Computer Science BSc

COC251

B614743

LEARNING A NEW LANGUAGE USING A VR ENVIRONMENT

by

Peter Smith

Supervisor: Dr Firat Batmaz

Department of Computer Science
Loughborough University

April 2020

Abstract

This project aimed to create a tool in virtual reality that could be used to aid language learning, specifically to help vocabulary memorisation.

Lessons were drawn from research, existing tools and a questionnaire that polled the public for their opinions. These lessons examined how languages are learned, how memory works and how technology like virtual reality can aid learning. The project used these to explore what techniques and features maximise the quality of learning, and these in turn were used to define the requirements and design the functionality.

An application was developed for mobile VR where a set of words are learned through exploration and exposure to household items, being shown the foreign word and how it sounds alongside a realistic 3D model of that object. The application allows users to practice by reviewing and revisiting these words and related questions regularly, intelligently spacing questions apart to maximise the memory retention. The application can be used to challenge and test the user by modifying different settings and by showing the user statistics of their performance. The data that is gathered can be analysed and used for comparison and for measuring various variables.

The app is a successful tool that can be used to learn and practice new words in foreign languages, and the resultant usage data can be analysed for use to benefit the learner, the teacher, or to improve the field by improving the app itself or by being used as a research tool.

Acknowledgements

A lot of work has been done to write this report, and there are some major people who have helped to support this. My project supervisor, Firat Batmaz, has provided immeasurable wisdom and advice to this project, and it has been extremely helpful to have him to support this project. While researching, I reached out to several experts who helped guide me to some relevant subject areas, and these ideas really impacted where the project was headed – thank you to James Moran and my school's Psychology Teacher, Chris Asker. I also want to thank the Language Centre at Loughborough University for giving their expert opinions on the application. I also want to thank friends and family for supporting me and proof-reading my draft reports.

Table of Contents

List of Figures

List of Tables

List of Code Listings

Glossary of Terms

API	Application Program Interface
L1	First Language. A learner's native first language.
L2	Second Language. Languages that aren't the learner's native first language.
UI	User Interface
VR	Virtual Reality

Chapter 1 – Introduction

This section details the aims and objectives of this project and will also outline the report. Time constraints and other risks that may impact the report are also detailed and how they are mitigated is discussed.

1.1 Background

A second language is a difficult skill to learn, and an even trickier one to master. Every language is different, but there are elements of them that are always the same. Some of the aspects which are always important are vocabulary, grammar and pronunciation. Vocabulary is often one of the most time-consuming to learn. The amount of words to comfortably speak a language is often in the tens of thousands – and this is not counting every word in that language. There is established research into the effective learning of vocabulary – and additionally there are many established tools already that seek to help the learning of vocabulary.

Virtual Reality (VR) is a quickly growing industry, and much of its popularity has come from the video games industry. However, there is substantial potential for education to use virtual reality to enhance the learning process. While there are numerous ways in which virtual reality can be applied to enhance language learning, especially vocabulary learning, there is a lack of effective tools that explore language education using VR. The few existing tools have flaws and leave large opportunities for innovation.

1.2 Motivation

Language learning, memory and virtual reality have all been interests of mine for years. Finding a project that so neatly combines all three of these is a wonderful opportunity for me. Studying Psychology at school introduced me to the many different theories behind memory, and in the process of this project I hope to learn more. Using various VR systems, I've developed a great interest in the technology and its potential to impact many different subject areas and industry sectors. I've been a fan of language learning for years, first learning French and Russian at school, then Swedish in my spare time. Most recently, I have been learning Japanese, by attending weekly lessons and supplementing my learning using textbooks, mobile apps and websites.

This project can explore how virtual reality can be used to implement language learning, which could examine ways of improving the quality of language education, as well as innovating upon existing ways of using virtual reality.

1.3 Aims and Objectives

This project aims to create an application in virtual reality to assist with second language learning, particularly with vocabulary learning.

This aim generates the following project objectives:

1) To study existing research and case studies to determine:
    a. Whether aspects of the research have been done before
    b. If there is any recognised 'best practice' in the project area
    c. If there are any information gaps in the project area
    d. Whether any previous studies use methods that could be used in the project
2) To identify
    a. Any existing tools and features that can be analysed and used in the project
    b. Whether there is a public desire for this tool to be made
    c. What software to use and why
    d. What hardware to use and why
3) To create a language learning tool based in Virtual Reality
4) To evaluate the project to
    a. Judge how successful it was
    b. State what could be achieved given a longer timeframe
    c. Propose improvements on the existing work
    d. Ascertain what future work can be done to further the field

1.4 Risks

This section details any risks to the completion of the project and discusses how these issues were mitigated.

1.4.1 Time Constraints

There were a couple of time constraints which may have impacted the final deliverable. One of these was the Semester 1 Exam Period in January. A significant amount of time was devoted to studying and revising for these exams, so that I would perform well in them. I mitigated this by spending a few hours each day working on the project. There was also coursework for several modules that needed to be completed, and this was mitigated in a similar way, by splitting my time each day sensibly.

1.4.2 Working with New Technologies

In this project I worked with initially unfamiliar technologies – both hardware and software – that I had not previously worked with. This may have affected the speed and quality of my work, potentially affecting the final deliverable. I started learning how to use these technologies early in the project – before I needed to implement anything. I dedicated some time each day to developing my skills in order to mitigate the risk that this posed.

1.4.3 Bugs

This project involved a significant amount of software development. There was a risk that there could be errors that did not have an obvious cause or solution. These 'bugs' would have needed significant time and effort devoted to them in order to be solved. To mitigate this, when I encountered a bug that did not have a simple or obvious solution, I first turned to the official documentation to track down any discrepancy. Next, I searched online to see if anyone had faced a similar issue and looked at their solution. Then, I would have approached an expert on the technology for their opinion. These steps helped decrease any time spent on an issue.

1.4.4 Pandemic

Within the last few months of the project, from mid-March, the COVID-19 pandemic became very serious in the UK and caused a great deal of uncertainty and disruption. As my project is largely desk- and computer-based, this didn't cause a great deal of disruption to my workspace, however the period of uncertainty and confusion may have stalled the progress of the project, as well as distracting from the project.

There was a brief period when in-person university 'content' was cancelled and I spent this time focussing on my project, watching the online replacement lectures when taking a break.

After receiving communication from the university confirming that the project would continue, at which point the country was already in 'lockdown', I began spending each day working solely on the project and the report.

1.5 Research Methodology

This section outlines the different methods used in this project and discusses how they are applicable to this project.

1.5.1 Literature Review

A literature review is an examination and evaluation of all the available literature on a given subject. It details the newest research and summarises the information in the literature. It analyses the information, mainly by identifying gaps in the current knowledge. A literature review is intended to show an in-depth knowledge of the subject. (Royal Literary Fund, accessed 2019)

This is applicable to this report as all the current research done into VR and language learning can be analysed and examined. This provides a good basis for further work in the area, as well as fulfils Aim 1.

