IT UNIVERSITY OF COPENHAGEN

Audio in digital learning environments

The effects on user experience when introducing audio in a customizable digital learning environment

BY JONAS KASTRUP KROHN & GUSTAV NØRGAARD PEDERSEN

jkak@itu.dk, gupe@itu.dk

Supervisor: Mircea Lungu

STADS Code: KISPECI1SE

Abstract

Language learning has gradually been digitized in conjunction with the advancements of technology. Furthermore, much research has been conducted on the impact of digital learning environments in relation language learning. However, some gaps appear in the research. The purpose of this paper is to investigate how learners' experience is influenced by exposing them to audio-based language exercises in a digital environment. Using questionnaires, this study analyzed how two groups of participants (*n=22*), one exposed to audio- and text-based exercises and one exposed to text-based exercises, both over a period of five days using the digital learning environment Zeeguu, rated and described their learning experiences and perceived learning outcomes. The exposure to audio-based exercises was found to facilitate a greater perceived learning outcome than that of text-based exercises. Furthermore, it was found that the increased diversity of the exercise pool increased level of entertainment and motivation amongst participants. The study implies that exposing learners to audio-based exercises increase the level of fun, motivation, and perceived learning outcome compared to that of learners only exposed to text-based exercises. Also, it was found that the quality of text-to-speech software greatly impact the learning experience.

Contents

1. Introduction

Language learning has gradually been digitized following technological advancements. Millions of people worldwide are currently learning a language and many of these do so through digital learning environments (Duolingo b, 2022; Babbel a, 2022). As language learning is an essential part of being human, researchers have sought to uncover the best methods for language learning and the effects these methods have on the learner's outcome. Moreso, the emphasis of these studies has turned to technological devices and learning environments as they've become a highly integrated part of everyday life. Many studies uncover how to increase motivation (Bárcena et al., 2016; Zhang & Zou, 2021; Schouten et al., 2016; Barata et al., 2015), or how introducing culturally relevant material increases learning outcome (Henry & Zerwekh, 2002; Bahrani, 2011), and others how language learners use personal devices to further their gains (Ko, 2017; Kim et al., 2013). Several established companies revolve solely around creating language learning environments that help the learner achieve the best possible results. They do so by utilizing knowledge accumulated through decades of research and applying it in digital settings.

One such system is Zeeguu, a digital language learning environment that provides the learner with customizable learning content and an opportunity to expand their vocabulary in a given language through language exercises (Zeeguu, 2022). Generally, digital learning environments rely on static and predefined lessons and content, but with Zeeguu, the content is chosen freely by the user from real-life articles in the target language.

Zeeguu started as a research project in 2016 and has since continuously been improved upon (Zeeguu, 2022). However, it does still have limitations. Users practice the words they have found to be difficult when reading articles, but the exercises only allow users to practice the text-based aspect, i.e., reading and writing. Research have found, that exposing a learner to audial input is beneficial, even potentially crucial to creating a complete understanding of the target language (Briton & Gaskill, 1978; Koondhar et al., 2018; Chen & Chang, 2011). Pioneers in commercialized digital language learning such as Duolingo or Babbel utilize audio features in varying ways to engage the learner audially (Duolingo a, 2022; Babbel a, 2022). Despite audio being an industry standard for language practice, not much research has been published on the benefits of introducing audio-based exercises in a digital learning environment. Therefore, this paper sets out to uncover the effects of introducing exercises that challenge the audial aspect of language learning. Initial literature exploration and knowledge gap lead us to the following research questions:

RQ1: How are users' learning experience affected by being exposed to audio-based exercises in a digital language learning environment with customizable learning material?

RQ2: How do users only exposed to text-based exercises rate and describe their learning experience, perceived learning outcome, exercise difficulty, and motivation compared to users exposed to both text-based and audio-based exercises?

To answer these questions, we created two audio-based exercises for Zeeguu with the intention of challenging the learners audially and improving their listening skills and phonemic awareness. To study this, an experiment was designed consisting of two groups of participants that were asked to complete exercise sessions daily on Zeeguu for a period of five days. Qualitative and quantitative data was collected through two elaborate questionnaires. Both questionnaires had five statements which participants from both groups had to rate and then elaborate on their rating, and the audio group had to answer an additional eight statements related to audio. The audio group were exposed to both audio-based and text-based exercises and the control group were only exposed to the original text-based exercises. With data collected from these questionnaires we hope to understand the impact of introducing audio-based exercises in a digital language learning environment and its effect on how participants perceive their language learning experience.

2. What is Zeeguu?

Zeeguu is an open-source language learning environment that allows users to practice a language using personalized learning material (Zeeguu, 2022). On Zeeguu, the web is crawled daily, and the user is presented with articles in their target language, or they can filter based on personal preference and a difficulty score calculated by an internal scoring algorithm in Zeeguu. The users can translate words they find difficult during reading. These words are then saved to be used in future sessions of vocabulary exercises. The exercises are all text-based and neither of them include any audio material during the session. The exercises will be presented in section 2.2.

2.1. How is Zeeguu different?

The high degree of customizability distinguishes Zeeguu from commonly known platforms such as Duolingo or Babbel and provides an interesting setting for studying the effects of including audio-based exercises in a customizable learning environment - something which is sparsely documented - as will be shown in section 4. We chose to conduct our research for this topic within Zeeguu, as the platform is an example of a custom-content learning environment that has already been proven to increase the learning outcome for high school students studying French (Lungu et al., 2018).

2.2. Existing exercises

This section will present the existing exercises of Zeeguu. There were three existing exercises with a minimalistic design and few interactive elements. They are fast to complete, intuitive, and simple for the user to engage with.

2.2.1. Find word in context

"Find word in context" presents the learner with a text paragraph in the language they're learning from an article they've read and a word from this text, translated to their native language. The user must then identify and click the translated word's corresponding translation in the paragraph or write it in the input field.
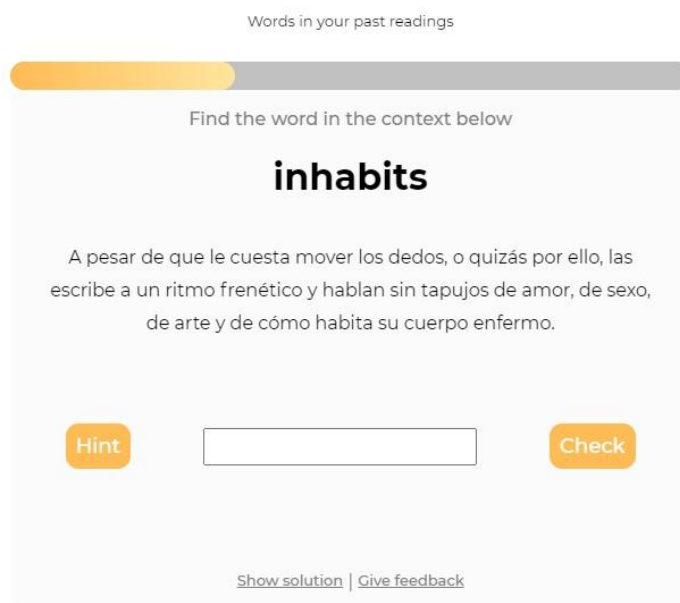
Figure 1. Find word in context exercise

2.2.2. Multiple choice

The "Multiple choice" exercise also gives the learner a text paragraph in the target language, but with a word removed and replaced with a line. The learner must then pick the correct word from the three options presented.
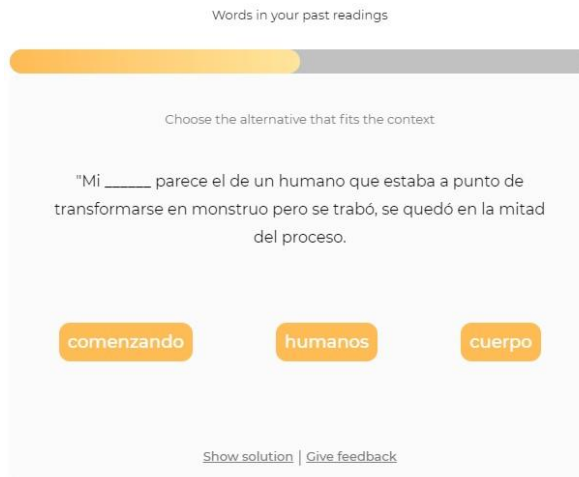
Figure 2. Multiple choice exercise

