read from the web yourself? (Scale, 1 = No, not at all – 5 = Yes, very much)
- What were the advantages of finding your own articles?
- What were the disadvantages of finding your own articles?

Final comments

- Can you see a way to improve The Zeeguu Reader?
- Would you recommend The Zeeguu Reader to others? (Scale, 1 = No, not at all – 5 = Yes, very much so)
- Do you have any other comments?
- Can we follow up in case we need further clarifications? (Yes, No)

C. Survey Answers

C.1. Initial Survey Answers

All answers from the 280 people who signed up can be found as "Anonymized initial survey answers (complete list).xlsx" at: <https://github.com/fribl/DataProcessing>

Answers from the 50 participants can be found as "Anonymized initial survey answers (participants).xlsx" at: <https://github.com/fribl/DataProcessing>

C.2. Issue Detection Survey Answers

The answers can be found as "Anonymized issue detection survey answers.xlsx" at:
<https://github.com/fribl/DataProcessing>

Findings from this survey

All five participants said the installation was easy, though one guy had issues reading articles in Hungarian. This was a language-specific bug, which was fixed. Another pointed out that sorting recommended articles on zeeguu.org was not working. This was also a bug, which was quickly resolved.

C.3. Final Survey Pilot Answers

The answers can be found as "Anonymized pilot final survey answers.xlsx" at:
<https://github.com/fribl/DataProcessing>

C.4. Final Survey Answers

The answers can be found as "Anonymized final survey answers.xlsx" and "Anonymized survey for pre-existing users.xlsx" at:
<https://github.com/fribl/DataProcessing>

D. Queries

A summary of each participants activity can be found as "Anonymized user activity data.xlsx" at: <https://github.com/fribl/DataProcessing>

The following user_ids are part of the experiment as new users:

3562, 3564, 3565, 3569, 3567, 3571, 3582, 3576, 3597, 3588, 3583, 3575, 3585, 3586, 3577, 3589, 3603, 3598, 3610, 3601, 3600, 3599, 3606, 3612,3615 ,3619, 3616, 3630, 3627, 3634, 3639, 3638, 3651, 3654, 3660, 3617, 3684, 3629, 3609, 3633, 3700, 3698, 3664, 3642, 3697, 3714, 3581, 3691, 3703, 3703.

The following user_ids are part of the experiment as new pre-existing users:
2705, 3524, 3526, 3572.

An anonymized version of the database can be found on GitHub. See the September 2022 database release at: <https://github.com/zeeguu-ecosystem/Data-Releases>

To replicate our findings the queries provided in this section can be used to fetch data from the anonymized database.

Two week activity data for new users

For each user id, run the following query with id, start and end date from this file: "Anonymized user activity data.xlsx":

```
1 SELECT * FROM user_activity_data
  WHERE time >= 'start date'
3 AND time <= 'end date +1'
  AND user_id=id;
```

Reading session data for pre-existing users

For each user id, run the following query with id, start and end date from this file: "Anonymized activity data pre-existing participants.xlsx":

```
SELECT * FROM user_reading_session
2 WHERE user_id= id
  AND start_time >= 'start date - 1'
4 AND last_action_time <= 'end date + 1';
```

Reading duration (new users)

For each user id, run the following query with ID and end date from this file: "Anonymized activity data pre-existing participants.xlsx":

```
1 SELECT SUM(duration) FROM user_reading_session
  WHERE user_id=id
3 AND start_time >= 'start date - 1'
  AND last_action_time <= 'end date + 1';
```

Exercise duration (new users)

For each user id, run the following query with id and end end date from this file: "Anonymized user activity data.xlsx":

```
SELECT SUM(duration) FROM user_exercise_session WHERE user_id=id
2 AND last_action_time <= 'end date + 1';
```

RSS feed data

The data from this query is further processed here: <https://github.com/fribl/DataProcessing>

```
SELECT id, rss_feed_id, FROM article
2 WHERE rss_feed_id IS NULL;
```

Reading session data after the experiment (new users only)

For each user id, run the following queries with id and end date from this file: "Anonymized user activity data.xlsx":

```
SELECT * FROM user_reading_session
2 WHERE user_id=id
  AND start_time >= 'end date - 1';
```

```
1 SELECT SUM(duration) FROM user_reading_session WHERE user_id=id
AND start_time >= 'end date - 1';
```

The data has been processed in a Jupyter Notebook:
<https://github.com/fribl/DataProcessing/blob/main/analysis.ipynb>, in Excel, and Google Sheets.

D.1. RSS Feed Data

The data used for Figure 30 can be found in "Anonymized RSS feed data.xlsx" at:
<https://github.com/fribl/DataProcessing>

E. Test Websites

E.1. Top 5 News Sites in Denmark, France, Germany, and UK

Top 5 news sites by country			
Danish	French	German	United Kingdom
bt.dk	leFigaro.fr	bild.de	bbc.co.uk
ekstrabladet.dk	leMonde.fr	spiegel.de	dailymail.co.uk
tv2.dk	ouest-france.fr	tagesschau.de	theguardian.com
dr.dk	20minutes.fr	focus.de	express.co.uk
bold.dk	l'Equipe.fr	welt.de	thesun.co.uk

Table 5: Top 5 News Sites in Denmark[28], France[1], Germany[69], and UK[73]

E.2. Top 10 Norwegian News Sites

Website	Is the main image found by Readability?	Is the main image found by our function getMainImage?
vg.no	No	Yes
nrk.no	No	Yes
dagbladet.no	Yes	Yes
tv2.no	Yes	Yes
nettavisen.no	No	Yes
e24.no	Yes	Yes
abcnyheter.no	No	Yes
aftenposten.no	Yes	Yes
dn.no	No	Yes
tek.no	Yes	Yes

Table 6: Top 10 Norwegian News Sites as of May 2022 [72]

E.3. Most Visited Danish Websites, and Whether They Needed Individualized Cleaning

Most visited Danish news sites in April 2022 [28]	Does it need individualized cleaning?
Ekstrabladet.dk	Minor changes
Bt.dk	Minor changes
Tv2.dk	No
Dr.dk	Major changes
Berlingske.dk	Major changes
Bold.dk	Minor changes
Politiken.dk	Minor changes
Seoghoer.dk	Major changess
Jyllands-posten.dk	No
Billedbladet.dk	Minor changes

Table 7: "No", meaning that the websites are cleaned completely by Readability and our general cleaning functions. "Minor changes", meaning that the websites could be improved by individualized cleaning. "Major changes" meaning the website needs to be cleaned because of problematic errors.

E.4. Most Recent Read Articles On Zeeguu.org and Whether Images are Fetched

Website	Is the main image found by Readability?	Is the main image found by our function getMainImage?
LeFigaro.fr	No	Yes
LeMonde.fr	Yes	Yes
Lexpress.fr	Yes	Yes
Bt.dk	No	Yes
Ing.dk	No	Yes
Marianne.net	No	Yes
Dr.dk	Yes	Yes
Nu.nl	No	Yes
Lequipe.fr	No	Yes
Telegraaf.nl	No	Yes
Theguardian.com	No	Yes
Bbc.com	Yes	Yes
Politiken.dk	No	Yes
Wired.com	No	Yes
Opendemocracy.net	No	Yes
Cnn.com	No	Yes
Faz.net	No	Yes
Trouw.nl	No	Yes

Table 8: The list was fetched in February 2022 and displays the sources of the 5000 most recent read articles on zeeguu.org and whether Readability and our funtion getMainImage can fetch the main image.