1.5.2 Questionnaires/Surveys

Questionnaires are a research methodology that can be used to gather large amounts of quantitative data in a cheap and quick way. It consists of a list of questions designed to gather information. The questions can be closed or open – closed questions structure the answer by only allowing responses which fit into pre-decided categories (such as yes/no or rating in a scale) whilst open questions allow for any answer, with space for as much detail as needed. Closed questions provide good quantitative data which can be analysed very easily, though they can lack detail. Open questions provide detailed qualitative data but can be harder to analyse without subjectivity. (McLeod, 2018)

A questionnaire can be used in order to gain information about the public's opinion on VR language learning tools, as well as data about what medium the public thinks is the most effective. To prevent gathering irrelevant information, designing the questionnaire carefully will be important. This can be done by not including too many open questions and mainly including closed, on-topic questions to gain relevant quantitative data. This will also fulfil Aim 2b.

1.6 Report Outline

The project has been broken down into the following chapters:

1) Chapter 1 – Introduction. This chapter outlines the project and its aims, objectives and risks. It also establishes the methodologies this project uses.

2) Chapter 2 – Requirements. This chapter draws together a list of requirements. It also reviews existing literature, hardware and software. A survey is also run to gain an insight into public opinion. Lessons are drawn which contribute to the requirements.

3) Chapter 3 – Design. This section details the design stage, and how each element of the final application is planned and designed. It also details and explains each design decision made and why.

4) Chapter 4 – Implementation. This section details the implementation stage and the issues faced while implementing a solution.

5) Chapter 5 – Testing. This section covers the testing of the application.

6) Chapter 6 – Conclusion and Evaluation. This section evaluates the results from the testing, discusses any limitations the application has, and how future work can build upon the development.

Chapter 2 – Requirements

This chapter completes a literature review which examines the extensive existing research in the main fields of the project. This chapter performs a gap analysis, reviewing existing hardware and software. A survey is also run to gain an insight into public opinion. This chapter draws together a list of requirements using lessons drawn from these.

2.1 Literature Review

This section reviews the extensive existing research in the main fields of the project, including language education and virtual reality. The analysis of my findings is used as evidence throughout the project to back up any decisions made. This also fulfils Aim 1.

2.1.1 Language Learning

This section explores different approaches and theories on how learners learn most effectively, including some theories specific to language learning.

2.1.1.1 Second Language Acquisition

There is a considerable body of work done into Second Language Acquisition (SLA), which is the process by which a learner would improve their performance in a second language, or L2. A third or fourth language (and so on) would also be called L2, as they are additional languages that are not the learner's native language. A learner's native language is often called L1.

Stephen Krashen breaks his theory of *Second Language Acquisition* (1983) into five hypotheses:

- The Acquisition-Learning distinction
- The Natural Order hypothesis
- The Monitor hypothesis
- The Input hypothesis
- The Affective Filter hypothesis

The Acquisition-Learning Distinction

The Acquisition-Learning hypothesis lays out a distinction between two independent systems of foreign language performance – the acquired system and the learned system. The acquired system or 'acquisition' is a subconscious process of 'picking up' a language. It is very similar to the process children undergo when they acquire their first language. Krashen says that acquisition requires more

meaningful interaction in the target language – more natural communication – in which the speaker is focussed on communicating, rather than the form of the utterances.

The learned system, or 'learning' is, in contrast, the conscious processes by which the learner learns about language and is the result of a more formal education. An example of this is learning about grammar rules explicitly (Schutz, 2019).

The Natural Order Hypothesis

The Natural Order hypothesis explains that "the acquisition of grammatical structures proceeds in a predictable order." Krashen goes on to explain that for any given language, learners will tend to learn certain grammatical structures early, and others later. He cites the studies of Dulay and Burt (1974, 1975) which reported that children acquiring English as a second language showed a 'natural order' for grammatical rules, regardless of their first language.

The Monitor Hypothesis

This hypothesis builds on the acquisition-learning distinction, arguing that 'formal rules, or conscious learning, play only a limited role in second language performance'. Instead, Krashen explains that the conscious rules (or the monitor) are used instead to influence the acquisition of the language. According to Krashen (1983), there are three conditions that must be met for these conscious rules to be useful when communicating:

- The speaker must have enough time to think about the rules
- The speaker must also be focussed on form
- The speaker must know the rule

Krashen points out that the conscious rules allow the speaker to use certain items that may not have been acquired. Therefore, use of the monitor can lead to more items that are "late-acquired" in the natural order.

The Input Hypothesis

This hypothesis attempts to explain how second language acquisition takes place. According to Krashen, the learner progresses along the 'natural order' when they receive a second language 'input' that is one step beyond their current stage of competence.

'Comprehensible input' is the language input that is one level above that of the learners' current level. The input hypothesis states that "a necessary (but not sufficient) condition to move from stage i to stage $i + 1$ is that the acquirer understand input that contains $i + 1$, where 'understand' means that the acquirer is focussed on the meaning and not on the form of the message." In other

words, when a student is faced with language that is a little further along the natural order than they are, they can still understand, and that is when they acquire.

The Affective Filter Hypothesis

This hypothesis breaks down how any affective factors will interact with the SLA process. According to Krashen, a 'variety of these affective variables relate to success in SLA', and he sorts these variables into three groups:

- Learners with high motivation generally do better
- Learners with self-confidence and a good self-image generally do better
- Learners with low anxiety generally do better

Summary

According to Krashen, picking up (or acquiring) a language is distinguishable from formally learning a language. Formal learning can inform and support a language's acquisition, but acquisition is how one improves their understanding of the language. Krashen's theory does imply that formal education is only of limited use in learning a second language.

In formal education, Krashen's theory implies that a majority of a learner's progress doesn't come through learning grammar rules, but rather through general, natural communication as part of the class. Especially if the class is taught in the second language and the learner is interested in the topic, the learner will acquire the language by being faced with that language input and trying to communicate using it back.

Krashen's theory also introduces the idea of a natural order of progression – the natural order in which a learner acquires various grammatical structures – and the idea of comprehensible input. Comprehensible input is the idea that if a learner is faced with language that is one level further along this 'natural order' than they are, the learner will still be able to understand, and this greatly improves their acquisition of the next level.

This theory does leave the question: does formal language teaching help? If learning only plays a limited role, how can education adapt to become more effective?

2.1.1.2 Language Teaching

Whilst it is beneficial to know how language can be learned more effectively, there has also been a substantial amount of research devoted to language teaching, and how to do this most effectively.

In “Approaches and methods in language teaching” (1986), Richards evaluates many different language teaching methods. He examines how these methods differ, and the key issues that each method is developed to solve. He generates certain questions which can be used to prompt development and innovation in language teaching:

“

1) What should the goals of language teaching be? Should a language course try to teach conversational proficiency, reading, translation, or some other skill?
2) What is the basic nature of language and how will this affect teaching method?
3) What are the principles for the selection of language content in language teaching?
4) What principles of organisation, sequencing, and presentation best facilitate learning?
5) What should the role of the native language be?
6) What processes do learners use in mastering a language, and can these be incorporated into a method?
7) What teaching techniques and activities work best and under what circumstances?

“ (Richards, 1986)

These questions can help to refine our teaching method to a more focussed point. Motivation in learning is also an important factor to consider, as also pointed out in Krashen’s affective filter hypothesis. Rustamovna (2019) asserts that the learner’s curiosity is the main driving force behind learning. She says that “the key is to let students focus on exploring an area which interests them and learn about it for themselves.” This shows that when teaching language, if there is something interesting, or something in which the learner can explore of their own volition, then there will be motivation, which will impact their learning.