2.2.3. Match

Lastly, "Match" presents the learner three words in their target language and three words in their native language, which they then must correctly pair. After correctly pairing the words, the user can click the foreign language word to have it read aloud.
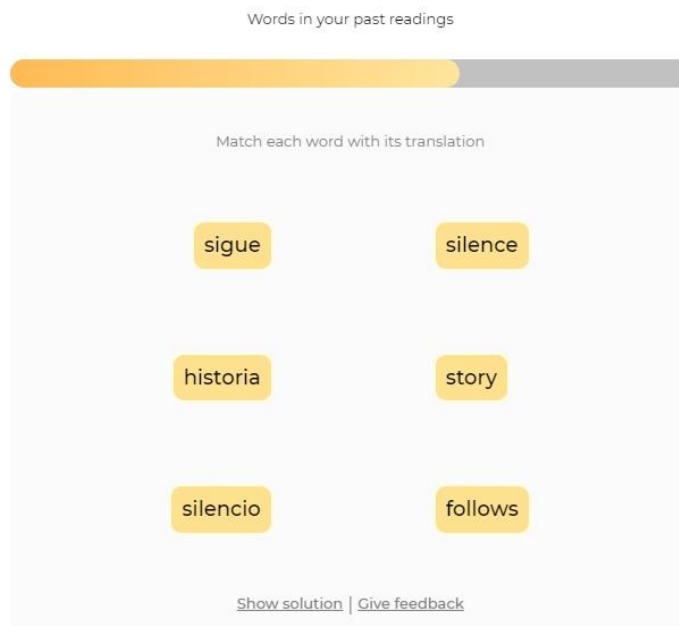
Figure 3. Match exercise

2.3. Existing exercise shortcomings

Similarly for all three exercises, when the correct word is clicked, written, or the participant has chosen "show solution", the corresponding translation will be shown, and they can click the speaker icon to have it read aloud. However, none of these exercises directly challenge the phonemic or listening capabilities of the learner. Several research articles relating to language learning showed that differentiating between exercise types and thereby increasing multimodality improves learner motivation (Griffith & Olson, 1992; Eutsler et al., 2020). Furthermore, the articles show that introducing audio-based exercises, thereby challenging the learner's

phonemic awareness, improves the ability to understand and distinguish spoken words (Griffith & Olson, 1992; Howles et al., 2014). This led to the assumption that a need for introducing exercises that support developing these aspects of language learning, would improve learning outcome.

3. Why include audio?

There exists a multitude of scientific papers that all contribute to the overall point that the inclusion of audio content or exercises is beneficial for the learner of a new language (Gjilakhani, 2016; Kondrateva, 2016; Renukadevi, 2016). In addition, Eutsler et al. outlines six domains in literacy, with the most fundamental one being phonemic awareness, the skill of combining certain sounds with certain letters (Eutsler et al., 2020). Phonemic awareness is considered a vital starting point for new learners and should ideally be prioritized highly, as it lays the foundation for mastering other parts of literacy (Griffith & Olsen, 1992; Briton & Gaskill, 1978). Koondhar references Krashen's learning theory, which argues that listening is the key for successful language learning and Chen & Chang describes listening as the most important communicative skill in relation to language learning, but ironically the hardest one to master (Koondhar et al., 2018; Chen & Chang, 2011).

3.1. Individual needs and increased stimulation

Another reason for including audio as an additional area of stimulation for the learner, is that learners often have individual needs when learning languages (Schouten et al., 2016; Fleming & Mills, 1992). Furthermore, Schouten puts the concept of audio in language learning into a digital context, citing the multimodal nature of digital media as a core strength (Schouten et al., 2016). The paper mentions how video, audio, and written text presented in conjunction can be beneficial, because it makes it easier to accommodate the learning experience to a learner's preferences (ibid.). Just as importantly, Toland and Thomas found that introducing audio files preceding or following classes had beneficial comprehension results for the students (Toland & Thomas, 2015). Many other favorable effects have also been shown to be caused by implementing audio in learning material (Harwood, 2014; Henry & Zerwekh, 2002; Eutsler et al., 2020; Howles, 2014; Zou & Zhang, 2021).

Conclusively, several reasons for the addition of audio-based exercises exist, including but not limited to practicing listening and phonemic awareness, which are fundamental skills in efficient language acquisition. Diversifying the learning environment from only text-based exercises to text- and audio-based will ensure that more learner-types' demands are accommodated, which in turn also could result in increased self-efficacy, autonomy, and motivation (Bauer, 2004).

4. Related work

When studying a foreign language there are several factors which the learner can't control but still impact their learning outcome. These factors are all part of the learning environment in which the learner seeks to achieve their goal. A learning environment consists of input, in the form of texts, video, audio, and/or combinations of these which the learner consumes. The curation of the correct learning material for the learner is important as any material not congruent with the learner's skill can inhibit their learning outcome and motivation (Zhang & Zou, 2021). The learning environment also contains features that allow the learner to practice, whether this be through communicative interactions, analog exercises, or digital exercises (Al-Khaza`leh, 2020; Witten et al., 2007). In summation: for the learner to completely master the desired language they must be exposed to a variety of opportunities that allows them to practice and consume every aspect of language learning.

As this study focuses on the experience of language learners when exposed to exercises that present audial challenges, we will not go into further details with elements not relating to this. However, as research on the effects of audio-based exercises in digital learning environments with customizable learning material is still very sparse, which was shown in the previous section and will be elaborated upon in this section, we will include studies which impose strategies for implementing new sensory input - thereby increasing modality - and/or exercises in other settings and studies relating to the advancements of digital learning environments. Consequently, some of the included studies show the causality between improving the quality and/or diversity of exercises and content and the positive impact on learners'' learning experience, but don't focus explicitly on audio.

4.1. Learning environments

Environments with a palette of different exercises already exist, such as Duolingo and Babbel, but not many public studies about the effect of these are available and the effects of these platforms' exercises are only showcased on the respective websites (Babbel a, 2022, Sporn et al., 2020; Duolingo b, 2022). Some studies with transparent scientific methods have been conducted, but only one of these is published by unaffiliated researchers (Eqbali & Nushi, 2017). Furthermore, the effects of exercises challenging specific aspects of language learning are not mentioned.

4.1.1. *Modifying learning environments through data forms*

Another way to modify the learning environment is through the data form on which the learning material is presented. Thomas and Toland, in an attempt to have learners review instructional material after classes, found that introducing learning material in audio format increased learners' motivation and perceived learning outcome (Toland & Thomas, 2015). Similarly other studies found motivation and learning outcome to increase when learners were exposed to audiovisual content in the learning environment (Howles, 2014; Zou & Zhang, 2021). While these studies don't focus on the implications of audio-based exercises, they contribute with findings that support the claim that modifying learning material and increasing multimodality results in positive gains for language learners.

4.1.2. *Designing the perfect learning environment*

Designing a perfect learning environment is difficult, and most likely impossible, as many different aspects of language learning can be modified, and every learner has different learning styles and needs (Fleming & Mills, 1992; Schouten et al., 2016). However, a way to accommodate the problem of often-predefined learning material and exercises a learning environment provides, is proposed by Witten et al. who created a platform that generates exercises based on the material submitted by the teacher (Witten et al., 2007). This approach aimed to improve the output of exercises and learning material but the exercises and learning material was text-based and therefore not multimodal (ibid.). It findings from the study did however indicate that learning outcome was increased.

4.2. *Implementing new exercises and exercises settings*

When we describe exercises in the context of related work it is not exclusively related to the implementation of new exercises in a vacuum. Some of the papers included in this section present many other learning environment modifications that are also implemented with the new exercises as part of a larger learning environment update. It, therefore, becomes difficult to identify and isolate the effects of every respective element introduced. The exercises designed and implemented in this study serve the purpose of positively impacting the learners' experience whilst allowing us to study the correlation between the new exercises and learner experience.

4.2.1. Diversity and learner motivation

As shown earlier, there is a correlation between the diversity of the learning environment's content and learners' motivation. Eutsler et al. conducted a literature review of 61 studies relating to the influence of mobile assisted language learning in elementary classrooms regarding literacy. In these studies learners were allowed, to varying degrees, to migrate their analog studies to mobile devices with various software installed (Eutsler et al., 2020). The researchers found that this generally increased the learners' motivation and self-efficacy, thereby increasing their time spent studying. However, not all studies included reported literacy gains (ibid.). Nevertheless, other studies support that integrating personal devices in learning environments is an enormous undertaking and often completely changes the nature of the learning experience, but that it does produce motivational, and autonomous benefits, which generally leads to better learning outcomes (Cordereo, 2017; Cumming-Potvin & Sanform, 2015).

Another study, with an analog approach, in which casual exercises amongst family members in the target language was incorporated during dinner time found that introducing additional domains for participants to practice increased their learning outcome (Bauer, 2004). Even though this approach is analog, it attributes to the fact that innovating the learning environment can positively impact learner self-confidence, motivation, and learning outcome (ibid.). Yet another study supports this claim and shows that increasing the exercise pool and emerging the learner into a rich customizable learning environment improves learner autonomy and motivation (Ostanina-Olszewsk, 2018).

4.3. Adding audio to digital learning materials

As found in the previous sections, audial input is tremendously important when learning a new language. Fortunately, there has been a large amount of research conducted on this topic and the benefits of increasing modality in a digital learning environment. Some studies investigate audio in relation to language learning, but not specifically audio-based exercises and its effect on learning experience. However, they find that integrating audio in conjunction with other input increases learners' learning outcome (Toland & Thomas, 2015; Witten et al., 2007).

Harwood found that 82% of study participants ($n=306$), agreed or strongly agreed to liking a mix of audio, video, and text, clearly showing that increasing multimodality, and including audio, have a positive impact on learners' perceived learning experience (Harwood, 2014). A finding that is further supported by Henry and Zerwekh who also found increasing modality heightens learner engagement and motivation (Henry & Zerwekh, 2002).

This section shows that numerous related studies exist on the topics of multimodality, digitizing learning environments and creating an experience that improves learners' motivation, autonomy, and learning outcome. Some of these articles also describe ways in which exercise structure has changed, but none of them report directly and isolated on the impact audio had on learners' experience.

5. Methodology

This section will detail how the audio-based exercises were designed and what considerations influenced this process. Furthermore, we will present the design of the questionnaires and any other methodological appliances introduced in the pursuit of answering the research questions. Designing the audio-based exercises was an essential part of the project. The new exercises needed to fit certain criteria to provide the most optimal environment for test subjects. As the language learning platform already had pre-existing exercises, the new exercises' design should fit the established standard, both visually, thematically and in terms of complexity. The source code of the Zeeguu system can be found on GitHub (GitHub a, 2022).

5.1. Initial design of audio exercises

When designing the audio-based exercises it became apparent that it might distort participant experience if they contained elements that the text-based ones didn't - other than audio. Therefore, effort was put into designing the audio-based exercises to be coherent with existing design thematic. With these considerations in mind, we then created the first iteration of audio-based exercises based on the existing design system.
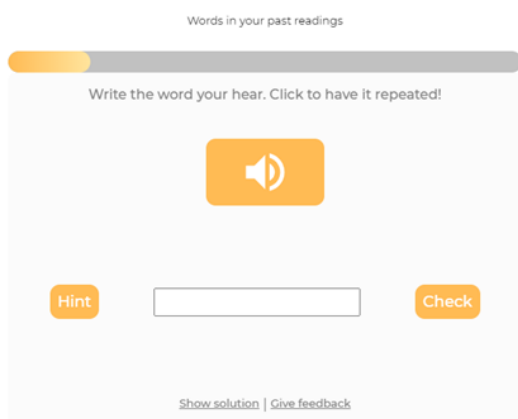
Figure 4. Initial spell what you hear exercise   Figure 5. Multiple choice audio exercise

For the audio element to become a significant part of the exercise experience we chose to design, develop, and implement two audio-based exercises. Introducing just one audio-based exercise might not have had a quantifiable impact on the participant experience whereas introducing three or more might have overshadowed the text-based element.

Figure 4 shows the exercise "Spell what you hear". The large button with a speaker icon reads the word aloud. The user can click it multiple times to have the word repeated. The user must then write the word spoken in the input field and click the "check" button to have their answer validated. Figure 5 shows the exercise "Multiple choice audio" which is heavily inspired by the "Multiple choice" exercise, but instead of having three clickable buttons with words on them, we created three buttons which would read aloud the respective word when clicked. The user must then click the button which reads aloud the word that fits in the context provided, and then confirm their selection by clicking the "check" button.

5.1.1. Audio-exercise one tweaks

After the initial design and implementation, we conducted some simple usability tests to collect data on the new exercises created. The usability studies consisted of four users completing two exercise sessions, whilst "thinking out loud". This gave us insight about the usability of the new exercises and was done to ensure that the audio-based exercises were as intuitive to users as the existing ones. However, it was found that "Spell what you hear" was significantly more difficult for users to complete as no context was provided. This prompted us to provide context from the original article, like that of "find word in context".

We then conducted a second round of testing with new users who reported to have an easier time completing "Spell what you hear", but it being more difficult than the existing text-based exercises (figure 6). Following the design constraints and reworking the exercises allowed us to expose participants to audio in the least invasive way possible.
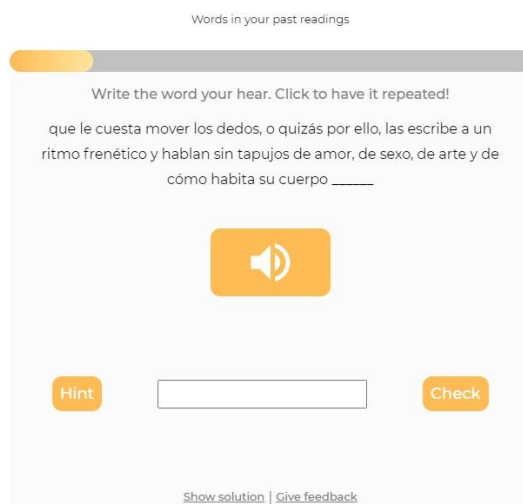

Figure 6. Redesigned spell what you hear exercise

5.2. Experiment description

When designing an experiment, it needs to be congruent with what the research question seeks to uncover and generate knowledge about. This section will describe the experiment design and how it allowed us to discover meaningful knowledge that will aid in answering this paper's research questions. Any methodological shortcomings and imperfections and how they influenced the experiment will be presented in the discussion section.

5.2.1. General study design

The study took place during the summer of 2022. The goal was to gain insight into how the introduction of audio-based exercises in the customizable learning environment Zeeguu influences the learning experience of the participants. It must however be explicitly stated that the intention of the study has not been to examine whether the exercises had an actual positive impact on learners' language competencies, as there already exist plentiful research on the importance of utilizing audial input when learning any language. Rather, the intention is to disclose how and if the participants perceived their experiences to have been influenced by introducing audio-based exercises and in what way.