When handling language teaching, there are many aspects to balance. As pointed out by Richards (1986), different methods have different goals: to teach reading, or to teach conversational proficiency, for example. Vocabulary is defined as ‘all the words that exist in a particular language or subject’ (Cambridge Dictionary, accessed 2019). Could teaching vocabulary be one of these possible goals, or is it more important than that? How important is it to teach vocabulary?

Vocabulary in Language Teaching

Norbert Schmitt is a prolific researcher on the topic of vocabulary. In a 2008 review article, Schmitt concludes that in order to function sufficiently in English, a learner must have a large vocabulary of around 8000 words, whilst a 20-year-old native English-speaker is expected to have a vocabulary of around 20,000 words.

Wilkins (1972) asserts that "while without grammar very little can be conveyed, without vocabulary nothing can be conveyed" and according to Lessard-Clouston in his book *Teaching Vocabulary* (2013), "without sufficient vocabulary, students cannot understand others or express their own ideas". Using English as an example foreign language, Lessard-Clouston continues by saying that as students develop greater fluency, it is important for them to learn a more productive vocabulary, as well as develop their own vocabulary learning strategies. As he noted, "teaching vocabulary helps students understand and communicate with others".

Schmitt (2002) says that words are not learned immediately, but rather more gradually over numerous exposures. According to Schmitt, there is a distinction between being able to understand a word (receptive knowledge) and being able to use a word (productive knowledge). Schmitt recalls an anecdote as an example:

"An example of knowing a word productively (at least in speaking mode) but not receptively in the written mode happened to me with a word connected with law. I had often heard and verbally used a word describing the formal charging of a criminal with a crime or offense. I never had the occasion to write this word, although I assumed from its pronunciation (/ɪnˈdʌɪt/) that the spelling was 'indite.' At the same time, I had occasionally seen the word indict. I did not know what it meant but assumed that it rhymed with predict. It was only later that I figured out that indict was the spelling for the word I had used for years to talk about law." Schmitt (2002, p. 5)

This shows that vocabulary is not as easy to learn as simply presenting the written or spoken word alone. There are more aspects to a word that must be considered. Nation (1990, p. 31) establishes a list of these aspects that is necessary for a learner to master and fully know a word. Schmitt (2000) cites this and calls them the types of word knowledge. The list is:

- The meaning of the word
- The written form of the word
- The spoken form of the word
- The grammatical behaviour of the word

- The collocations of the word (other words that co-occur more frequently with the word)
- The register of the word (how the word is used in different circumstances)
- The associations of the word
- The frequency of the word

Vocabulary is an important bedrock of a language, and in the case of English, a large magnitude of words must be known in order to be able to function well. It is not difficult to assume the same of other languages, to some degree.

2.1.1.3 Experiential Learning

Whilst looking specifically at language education is relevant and can help us, looking more broadly at theories of education in general may help show some useful techniques. Experiential learning theory (ELT) was proposed by Kolb in 2000. It defines learning as "the process whereby knowledge is created through the transformation of experience." The model uses experiences as the key idea in developing and learning new concepts.

Kolb et al. (2000) summarises the model succinctly. The ELT model has two ways of grasping and understanding new experience:

- Concrete Experience (CE)
- Abstract Conceptualisation (AC)

The model also has two ways of transforming this experience:

- Reflective Observation (RO)
- Active Experimentation (AE)

These ideas interact in a cycle, to fuel learning. This can be seen in Figure 1.

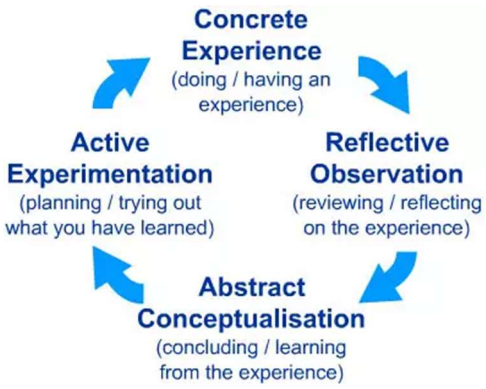


Figure 1. Kolb's Experiential Learning Theory (McLeod, 2017b)

The cycle in Figure 1 shows that concrete experiences are reflected upon. These reflections are merged into various conclusions or abstract concepts. These concepts can be used as the basis for active testing, which guide the learner towards new experiences. The key to the model is that in order to learn using experiences, the learner must reflect upon them.

In Tanaka et al. (2003), experiential language learning was used, and this led to a significant improvement in the participant's L2 proficiency. Here, the key 'experience' was 'studying abroad' and taking an English (as a second language) program whilst abroad.

This method of learning could be utilised within an environment such as virtual reality, to improve learning.

2.1.1.4 Problem-Based Learning

Problem-based learning (PBL) is an approach to learning that is structured around an open question. When implementing this approach, students work to learn what they need to know to solve a problem. Hmelo-Silver (2004) describes PBL as "focussed, experiential learning organised around the investigation, explanation and resolution of meaningful problems."

PBL is closely linked to experiential learning and follows a similar cycle, as displayed in Figure 2. After being presented with the problem, students analyse the problem and identify the relevant facts. This helps reframe the scenario and aids understanding, which leads to the students generating hypotheses on how to solve the problem. Then the students identify any gaps in their knowledge – this is an important step in PBL. These knowledge gaps become "learning issues" that the students research during their self-directed learning (SDL), according to Hmelo-Silver. After this SDL, the students apply their new knowledge and evaluate the hypotheses. After each problem is solved, "students reflect on the abstract knowledge gained".

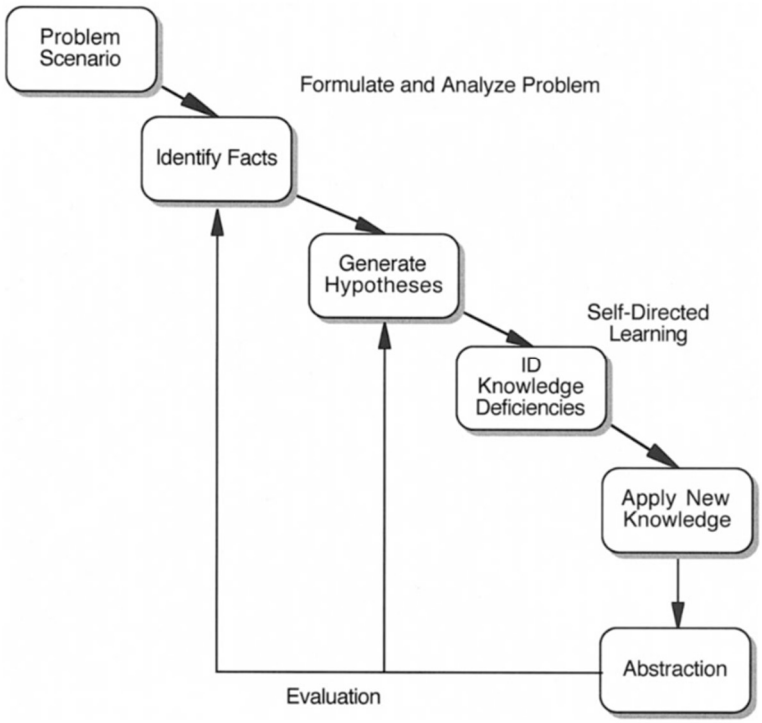


Figure 2. The Problem-based Learning cycle (Hmelo-Silver, 2004)

A teacher is more of a guide to the students, helping them through the cycle, as well as helping them learn the 'cognitive skills needed' to solve the problem. Students are generally self-directed and manage their own learning, they gain skills that can be applied broadly across any learning.