To achieve this, we created two groups on the Zeeguu platform. The audio group would be exposed to the new audio-based exercises and the existing text-based exercises. The control group would only be exposed to the text-based exercises. The participants for each group were chosen randomly from a list of participants. The participants were asked to complete a pre-study survey in which we collected general participant information. Post-study the participants were asked to complete a questionnaire from which we collected their experience data.

5.3. Participant acquisition

Participants were recruited through posts on the social media Facebook and LinkedIn and had no prior knowledge of Zeeguu (See Appendix A). Initially 34 participants chose to join the study. 12 of these either never started the study, completed the study, or answered the final questionnaire. This leaves an *n* of 22 participants, and an *n* of 11 for each of the groups.

All participants who completed the study have Danish as their native language and the age groups with the biggest representations were 26-30 (50%), 21-25 (27.3%), and 51-55 and 56+ (9.1%). The most learned

languages were Spanish (23%), German (18%), and French (14%). Furthermore, participants' emails were recorded so that we can match the participant answers in the final questionnaire with data collected in the post-study questionnaires if necessary. Beyond the above-mentioned information, no other personal information was obtained from the users, as it isn't necessary to answer the research questions of this study.

5.4. Questionnaire design

This section presents the primary data collection approach. We'll go into detail about how and why the questionnaire questions and statements were formulated as they were, and the knowledge these design decisions are based upon.

The questionnaires' design was inspired with the principles described by Hinderks et al. in mind (Hinderks, 2017). They outline a user experience questionnaire that allows for the capture of meaningful data relating to the experience of the user by using an ordinal scale whilst capturing the impressions, feelings, and attitudes of the participants (ibid.). However, we did make some alterations to the questionnaire. Firstly, we expanded the ordinal scale from a seven-point scale to a ten-point scale, thereby granting the participants a larger scale for rating their experience. This ten-point scale should be understood as, five being neutral, one being "I don't agree", and ten being "I fully agree".

5.4.1. General considerations

As there are two groups of participants in this study two questionnaires were needed. The purpose of the questionnaires is to allow us to collect data about the participants' language learning experiences when using Zeeguu. When designing a questionnaire, it is important to maintain a simple language and phrase the questions in a way that facilitates a tangible frame for the participant to answer in relation to the scope of the study (Hinderks, 2017). Furthermore, for the purpose of the questionnaire design, it is important to include open-ended questions to facilitate a deeper understanding as to the learners' learning experience through the capture of *thick data* (Hinderks, 2017; Wang, 2013). This was achieved by adding descriptive open-ended questions allowing the participants to describe and elaborate on their experiences and ratings to the statements.

Five the statements were incorporated in both questionnaires to establish a common ground making data from the respective groups comparable. This would in turn allow for a deeper understanding of the impact from introducing audio in the exercise environment through a comparable analysis of the data collected.

The general information collected through the initial contact information survey might aid in the discovery of recurring themes and patterns during analysis. However, it was deemed unnecessary to collect any additional general participant information as it wouldn't have any impact on the quality of the study and served no purpose qua the design of the scope.

5.4.2. Audio group questionnaire

The questionnaire contained 13 statements, each with supplementing open descriptive paragraphs to capture detailed information about the statement in question. The participants had to rank to what degree they agreed with the statements on an ordinal-scale from 1-10.

The statements/questions were: 1) The Zeeguu exercises makes me more likely to study a foreign language often., 2) I believe completing Zeeguu exercises regularly would make me learn more words in the language I'm studying., 3) I believe completing Zeeguu exercises regularly would make me remember learned words more efficiently., 4) How fun did you find the exercises?, 5) How difficult did you find the exercises?, 6) The inclusion of audio-based exercises on Zeeguu makes me likely to study more often using the Zeeguu exercises., 7) I believe using Zeeguu with the inclusion of audio-based exercises would make me learn more words (versus no audio-based exercises), 8) I believe using Zeeguu with the inclusion of audio-based exercises would make me remember words more efficiently (versus no audio-based exercises)., 9) The audio-based exercises improved my exercise experience., 10) How difficult did you find this exercise compared to the existing text-based ones? ("Spell what you hear")., 11) How difficult did you find this exercise compared to the existing text-based ones? ("Multiple choice audio")., 12) The audio-based exercises are helpful for improving my pronunciation of words in the language I'm trying to learn., and 13) The audio-based exercises are helpful for improving my listening skills in the language I'm trying to learn. The questions comparing the exercises (questions 5 and 6) included figures of the respective exercises to aid the learner recall their experience better, hopefully generating more elaborate and precise answers.

Additionally, there were four more questions giving the participant the opportunity to elaborate: 1) What would improve your learning outcome when completing Zeeguu exercises?, 2) What changes/improvements to the Zeeguu exercises would make you more motivated to learn?, 3) What improvement would make the Zeeguu exercises more fun?, 4) Please elaborate if you have further comments regarding your experience with the exercises in general.

5.4.3. *Control group questionnaire*

The control group questionnaire had the same structure as the audio group's questionnaire, but any questions regarding the audio-based exercises were removed. Therefore, participants from this group were presented questions one through five, and the four follow up questions

5.5. *Distinguishing user-interface data from user-experience data*

This short section will describe what measures were deployed to minimize unwanted user-interface (UI) data. This is done since UI data isn't congruent with the scope of this study. However, due to the design and scope of the study, it has been assessed that there was a significant risk that participants would comment on UI aspects of Zeeguu, why countermeasures were implemented.

As shown in section 5.4.2 and 5.4.3, the only questions focusing on the usability of the new exercises are related to their difficulty level, i.e., user-experience aspects and not visual elements. The questions supplementing the statements are phrased in such a way it lets the participant describe their experience, how they perceive the exercises, and elaborate on their rating of the respective statements. Thereby we created a setting that nudged the respondents to elaborate on UX aspects of the exercises and not the UI aspects.

Even though the risk has been addressed and countermeasures implemented it isn't guaranteed that participants won't provide UI related feedback and/or observations to the questions in the questionnaire. However, whenever or if this is the case, these participants' observations will be defined as outliers in the data preparation-phase and therefore discarded and not included in the analysis. If however, it's the case that the remaining comment contains valuable information the comment will still be included, but any information that is related to UI will be excluded.

5.6. *Data analysis*

Data for this study is derived from the questionnaires:

1. Audio group questionnaire statement rating
2. Audio group questionnaire elaboration questions
3. Control group questionnaire statement rating
4. Control group questionnaire elaboration questions

With an *n* of 22 participants each statement and following elaboration resulted in a considerable amount of data. However, to be congruent with the scope of the study, our focus was mainly on ratings of statements, respective elaboration, any UX related comments, and how participants experienced and perceived language learning on the platform. Furthermore, focus was also put on any differentiation between the control group and the audio group in relation to their rating and comments.

6. Analysis

In the following section we will analyze the results from the questionnaires which include participants' rating of their language learning experience and any additional remarks they've made. These results will then lay a foundation for an analysis in which we elaborate and deduce the observations made in relation to the research questions. We will present all notable observations in the following subsections and finally summarize how they contribute to answering the research questions. Furthermore, the y-axis of the frequency histograms from Google forms had differentiating values, why we created new diagrams with the same y-axis value using the data collected (See Appendix B). The data is the same and all diagrams can be reconstructed by creating frequency histograms from the data. All data collected is available in a public repository on GitHub (GitHub b, 2022).