Adamidi et al. (2017) used a Computer-supported environment to teach a second language with PBL. The results suggested that participation in the tasks developed language learning skills to a great extent.

There is also often a collaborative element to PBL and working as a team to solve the problems.

2.1.1.5 Discussion

There is a significant amount of research into the effective teaching and learning of languages. Krashen's theory of SLA is important as it helps explain the relationship between education and acquiring a language. A key element to take away is the idea of comprehensible input – that if a learner is challenged with any language input that is one level further along the language's 'natural order' than they are, the learner will still be able to understand and acquire that language input.

When it comes to language teaching, different methods have different goals – and when developing a new method, it's important to know the goals and build the teaching method around them. Research also shows that a learner's motivation and curiosity play important roles in how well the learner performs.

Vocabulary is a very important part of language and there are multiple aspects to each word – it isn't as simple as knowing the translation. There are different elements to each word – its pronunciation, spelling and use in different contexts are a few key ones. There is a distinct difference between being able to understand a word and being able to use a word in a natural way. Due to this added complexity, research has shown that it takes numerous exposures to a word in order to learn and 'master' it.

Experiential learning (EL) is a method where students learn using experiences by reflecting upon them and using what happened in one experience to inform them in the next experience. This also can involve the students using previous experiences to test themselves and experiment as a method of learning.

Problem-based learning (PBL) is a specific type of experiential learning, which involves posing an open problem to students and letting them have control and find the solution themselves. They hypothesise based on what they know, identify what they don't yet know, explore the subject area in order to learn, in order to test their hypotheses and find the solution.

As this project will be developing a method through which students will be learning a second language, it is important to take these lessons into account –the ideas posed by Richards will need to be reflected upon, especially with regards to what goals this project will be working towards. The proposed learning methods will also need to be used in such a way to provide an effective teaching and learning method. This will likely utilise a combination of the explored methods, as there are numerous ways in which concepts like comprehensible input and problem-based input can be combined.

This research thoroughly explores theories on how learning and education can be done effectively, as well as theories on how a learner can improve at a language. Though they do leave some questions, for example, once an element of a language, like a word, has been learned, how do we make sure that it is not forgotten?

2.1.2 Vocabulary Memorisation

This section looks more closely at how vocabulary can be memorised and maintained over long periods of time. It examines several established vocabulary memorisation techniques, including several theories of memory and how they can be applied to vocabulary memorisation.

2.1.2.1 Repetition

The Multi-Store Model of Memory (MSMM) by Atkinson & Shiffrin (1968) proposes that memory can be broken down into a structure of three stores:

- Sensory Register (SR)
- Short-Term Memory (STM)
- Long-Term Memory (LTM) (McLeod, 2017a)

Information from the senses is stored in the SR, but it's only stored for a fraction of second before being forgotten. If attention is paid to any piece of sensory information, it will move into STM. The STM is said to only have a capacity for 5-9 items, and a duration of around 30 seconds. (Turner, accessed 2019)

If information in the STM is rehearsed, it moves to the LTM. If this 'maintenance rehearsal' does not occur in the STM, the information is forgotten. Maintenance rehearsal that is mentioned is essentially repetition.

The flow of information through the MSMM can be described with the flowchart in Figure 3.

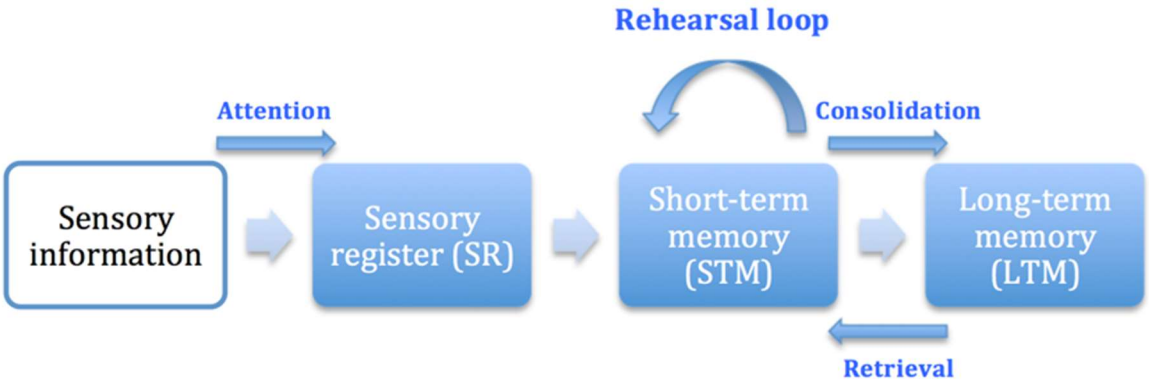


Figure 3. Multi-store Model of Memory (Turner, accessed 2019)

Atkinson and Shiffrin (1968) also cite evidence from Hebb (1961) and Melton (1963), saying that they show that rote repetition is good for the LTM (Atkinson and Shiffrin, 1968).

Glanzer & Cunitz (1966) builds on the MSMM and shows that repetition is a viable way to remember words. They presented participants with a list of words and asked them to learn them. The participants were split into two groups – one recalled the words immediately after learning, and the other group was asked to recall after spending 30 seconds doing a distracting task.

They found that words that were presented early in the list were consistently remembered well. Glanzer and Cunitz explained this was an example of the primacy effect – that these words were put into long term memory because the participant had time to rehearse each word.

They found that words towards the end of the list were also remembered well when recalled immediately – an example of the recency effect. These words were still in the participants' STM when asked to recall. Participants who had taken part in the distraction task could not remember these words well, as it had been too long for the words to remain in the STM.

Glanzer and Cunitz demonstrates that rehearsal is a viable method for entering information into the LTM, with the example of the primacy effect. When given the opportunity to rehearse words, they will enter the LTM. (McLeod, 2008)

2.1.2.2 Elaboration

Another theory of memory is the Levels of Processing model by Craik & Lockhart (1972). This model focusses on the depth of processing, arguing that the deeper information is processed, the longer the memory will be retained (McLeod, 2007).

Depth, in this case, is defined by Craik and Lockhart as:

"the meaningfulness extracted from the stimulus rather than in terms of the number of analyses performed upon it" (1972, p.48)

This means that a student will remember information better if they have processed it more deeply, or in other words, if the student has extracted more meaning from the stimulus. This is an unstructured model and doesn't focus on stores or structures like the STM or LTM and instead concentrates on the processes involving memory.

Shallow processing can take two forms – structural and phonemic processing. Structural processing is when a student encodes information using the physical characteristics of an item, whilst phonemic processing is when a student encodes using the sound of an item. Shallow processing involves maintenance rehearsal and results in quite a short-term retention.

Deep processing involves semantic processing, which occurs when encoding uses the meaning of a word and relating it to similar words with similar meanings. If the information a student wants to

remember has multiple links to their existing memories / knowledge, then it will be easier to remember it for longer. Deep processing involves elaboration rehearsal which involves more meaningful analysis and results in more long-term retention (McLeod, 2007).

Keyword Method

Raugh et al. (1977) investigated the effectiveness of the 'mnemonic keyword method', when teaching a Russian vocabulary to English-speaking students. This keyword method "divides the study of a vocabulary item into two stages. The first stage requires the student to associate the spoken Russian word with an English word (the keyword) that sounds like some part of the foreign word; the second stage requires the student to form a mental image of the keyword 'interacting' with the English translation." An example of this is the Spanish word for letter – carte. Our keyword would be cart (like a minecart), and our mental image could be a cart full of letters.

This keyword method draws multiple links between the foreign word and the English translation using both a similarly sounding keyword and a mental image. This combination involves deeper processing than simple rote repetition. This method uses elaboration to be a more effective tool.

In Raugh's study, the findings indicated that this keyword method was highly effective. These findings show that drawing links between the foreign word and processing the information at a deeper level is an effective strategy. It also demonstrates that relating the new word to our existing knowledge, as well as generating mental images, are useful skills.

2.1.2.3 Retrieval Practice

Weinstein et al. (2010) ran an experiment to compare three different study strategies – re-reading, answering questions, and generating questions. The study found that answering and generating questions gave better performance than re-reading. Weinstein concludes that actively recalling information is more effective than merely reviewing it.

When this principle is applied to vocabulary learning strategies, it is possible to look at the native word and actively recall the foreign word, or vice versa, rather than just read through a bilingual list of words.

2.1.2.4 Flashcards

Flashcards are a common way to implement retrieval practice. When flashcards are used to study and memorise, the learner often 'drops' (put aside and stop studying) items that they think that they know. Kornell et al. (2008) asserts that there is a convincing logic to it – by dropping items that seem well-learned, it "creates more opportunities for the remaining items to be studied." The study investigates whether dropping flashcards is effective.

The study finds that dropping flashcards had a small but consistently negative effect on learning. This is because flashcards that are well-learned can still be forgotten and dropped cards should still be returned to later.

2.1.2.5 Spaced Repetition

The spacing effect, according to Kornell (2009), is the idea that 'spacing learning events apart results in more long-term learning than massing them together'. This doesn't take any more time or effort on the student's part and relies on a more effective time distribution.

It is often used alongside flashcards. Kornell et al. (2008) makes the additional point that returning to dropped flashcards leads to more opportunities for the spacing effect, and that the more cards are dropped, the spacing between remaining cards decreases.

In 2009, Kornell published another study investigating the spacing effect, including what should be studied in a session: using the spacing or massing methods. The spacing condition was repeatedly returning to a large stack of flashcards, whilst massing was splitting flashcards into smaller stacks and studying them one stack at a time. This study found that spacing a large stack of flashcards over multiple days had a large positive effect on memory.

Schimanke (2018) says that a goal of spaced repetition is to look for the 'longest possible intervals that do not lead to forgetting'. This can be seen in Figure 4, where reminders are given when the chance of remembering is projected to fall below 90%. It's also worth noting that the rate of decrease of this value slows after each reminder due to the spacing effect.

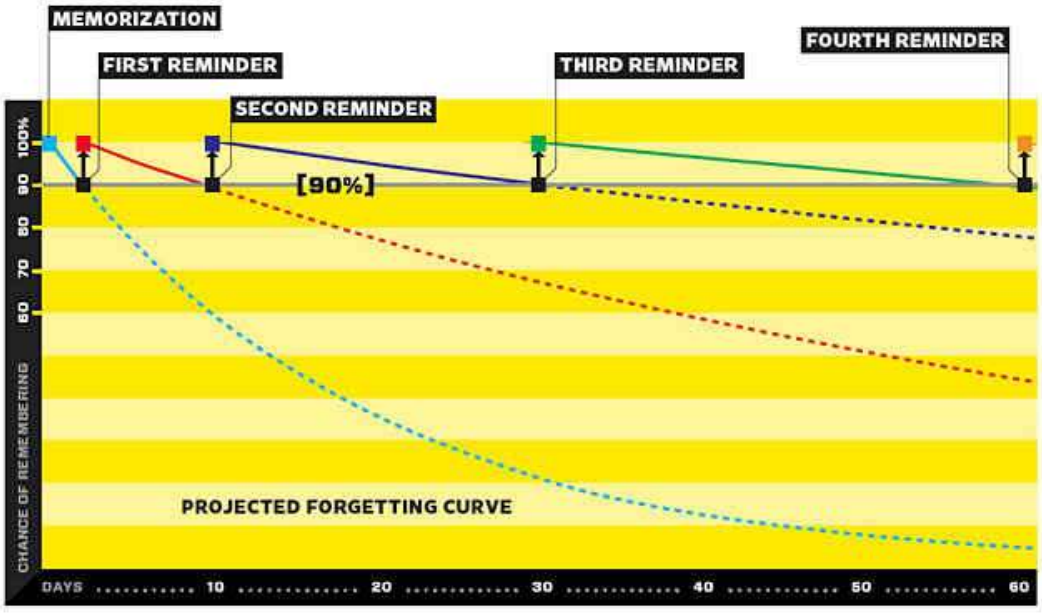


Figure 4. Spaced Repetition (Gwern.net, 2019)

Leitner System

The Leitner system is a spaced repetition algorithm that can be used alongside flashcards. This system is said to be used by 'several electronic flashcard programs'. The algorithm is summarised by Settles et al. (2016):

"The main idea is to have a few boxes that correspond to different practice intervals: 1-day, 2-day, 4-day, and so on. All cards start out in the 1-day box, and if the student can remember an item after one day, it gets "promoted" to the 2-day box. Two days later, if she remembers it again, it gets promoted to the 4-day box, etc. Conversely, if she is incorrect, the card gets "demoted" to a shorter interval box."

This algorithm can be demonstrated in Figure 5.

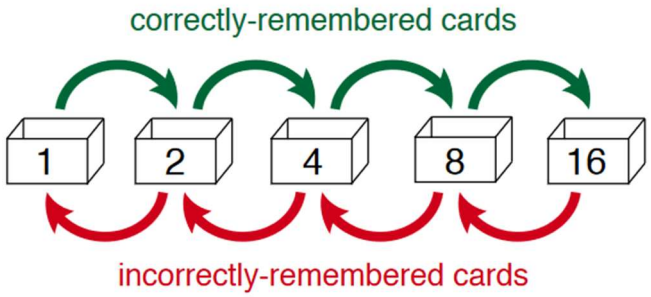


Figure 5. Leitner system for flashcards (Settles et al. 2016)

2.1.2.6 Discussion

When it comes to memory and different techniques which can be used to remember vocabulary, there is a wealth of research and occasional contradictions. That said, each different theory has supporting evidence, and so there is likely a much more complex explanation for how memory works that hasn't been fully discovered yet.

That said, these methods have been shown to work. Repetition is the baseline for memorisation. An instinctual method for many, when a student wants to memorise information, they tend to repeat that information over and over, and eventually they rehearse this information enough that they can recall the information at a later time. This method does work. However, even though it is the intuitive method, it has been shown many times that it is not the best, particularly when it comes to memory retention. Within its body of research, it does establish that only a certain number of items can be learned effectively at a time (5-9), by the short-term memory.

A rival theory of memory suggests elaboration rehearsal as a superior method – by processing the information more deeply, it can be remembered for longer. By processing the information deeper, a

student can attach more meaning to the information. In other words, they create more links between our existing knowledge and the new information. This can be by imagining a story that links the information, or attaching words, images, sounds, or concepts to the information. This helps to craft a better-quality memory, which will be retained for longer.