6.1. Audio increased participants' willingness to use Zeeguu for studying

Observation 1: *We observed that participants exposed to audio-based exercises reported that the learning environment was more likely to make them study more.*
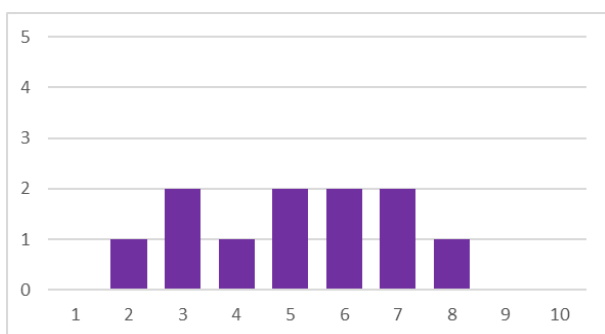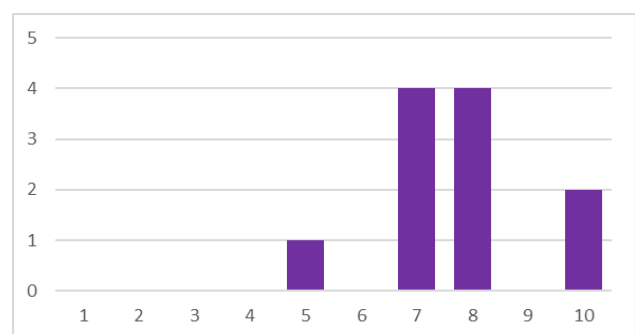

Figure 7. Frequency histogram of control group

Figure 8. Frequency histogram of audio group

Figure 7 and figure 8 show the respective groups' ratings. Participants in the audio group responded with a 7.73 mean and 1.42 standard deviation (SD) whereas participants in the control group averaged 5.1 and 1.92 SD when asked the question: "*The Zeeguu exercises makes me more likely to study a foreign language often.*" (See Appendix B). This observation is an indicator, which is supported by the data collected from other questions, that the audio group had an overall more motivating experience. However, this finding isn't surprising, as the Aiki study found that users expressed, they were missing variation in exercise types - participants from their study were exposed to the same exercise pool as the control group of this study (Inie & Lungu, 2021).

The rating of this specific question is influenced by multiple factors from everywhere in the learning environment. Therefore, it serves as a means for the participants to rate much of their overall experience. Later in the questionnaire the participants rate and elaborate on many other aspects of the learning environment and their learning experience, all of which can be seen as emergent parts of a whole.

Introducing audio-based exercises did increase the exercise variation, which might have inclined some participants to dismiss the wish of grammatical exercises, which participants from the Aiki study requested as they found the exercise sessions more varied and challenging (Inie & Lungu, 2021). This finding is further backed by the five participants from the control group that explicitly requests more variation in the exercise pool.

6.2. Audio increased difficulty

Observation 2: *Overall, users exposed to the audio-based exercises found the experience to be more difficult than the control group.*

Figure 9. Frequency histogram of control group
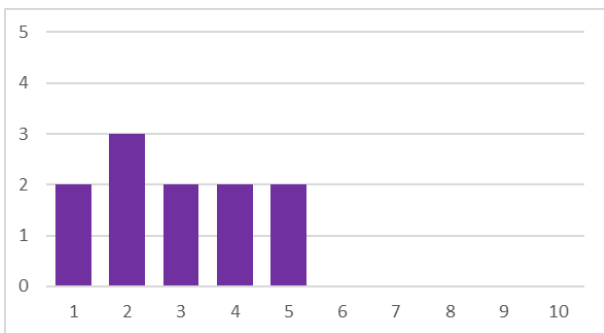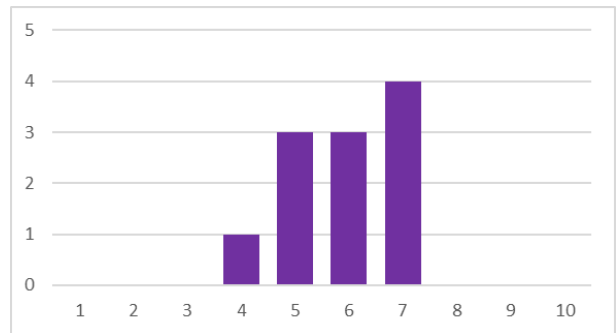

Figure 10. Frequency histogram of audio group

For the question: "*How difficult did you find the exercises?*", the audio group reported an average of 5.9 with a median of 6 whereas the control group only reported an average of 2.9 with a median of 3, answers shown in figure 9 and figure 10. This is a significant difference and the higher rating from the audio group can be

seen in direct correlation to the introduction of audio-based exercises. When participants from the audio group had to answer: "*How difficult did you find this exercise compared to the existing text-based ones*" they averaged a mean of 7.54 rating for "Spell what you hear" and "Multiple choice audio" was rated 5.63 (See Appendix B). Both exercises were rated significantly more difficult than the total difficulty of the control group in relation to the exercises they were exposed to. However, participants from the control group reported that the exercises were too easy: "*It was not challenging. The exercises were too easy.*", "*They* (the exercises) *were very straightforward and not difficult for me to complete*" (See Appendix C). These statements indicate that exercises that do not challenge the user might result in the user becoming bored and/or losing motivation. Furthermore, participants from the audio group elaborated on their experience: "*Not hard, but hard enough*", "*(...) I did not feel they were too difficult for my level, and I believe they have to be a little difficult if I am to learn from them.*", "*I think they weren't too difficult, but just the right amount.*" (See Appendix D). This further solidifies the assumption that a higher, but more appropriate, level of difficulty improves the learners' experience which support the findings by Zhang and Zou (Zhang & Zou, 2021).

An additional factor that possibly influenced the higher rating of difficulty in the audio group was that several participants reported on a technical aspect that made the experience more difficult: the text-to-speech voice was sometimes hard to understand. This could potentially be mitigated by implementing text-to-speech software of higher quality, something there will be elaborated upon later (See Appendix E).

6.3. Audio increased perceived level of fun

Observation 3: *Overall, users exposed to the audio-based exercises reported that their experiences were more fun than the control group.*
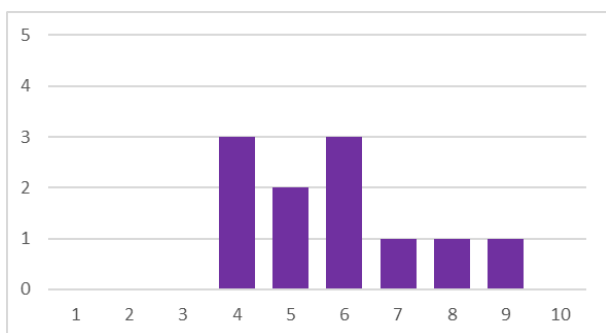
Figure 11. Frequency histogram of control group
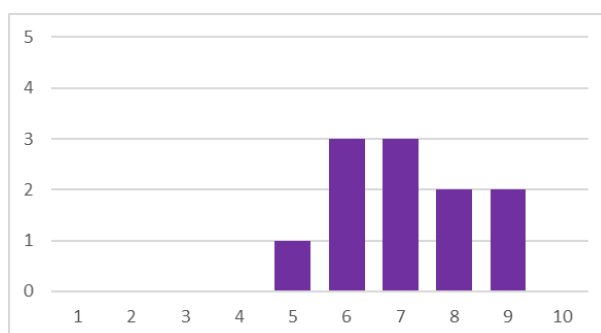
Figure 12. Frequency histogram of audio group

Participants in the audio group reported an average of 7.1 with a SD of 1.3 when asked to rate how fun they thought the exercises were overall, answers shown in figure 11 and figure 12. For the same question, the control group reported an average of 5.8 with a SD of 1.66 (See Appendix B). Around half of participants in the control group cited a lack of variance in the exercise types as the main reason for not scoring higher. Some of the

participants in the control group also reported low difficulty as making the exercises too little of a challenge, and therefore less fun (See Appendix F). Furthermore, the lower SD in the audio group indicates a higher level of agreement to the statement.

The audio group also had a small number of users comment on the lack of variance in the exercise types as a reason for a lower score. Since the audio group was presented with more types of exercises than the control group it's evident that fewer participants would feel a lack of variance. This is especially true as the control group exercise pool doesn't include audio-based exercises, which limits the sensory stimulation of the participants. As described earlier, when learning a language, different kinds of stimuli are important for the learner to master it completely. As Zeeguu is mainly focused on increasing the learners' vocabulary, this finding indicates that developing an exercise pool that stimulates multiple senses will have a positive impact on the acquisition of new words and the retention of these through higher learning motivation.