Once memories have been learned, it's necessary to keep them maintained, else they will be forgotten. There are different ideas that revolve around this as well, and the best methods in which to retain information once it has been learned.

Retrieval practice demonstrates that by actively recalling the information and 'retrieving' it out of their memory, a student can improve the quality of a memory. A useful tool for retrieval is flashcards, and it is possible to also utilise the spacing effect to improve the quality of memorisation further. This can be done by practicing retrieving information at varying intervals, which keeps the information 'fresh', and much less likely to be forgotten. This is shown to be an effective method for learning as well, as the regular revisions help to solidify the memory.

These methods can be used within the project to encourage the user to effectively learn and then remember the words for a long amount of time. It's also important to help the user attach more meaning to each word by presenting not only the item and its translation of the item, but the spelling, sound, and mental image of the item as well. Other interactions might be able to show the use of the word within a sentence.

This research has explored extensively various methods which can be used for memorising vocabulary and builds on our existing knowledge of how students learn languages. Much of the research so far has made no mention of computers and how they can be used to contribute towards a student's learning. Seeing as computers are such a multi-purpose and widespread technology, can they be used effectively alongside language learning?

2.1.3 Computer Assisted Language Learning

This section reviews existing research into how computers can be used as a tool within language learning.

According to Levy (1997), computer assisted language learning (CALL) is an inter-disciplinary field that has been influenced by many other disciplines. Computer science and language teaching/learning are expected fields to be impacted, but there is also psychology, artificial intelligence, human-computer interaction and linguistics, among others.

Levy et al. (2008) explains that the diversity of CALL as a discipline is thanks to many factors, but primarily due to the "range of technological tools available with the potential for use"; "an increasingly sophisticated understanding of how languages are learned"; "environmental factors" which lead to a variety of needs for different learners based on their situation; and "particular challenges that arise" thanks to the characteristics of the target language.

According to Levy et al. (2008), a considerable amount of the technological advancements in the last 30 years have some relation with communication and language. Levy says that language teachers have more and more options when it comes to technology that they can use. He also says that technologies that can be used go further than merely communication tools though – many generic tools are also useful – a word processor for writing, for example. Levy says that many implementations of CALL lead to language-learning websites. The websites can differ hugely in terms of scale and focus, which speaks to how flexible websites can be as a teaching medium.

2.1.3.1 Gamification

Gamification is when game design elements and various motivation principles are applied in non-game situations. This is usually done to help provide motivation to users, which helps to keep the users engaged with the system for longer periods and keep them motivated to return to the system on a more regular basis. (Osipov et al, 2015)

Rustamovna (2019) mentions that using games can lead to students learning without realising. She also says that gamification is an effective method to motivate learners. Motivation, as has been pointed out in Section 2.1.1.2, is an effective way to improve learning.

Gamification is a method that can support user engagement, increase user activity, or quality and productivity of actions, according to Hamari et al. (2016). These patterns are thought to occur due to "positive, intrinsically motivating 'gameful' experiences."

Galoyan et al. (2017) agrees with this research, saying that games are beneficial for "motivation, engagement and a mastery of content-related knowledge and skills." The study examines how game-based learning can be incorporated into a language course. Galoyan further examines the Game Network Analysis (GaNA) framework, that can allow for game-based learning to be implemented into a language context.

The GaNA is a combination of various frameworks designed for pedagogy and analysis. When adopting this framework, the teacher first picks a game and begins analysing it, examining it for any good elements for teaching, content or technology. Then, the teacher will design a curriculum for a unit or course by designing experiences for play (P), curricular activities (Ca), and opportunities or reflection (R) and discussion (D). These make up the PCaRD framework. The opportunities for reflection and discussion are like the reflection needed to learn from an experience in experiential learning.

These built-in experiences allow for students to inquire (I) further, communicate (C) with teachers and peers, construct (C) models to demonstrate understanding and express (E) themselves. This makes up the ICCE framework. These two frameworks together are integral to the GaNA framework.

Galoyan also provides an example lesson plan, which may be useful to reflect on when designing a 'lesson' within my application. This lesson plan can be found in Table 1.

Table 1. – Sample Lesson Plan using GaNA framework (Galoyan et al., 2017)

Sample Lesson Plan **Monopoly Game-Based Language Learning**
General Information • Level of English: Intermediate • Age: High-School • Course: English for General Purposes • Duration: 1 hour 35 min • Game Used: Monopoly (online version) • Link to the game: http://www.pogo.com/games/monopoly
Goals • Teach business-related vocabulary in English • Engage students in interactive curricular activities based on PCaRD and ICCE

- Help students practice and acquire 21st century knowledge and skills, such as problem solving, critical thinking, collaboration, communication, strategic thinking, and effective negotiation
- Enhance in-class communication and interaction in English

Learning Outcomes

- Use business-related English vocabulary in meaningful conversations
- Practice problem solving, critical thinking, collaboration skills, strategic thinking, and negotiation skills to communicate with peers and play the game effectively
- Make relevant connections between the game, classroom activities, and their lives

Anticipated Problems

- Some students might not be familiar with the game and need additional support from their peers and the teacher
- Some technical problems might occur while playing the game online

Activity	**Procedure**	**Objectives**	**Time**
Introduction	• Introduce the topic of the lesson and the game • Distribute handouts with the rules of the game and discuss them with the students	• To scaffold the learners into the following curricular activities	10 min
Vocabulary teaching	• Pre-teach some essential vocabulary from Monopoly • *Vocabulary from Monopoly* • *Monopoly* • *Property* • *Mortgage* • *Bankruptcy* • *Liability* • *Own, owner, ownership*	• To activate the learners' schema • To provide them with the necessary vocabulary for in-game communication and interaction and follow-up reflection and discussion activities	10 min

	• *Auction off* • *Income tax* • *Luxury tax* • *Jail* • *Real estate* • *Community*		
Play (**P**CaRD)	• Provide each group with a computer or ask them to use their mobile phones • Give the necessary instructions on how to access the game online • Get the students to play the game in groups of three or four(alternatively, they can play with a virtual opponent)	• To make students play the game collaboratively to engage them in **Inquiry** and **communication** (**I**CCE) • To engage them in **discussions** (PCaR**D**) where they can practice the target language and the game strategies	60 min
Curricular Activity and Reflection (P**CaR**D) **Creating and presenting a poster**	• Split the students into groups of 3-4 • Provide the students with poster paper and markers • Give instructions on how to create and present a poster sharing their experience with the game	• To provide an opportunity for **construction** and **expression** (IC**CE**) • To get the students to actively **reflect** (PCa**R**D) on their experiences with the game and make them make meaningful connections between the game play, curricular	15 min

		activities, and their personal lives	

2.1.3.2 Applied Behaviour Analysis

Applied Behaviour Analysis (ABA) is a field broadly relating to the use of data analysis to affect behaviour. It involves analysing behavioural data and applying that to inform and explore changes of behaviour.

ABA "offers a foundation for the […] application of environmental manipulations", according to Cooper (1982), and can be used within education to investigate many areas and achieve many goals. Cooper describes how ABA can be used to affect behaviour in a positive way, such as reinforcing positive academic or social behaviours. ABA can also be used in another way to gain more information about behaviours and contribute new research.