Overall, the response from the participants shows that adding audio exercises to the text-based ones will have a positive effect, because they create a more dynamic learning environment. As one user in the audio group puts it: "*It just seems like a fuller experience*" (See Appendix E). This is congruent with findings from studies presented earlier in this report. Firstly, when introducing exercises that stimulate the auditory sense the exercise pool becomes more satisfactory to users who are predisposed to audio-based learning. This is congruent with the findings of Fleming & Mills and Schouten et al. (Fleming & Mills, 1992; Schouten et al., 2016). Secondly, as mentioned earlier, it creates a more varied and challenging experience which in turn increases user motivation.

6.4. Audio increased learner motivation

Observation 4: *When asked about them, users exposed to audio-based exercises generally found them motivating.*

When the audio group was asked how much they agreed with the statement "*The inclusion of audio-based exercises on Zeeguu makes me likely to study more often using the Zeeguu exercises*", most users either completely agreed with the statement or somewhat agreed. An average of 6.8 were reported with a median of 8. However, the ratings had a SD of 3.15, the highest calculated SD from all statements from both groups (See Appendix B). The high SD can partly be attributed to one user who did not agree with the statement at all but reported that their answer was not an indication of the audio-based exercises having a negative effect, just that they had no positive effect for the participant's experience.
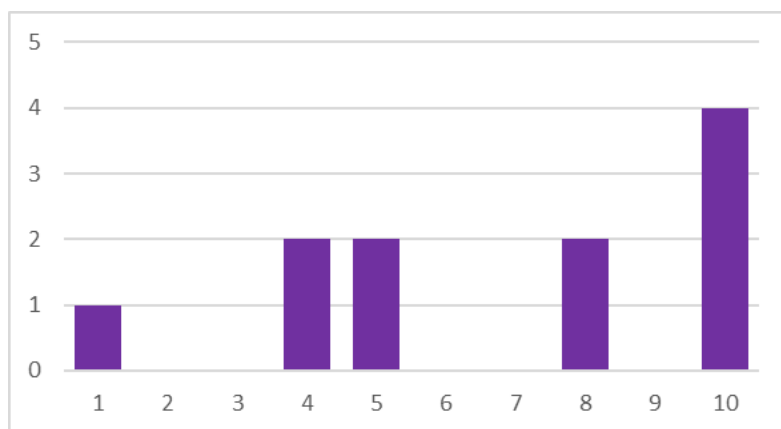
Figure 13. Frequency histogram of audio group

The average indicates that participants found the inclusion of audio-based exercises motivating and by analyzing the SD and the elaborative comment for the participant who rated '1' we can establish that one participant, who misunderstood the concept of the radial-scale (as '5.5' is truly neutral) resulted in a lower rating. Furthermore, there is five participants who rated significantly lower than the median. This again talks into the concept of accommodating multiple learner types by creating a diverse environment as four participants requested varying degrees of gamification. It can be argued that while introducing audio allow Zeeguu to accommodate more learner types, an increase in methods implemented to create an even more diverse environment can be applied to meet this criticism.

6.5. Audio increased perceived potential learning outcome

Observation 5: *Overall, users exposed to the audio-based exercises felt they had a better potential learning outcome doing the exercises, than users in the control group.*

Participants in the control group reported 6.72 average for the question "*I believe completing Zeeguu exercises regularly would make me learn more words in the language I'm studying*" and 7.45 for "*I believe completing Zeeguu exercises regularly would make me remember learned words more efficiently*".

Participants from the audio group reported 8.9 and 8.9 respectively and 7.27 to the question "*I believe using Zeeguu with the inclusion of audio-based exercises would make me learn more words (versus no audio-based exercises)*", 8.45 to "*I believe using Zeeguu with the inclusion of audio-based exercises would make me remember words more efficiently (versus no audio-based exercises).*", and 8 for "*The audio-based exercises improved my exercise experience.*" (See Appendix B).

The data implies that users from the control group find the Zeeguu environment to be beneficial to their learning, but even more so does the audio group. Additionally, the audio group rated that they believe

they will learn more words and the retention of these words to be higher than the control, meaning that they perceived their learning outcome over a given period to yield a higher.

The audio group is presented with questions which concern their perception of how they experience the learning outcome in contrast to if there were no audio-based exercises: "*I believe using Zeeguu with the inclusion of audio-based exercises would make me remember words more efficiently (versus no audio-based exercises)*" and "*I believe using Zeeguu with the inclusion of audio-based exercises would make me learn more words (versus no audio-based exercises)*". These questions are meaningful to introduce as the participants also have been exposed to text-based exercises, allowing them to imagine how the exercise experience would have been without audio-based exercises. The audio group ratings for these questions (averaged 8.45 and 7.27) indicates that they perceive their learning outcome as significantly greater when audio is incorporated in the exercise pool, thereby backing the scoring difference from the control group's rating of the question relating to number of words learned and how well they believe they will retain them (See Appendix B).

Through the analysis of the audio group questionnaire, it became apparent that the mix of written and audio-based exercises positively influenced the learning experience of the participants. They reported higher averaged ratings for all numeric statements and their comments were overall positive. Furthermore, for all five statements both groups rated, the control group had an averaged SD of 1.82 and the audio group had a 1.28 average SD, indicating a higher level of coherence across participants' experiences (See Appendix B).

6.6. Quality of text-to-speech impacted user experience

Observation 6: *The quality of the text-to-speech software is very impactful on user experience with the audio-based exercises.*

One of the main points of criticism deducted from the audio group was that some words could be difficult to hear properly, when spoken by the text-to-speech software. This frustrated some users, especially in relation to "Spell what you hear", where the user is completely reliant on understanding the spoken word. Five participants from the audio group mentioned at some point that they struggled with hearing the words properly (See Appendix D. Furthermore, it became apparent that the quality of the pronunciation is impacted by the language the user had chosen. German has letters that are pronounced differently than in other languages (such as a word starting with 'v' will be pronounced as a 'f'). One participant seemed especially frustrated with the pronunciation and several times addressed it by saying "*Poor machine pronunciation*" and "*The German pronunciation was poor.*" (See Appendix E). This comment is further backed by seven other comments where users to varying degrees criticize the German pronunciation. However, even though participants criticized the

pronunciation they still thought it was a good way to support and develop good pronunciation: "*It is fun, and it supports correct pronunciation.*" which is further backed by the differences reported to the question "how fun did you find the exercises" (See Appendix D). While the audio-based exercises clearly had a positive impact on the participants' learning experience it could arguably have had an even more positive effect if the text-to-speech software was of higher quality.

6.7. Audio's effect on users' experience

This section will put the findings of the analysis in relation to the first research question.

> **RQ1:** *How are users' learning experience affected by being exposed to audio-based exercises in a digital language learning environment with customizable learning material?*

Through this analysis it has been shown that introducing audio has a significant impact across multiple parameters relating to user experience. Though several participants reported they were challenged by the quality of the text-to-speech software, they responded overwhelmingly positively to the inclusion of audio-based exercises. As shown in the analysis, introducing audio-based exercises increased the general difficulty of the exercise sessions significantly, which can be seen as a result of more aspects of their language learning being challenged, but it must also be considered that the quality of the text-to-speech software have had an impact on these results. Furthermore, audio group users expressed that their perceived learning outcome, perceived benefits of using Zeeguu for studying, level of fun, and motivation was positively impacted by the audio-based exercises.

6.8. The control group's user-experience compared to the audio groups'

This section will present the results of the analysis and how it correlates to the second research question:

> **RQ2:** *How do users only exposed to text-based exercises rate and describe their learning experience, perceived learning outcome, exercise difficulty, and motivation compared to users exposed to both text-based and audio-based exercises?*

It has been shown that the introduction of audio has had a positive impact on the audio group's user experience. Furthermore, it was shown, through the results of both questionnaires, that participants exposed to audio rated their experience more positively than the control group across all statements. This includes but might not be

limited to, user motivation, likelihood to use Zeeguu for studying, perceived acquisition of number of new words and retention thereof, motivation, and level of fun was increased by introducing audio-based exercises. It was also found, despite the control groups' lower rating, that they did find that Zeeguu would improve their learning outcome and motivation for studying a language.

7. Discussion

The following section will discuss the methodological approaches applied throughout this study. This is done to identify any shortcomings, and to discuss how alternative methodological approaches could have impacted the results and data.

7.1. Questionnaire criticism and shortcomings

The data acquired through this study has been gathered through questionnaires and as such, makes up for the bulk of the methodological structure. Therefore, these questionnaires should undergo scrutiny when discussing and evaluating this study. First point of criticism for the statements is that the respondents, when asked to rate on a scale from 1-10 how much they agreed with a given statement, were given statements that were positively phrased, such as "*The Zeeguu exercises makes me more likely to study a foreign language often.*". The problem is that a respondent that doesn't feel compelled to study a language Zeeguu likely will disagree with the statement, and a different respondent who also disagrees with the statement but for a different reason, also would give it a low rating. Therefore, two respondents who disagree with the statement, but based on two different thoughts of reasoning, could end up giving the same rating. This means that when analyzing the quantitative data, i.e., rating of statements, it became clear the edge cases (ratings that had a relatively high SD) needed to be studied closer and taken into considerations if they prompted reason for doing so. Furthermore, by wording the statements positively we insinuate that the statement is correct. This might therefore result in the respondents being more inclined to answer in conjunction with the statement rather than their honest opinion.

To accommodate this problem, we allowed all respondents to elaborate their rating. This allowed us to understand the reasoning as to why a given respondent gave the respective rating. Through the analysis of the questionnaires, it also became apparent that respondents that rated either very low or very high for a given statement almost always explained their reasoning, which aided during the analysis of the data. However, it does not fully compensate for the positive phrasing, and ratings can therefore be expected to be higher than they would have been if the statements were neutral.

7.1.1. Ordinal scale applied

Another aspect of the questionnaires that could have minimized rating distortion, was adding a rating option of '0'. This could have proved to be beneficial as it would make a rating of five truly neutral whereas in the iteration of the questionnaire used in this study, 5.5 would be the actual true neutral. While we as researchers are aware of this, participants might have naturally gravitated towards five as a neutral answer due to the range of the ordinal scale, thereby distorting the ratings. This might have contributed to respondents giving a higher rating.

7.2. Learning environment shortcomings

Furthermore, when reading this study, it is of utmost importance to recognize the findings won't be applicable or replicable across all digital learning platforms. As mentioned, Zeeguu is a micro learning environment that focuses on increasing the learner's vocabulary, i.e., what these words mean, pronunciation (with the introduction of audio-based exercises) and the context of their usage (Zeeguu, 2022). This is different from many mainstream providers, including but not limited to Duolingo and Babbel (Duolingo a, 2022; Babbel a, 2022), who focus on increasing knowledge of the language across additional parameters, such as grammar. Though much material from these platforms revolves around audio, it is within a static context with pre-curated learning material, different practice facilities (Duolingo a, 2022).

7.2.1. Customizable learning material

Another aspect that is important to address is the learning material Zeeguu allows for. Using Zeeguu the learner is able to construct a custom library of learning content based on their interests by importing articles, or choosing from a pre-curated selection of articles, which they can read and select words from to practice from. This means Zeeguu provides an opportunity for the learner to practice on material solely selected by the learner themselves. While it has been shown that customizable material increases learner motivation the scope of this study does not encompass the effects of this aspect (though some data revolving this emerged through the analysis), thus it shouldn't be excluded that it might have affected the participants' experiences. Consequently, this further emphasizes that findings from this study might not be applicable to platforms with learning material provided differently than that of Zeeguu. The findings should hence be read in the context of this study and are not universally applicable across all platforms.

7.3. Varying exercise availability

One of the main findings in the analysis is that the users in the audio group reported having a more fun experience than users in the control group. In the control group, users reported a lack of variety in exercises as one of the reasons for their lower rating. To determine the full effect of having audio-based and text-based exercises mixed versus having only text-based exercises, it would likely have been beneficial for the study to make sure that the variety in exercises were the same. As an example, if two additional text-based exercises had been designed, implemented, and given to the control group, such that both groups were exposed to five different exercise types, instead of five for the audio group and only three for the control group. This might have resulted in the control group giving higher ratings to questions concerned with difficulty and amount of fun.

8. Future work

As mentioned in the previous section, the positive results of the audio group, in relation to perceived learning outcome and how beneficial they believed using Zeeguu is, might have been distorted by the small exercise roster of the control group. Therefore, designing and conducting a follow-up study revolving around a similar structure but with the introduction of additional text-based exercises for the control group, could either cement or challenge the findings of this study. Furthermore, the study should contain statements that are neutrally phrased to minimize distortion of participants' ratings.

Multiple participants reported that the text-to-speech software was of low quality, especially concerning languages with difficult pronunciation. However, as mentioned earlier, as Zeeguu's Learning content is highly customizable and vast, it would prove practically impossible to record all articles to provide the user more correct pronunciation. It could be argued that when auto generated text-to-speech software has been improved, a study with the same underlying setup could be conducted. It would in this case be interesting to compare the participants' ratings of the statements from this study and the new study.

Lastly, it became apparent that any conclusions of this paper would benefit from being substantiated from a larger volume of data. Therefore, re-iterating this study with more participants and possibly over a longer period while considering the previous criticism, could further elaborate on the findings.

9. Conclusion

This paper examined the effects of exposing users to audio-based exercises in a customizable language learning environment. Starting from the point of the supplementary research question: "*How do users only exposed to text-based exercises rate and describe their learning experience, perceived learning outcome, exercise difficulty, and motivation compared to users exposed to both text-based and audio-based exercises?*", the findings conclude that users only exposed to text-based exercises felt that their learning experience was less efficient, less challenging, and less fun and motivating than participants from the audio group. They felt their exercise sessions were more repetitive and reported lower likelihood of using the platform for language learning than the audio group participants. Furthermore, participants from the control group rated noticeably lower regarding perceived learning outcome in relation to vocabulary acquisition and retainment.

By answering this question, we can then answer our main research question: "*How are users' learning experience affected by being exposed to audio-based exercises in a digital language learning environment with customizable learning material?*". The study found that users' learning experience is largely affected positively by the exposure to audio-based exercises on the learning platform Zeeguu. Participants exposed to audio reported the level of difficulty to be more appropriately challenging, in contrast to the control group who reported it to be too easy. In relation to this, they also reported that they found the exercises more fun and motivating. They perceived their learning outcome to be greater than that of the control group in relation to the number of words learned and how well they'd retain them. They also highly agreed that audio-based exercises are helpful for improving listening skills and pronunciation. They addressed that the Zeeguu exercises make them more likely to study using Zeeguu than the control group.

Lastly, it can be concluded that a poor quality of text-to-speech software is likely to have a negative impact on user experience.