ABA can also be applied to "assessing the academic needs of students [and] evaluating individual students' responses" (Ardoin et al, 2016). This provides some areas that can be analysed, and then that analysis can be used to inform and improve the effectiveness of the teaching. It is clear that ABA can be applied to many more varied areas, and in each case, the analysis can be used to inform and improve the area.

In his book, Turner (2014) talks about using statistics in language education – whilst not specifically in the field of ABA, this is related to this project because it involves the application of data analysis to improve language education. He says that an increased number of language educators are "investigating some aspect of their own practice to gain a better understanding of and improve both learning and teaching".

2.1.3.3 Discussion

There is a good amount of research into the effectiveness of utilising computers in language learning, especially within the discipline of computer-assisted language learning. It is interesting to note that Levy et al. said that many implementations of this come in the form of websites, and that the variety of these websites attests to the flexibility of websites as a teaching medium. Websites are an effective method for teaching, as the internet is widely available, and the sheer number of technologies that can be used alongside them allow the number of possible implementations to be massive.

Gamification is an area with a wealth of information published about it. Gamification is when game elements are incorporated into learning, and this hugely increases the motivation, user activity and prolonged engagement of students. It also enhances the learning, leading to a higher performance.

The GaNA framework is also proposed which can help prepare a gamified lesson plan, which may be of use when designing and developing the learning strategies used within my project's application. The key idea behind the GaNA framework is that through playing a game in another language, that language is learned.

Applied Behaviour Analysis (ABA) can also make use of computers – the data that is gathered by computers can be analysed and used to improve the quality of education and digital tools. This data can be used towards reinforcing positive behaviours, assessing the needs of students and evaluating individual responses, and many more. ABA can also be used to gain a better understanding of language education, and this can be used to improve the quality of that education.

If the way that the computer is interacted with is drastically changed, for example if VR is used rather than a keyboard and mouse, the benefits gained from gamification and the ABA will still apply and can still be used.

2.1.4 Virtual Reality

This section investigates the relationship between virtual reality and language education, also examining the link between virtual reality and memory.

2.1.4.1 VR Adoption

Virtual Reality has seen a vast advancement in recent years, both in the development of its technology and the industry has improved as well, with a growing market. The market has grown from 2.2 billion dollars worldwide in 2017 to 6.2 billion dollars worldwide in 2019. It is also projected to grow to 16.3 billion dollars by the year 2022, as seen in Figure 6. (SuperData Research, 2019) The asterisks indicate projected statistics.

This demonstrates that virtual reality is growing in popularity and as more units and software are sold, the technology is becoming more available and is gaining more attention – which will hopefully lead to more and more applications taking advantage of the benefits of virtual reality.

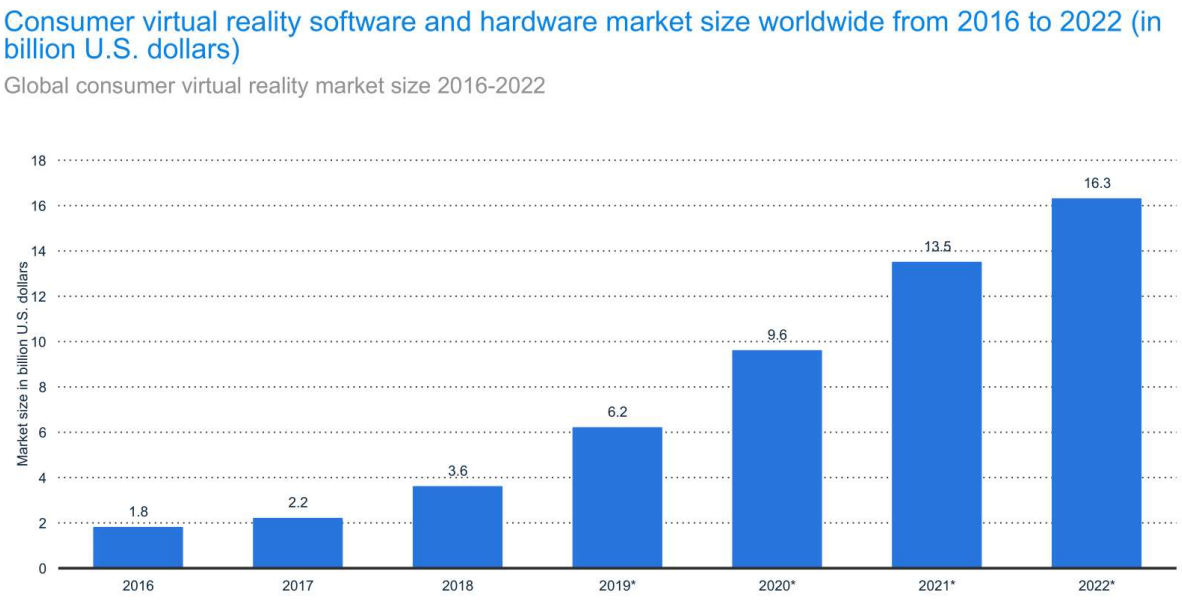


Figure 6. Forecast Virtual Reality Market Size 2016-2022 (SuperData Research, 2019)

There is also a predicted growth in the virtual reality software market specifically, growing from 1.9 billion dollars worldwide in 2019 to 6.4 billion dollars by the year 2022, as seen in Figure 7. This also shows that there is a predicted growth in the amount of software made for virtual reality.

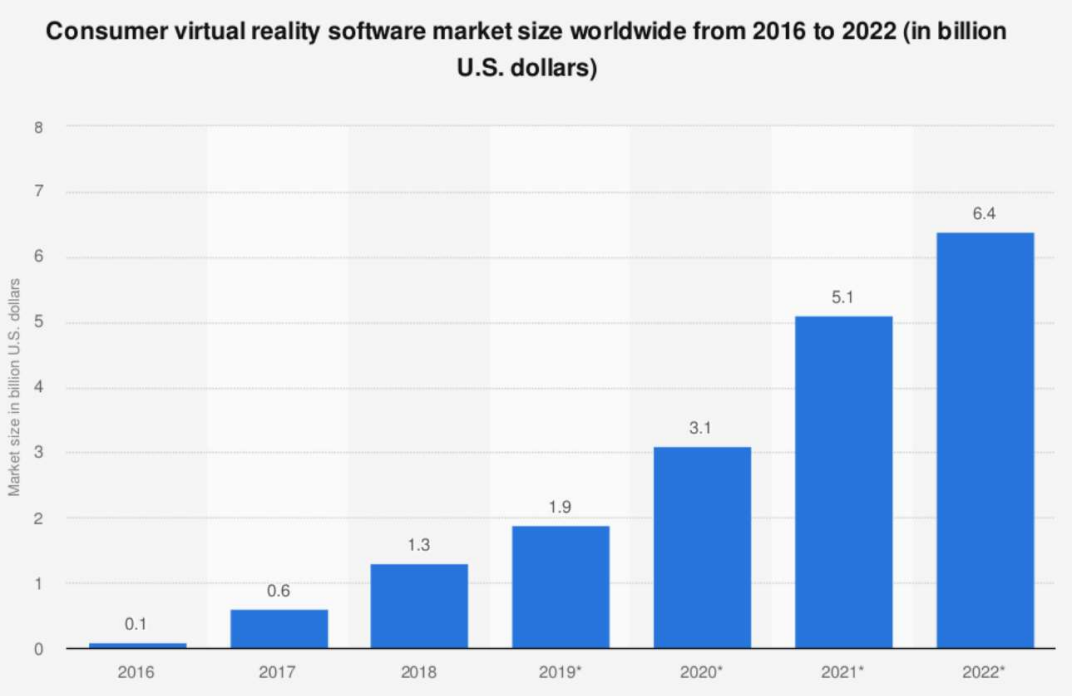


Figure 7. VR Software market size forecast 2016-2022 (SuperData Research; Statista estimates, 2019)