10.   References

Al-Khaza`leh, Bilal., (2020), Exploring Language Learning Strategies of Saudi EFL Learners at Shaqra University, Saudi Arabia, College of Science and Humanities, Shaqra University, Australian International Academic Centre, PTY.LTD, Advances in Language and Literacy Studies, Vol. 10, No. 4, pp 63-71

Babbel a., (2022), Learn a language, have a real conversation, in www.babbel.com, retrieved 15th of August (2022), from https://begin.babbel.com/en_gb/new-multilanguage-selection/

Bahrani, T., (2011), The role of audiovisual media news in language learning, Islamic Azad university, Iran, in English language teaching, vol. 4, no. 2, pp 260-266

Bárcena, E., Read, T., Kukulska-Hulme, A., (2016), Exploring the application of a conceptual framework in a social MALL app, in New perspectives on teaching and working with languages in the digital era, pp 223-232

Barata, G., Gama, S., Jorge, J., Gonclaves, D., (2015), Gamification for smarter learning: tales from the trenches, Springer. in Smart learning environments, vol. 2, no. 10, pp 1-23

Bauer, Angelika., Kulke, Florian., (2004), Language Exercises for Dinner: Aspects of Aphasia Management in Family Settings, Psychology Press LTD, Aphasiology, Vol. 18. No. 12, pp 1135-1160

Brinton, D., Gaskill, W., (1978), Using news broadcasts in the ESL/EFL classroom, Tesol Quarterly, vol. 12, no. 4, pp 403-413

Chang, Chi-Cheng., Chen, I-Jung., (2011), Content presentation modes in mobile language listening tasks: English proficiency as a moderator, Routledge - Taylor and Francis group, in Computer assisted language learning, vol. 24, no. 5, pp 451-470

Cordero, Kristina., Nussbaum, Miguel., Ibaseta, Valentina., Otaíza, Maria, José., Chiuminatto, Pablo., (2017), Read, Write, Touch: Co-Construction and Multiliteracies in a Third-Grade Digital Writing Exercise, Journal of Computer Assisted Learning (2018), pp 1-12

Cumming-Potvin, W., & Sanford, K., (2015)., Countering a 'back-to-basics' approach to teacher education: Multiliteracies and on-line discussions in a community of practice. Language and Literacy: A Canadian Educational e-journal, 17(1), 21–41.

Duolingo a., (2022), Duolingo approach, in www.duolingo.com, retrieved 2nd of August (2022), from https://www.duolingo.com/approach

Duolingo b., (2022), Duolingo research, in www.duolingo.com, retrieved 2nd of August (2022), from https://research.duolingo.com/

Eutsler, L., Mitchell, C., Stamm, B., Kogut, A., (2020), The influence of mobile technologies on preschool and elementary children's literacy achievement: a systematic review spanning 2007-2019, AECT, in Education tech research dev, vol. 68, pp 1739-1768

Eqbali, M, H., Nushi, M., (2017), Duolingo: A mobile application to assist second language learning, Shahid Beheshti University, in Teaching English with technology, vol. 1, pp 89-98

Fleming, N, D., Mills, C., (1992), Not another inventory, Rather a catalyst for reflection, Lincoln University, New Zealand, in To improve academy, vol. 11, pp 137-155

Gilakjani, Abbas., Sabouri, Narjes., (2016), The Significance of Listening Comprehension in English Language Teaching, Department of English Language Translation, Islamic Azad University, Iran, Theory and Practice in Language Studies, Vol. 6, No. 8, pp 1670-1677

GitHub a., (2022)., retrieved from https://github.com/zeeguu-ecosystem

GitHub b., (2022)., retrieved from https://github.com/zeeguu-ecosystem/studies

Griffith, P, L., Olson, M, W., (1992), Phonemic Awareness Helps Beginning Readers Break the Code, The Reading Teacher, vol. 45, no. 7, pp 516–523

Harwood, Chris., (2014), Personal Learning Environments: Using SymbalooEDU in Learning English for Academic Purposes, ESP Today, Ontario Institute for Studies in Education, University of Toronto, Canada, Vol 2., pp 199-215

Henry, George., Zerwekh, Robert., (2002), SEAsite: Web-based interactive learning resources for Southeast Asian languages and cultures, in CALICO journal, vol. 19, no. 3, pp 499-512

Hinderks, Andreas., Schrepp, Martin., Thomaschewski, Jörg., (2017), Design and Evaluation of a Short Version of the User Experience Questionnaire, International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 4, No. 6, pp 103-108

Howles, L., Lake, W, M., Garret-Rucks, P., (2014), Enhancing L2 reading comprehension with hypermedia texts: student perceptions,Equinox publishing, in Computer assisted language instruction consortium, vol. 32, no. 1, pp 26-51

Kim, D., Kim, D, J., Rueckert, D., Seo, D., (2013), Students' perceptions and experiences of mobile learning, in Language learning and technology, vol. 17, no. 3, pp 52-73

Ko., Myong-Hee., (2017), Learner perspectives regarding device type in technology-assisted language learning, Routledge - Taylor and Francis group, in Computer assisted language learning, vol. 30, no. 8, pp 844-863

Kondrateva, Irina., Safina, S, Minnisa., Valeev, A, Agzam., (2016) Listening as a Method of Learning a Foreign Language at the Non-Language Faculty of University, Kazan (Volga Region) Federal University, Russia, International Journal of Environmental & Science Education, Vol. 11, No. 6, pp 1049-1058

Koondhar, M., Elizabeth, M, A., Insaf, A, Siming., Tariq, H, Umrani., (2018), Language learning approaches: Unity in diversity, in Australian International Academic Centre, vol. 9, no. 6, pp 34-38

Lungu, Mircea, Filip., Brand, Luc van der., Chirtoaca, Dan., Avagyan, Martin., (2018), As We May Study: Towards the Web as a Personalized Language, Textbook. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI 2018, Montreal, QC, Canada, April 21-26, 2018, 338.

Inie, Nanna., Lungu, Mircea, Filip., (2021), Aiki - Turning Online Procrastination into Microlearning. In CHI '21 - Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, May 2021, Yokohama, Japan, May 8–13, 2021, Yokohama, Japan. ACM, NewYork, NY, USA, 21 pages

Ostania-Olszewska, Julia., (2018), Uniwersytet Pedagogiczny w Krakowie, Modern Technology in Language Learning and Teaching, pp 153-164

Paris, Tengku., Manap, Nurma., Abas, Hafiza., Ling, Lo., (2021), Mobile-Assisted Language Learning (MALL) in Language Learning, Journal of ASIAN Behavioural Studies, Vol. 6, No. 16, (2021) pp 61-73

Renukadevi, D., (2014), The Role of Listening in Language Acquisition; the Challenges and Strategies in Teaching Listening, International Journal of Education and Information Studies, Vol. 4, No. 1, pp 59-63

Schouten, G, M, Dylan., Smets, J, J, Nanja., Driessen, M., Fugri, K., Neerincx, M, A., Cremers, A, H, M., (2017), Requirements for a virtual environment to support the social participation education of low-literates, CrossMark, in Univ access inf soc, vol. 16, pp 681-698

Sporn, Z., Chanter, J., Meehan, D., (2020), Babbel language learning podcast, in iJAC, vol. 13, no. 3, p 43-49

Toland, H, S., Thomas, Simon., (2015), Imitating podcasts by providing audio content to support and enhance language learning, in The JALT CALL journal, pp 3-17

Witten, Ian., Wu, Shaoqun., Aquino, Rosane, Schenkel., Edwards, Arthur., Nichols, David., (2007), A Digital Library of Language Learning Exercises, Department of Computer Science, University of Waikato, New Zealand, iJET International Journal of Emerging Technologies in Learning, pp 1-7

Zeeguu., (2022), Zeeguu landing page, in www.zeeguu.org, retrieved 10th of June (2022), from https://www.zeeguu.org

Zhang, Ruofei., Zou, Di., (2021), A state-of-the art review of modes and effectiveness of multimedia input for second and foreign language learning, Routledge - Taylor and Francis group, in Computer assisted language learning, pp 1-28

Wang, Tricia., (2013), Big data needs thick data. Ethnography Matters 13 (2013).

11.   Appendices

11.1.   Appendix A - Social media post

11.2.   Appendix B - Questionnaire data

11.3.   Appendix C - Summary questionnaire data of the control group

11.4.   Appendix D - Summary questionnaire data of the audio group

11.5.   Appendix E - Individual questionnaire data of the audio group

11.6.   Appendix F - Individual questionnaire data of the control group