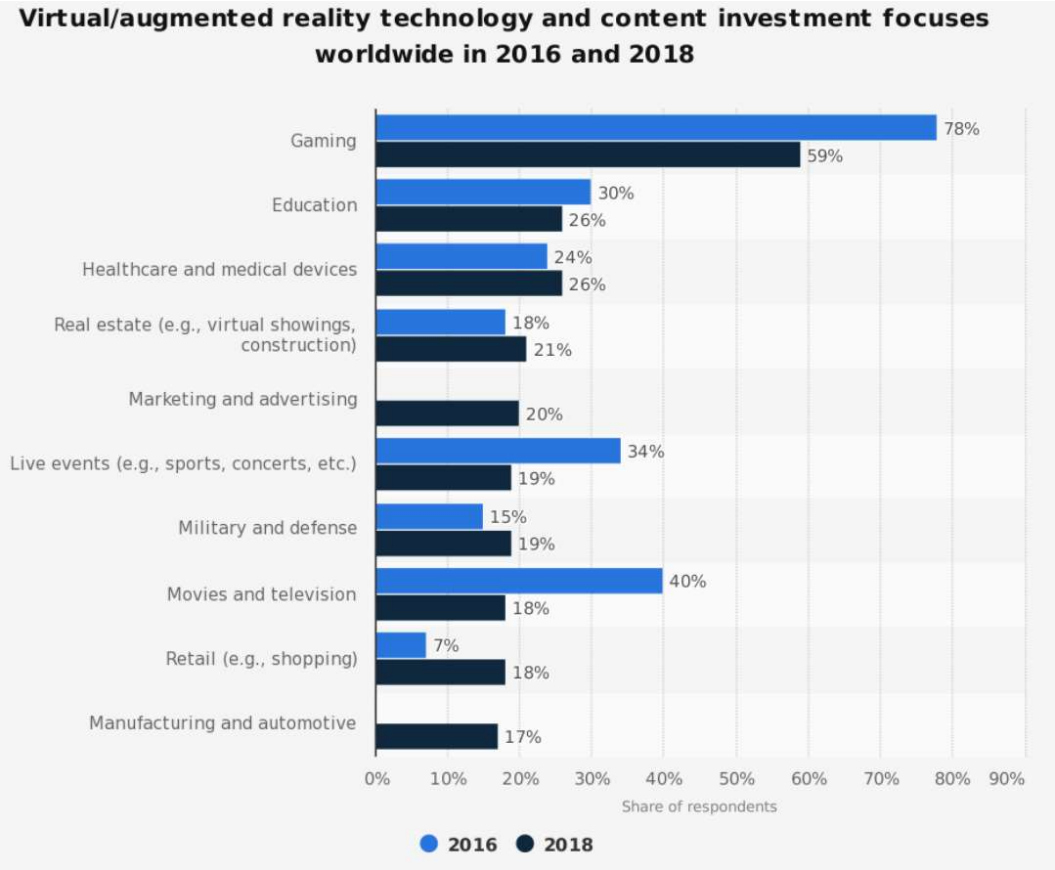


Figure 8. VR/AR investment focuses in 2016 and 2018 (Perkins Coie, 2018)

Whilst virtual reality is growing, education is not a key focus, and is only receiving partial focus. As Figure 8 shows, education was a focus for only 30% of investments in 2016, and only 26% in 2018. It isn't the biggest focus, however it is receiving attention and with the development and research into new applications of VR within education, this project can still make an impact, as well as encourage the growth of the education sector within the virtual reality industry.

2.1.4.2 VR and Memory

Virtual Reality inherently is a new and different medium in order to convey information. There are some different ways in which it can be interacted with, which lead to some interesting effects on memory.

Immersion

According to Krokos et al. (2019), 'virtual reality headsets, such as head-mounted displays (HMD), afford us a superior spatial awareness'. This study also looks at memory palaces and how virtual reality can implement a virtual memory palace. Memory palaces (also known as Method of Loci) are essentially a method to spatially organise thoughts and concepts, where one can later navigate an imagined structure to recall information. Krokos et al. investigates the link between immersion in virtual memory palaces and our ability to recall information.

Krokos compares information given via an HMD to that of a two-dimensional desktop display with a mouse-based interaction. The information given via HMD mimics a memory palace – a spatial representation of information. Within the scene, famous faces were placed in certain locations around the user – for example, Oprah Winfrey was standing at the top of a staircase. After a two-minute interval, the scene reappears with numbered boxes for where the faces were, and the participants were asked to recall which face was in each location.

The results showed that using the HMD increased recall accuracy significantly. This shows that using a virtual memory palace can improve recall. For instance, if there had been words instead of faces, or items with words attached, this could be a method to enhance vocabulary learning.

Presence

The concept of presence is defined by Lee (2004) as "a psychological state in which virtual objects are experienced as actual objects in either sensory or non-sensory ways." This is an important concept that is used to investigate virtual experiences, such as virtual reality. It's important to note that everyone will have a different level of presence, and it isn't an assumed constant that virtual reality induces high amounts of presence.

Bailey et al (2012) looks at how presence impacts memory. After showing participants an informational pro-environmental message within an immersive virtual environment, they were asked to complete a memory task about pro-environmental principles, outside of the virtual world. The participants also self-reported on their presence. The study found that as levels of presence increased, the participants recalled fewer correct examples.

One of the reasons that Bailey explains this result is through mediated arousal. According to Bailey, virtual environments can provide high sensory experiences – which can induce arousal. Excessive levels of arousal can hinder memory, and it is possible that the virtual environment in Bailey's study "enhanced a feeling of presence but may have hit a threshold of arousal that hindered memory."

Cho (2018) investigates how VR affects memory, as well as presence and motivation in language learning. Cho implemented a fully immersive environment, using several factors:

- The simulation should be interactive
- The simulation must be familiar
- The first-person perspective must be used to increase immersion levels

Cho found that immersive VR allows for a higher spatial presence than a desktop equivalent. The study also found that a significant increase of presence has a positive correlation with memory retention – directly contradicting the finding of Bailey. Cho does acknowledge that in certain conditions there was no significant correlation and that the relationship between presence and memory is an unpredictable one. Cho also cites several other studies, such as Mania & Chalmers (2001), which revealed a negatively correlated relationship.

Mania and Chalmers (2001) concluded that presence is not always a significant factor, and that other, more individual differences (such as participants' ability to remember certain types of information) are associated more strongly with memory.

2.1.4.3 VR in Education

Makransky et al (2018) ran a study to see the differences between using desktop (2D) or immersive (3D) VR in a virtual science learning simulation. He found that the immersive VR version (implemented using head-mounted displays) obtained higher scores and identified some of the largest differences occurring with regard to presence and motivation. According to Makransky, motivation is "one of the most important potential benefits of using simulations in education". Seen as there is existing research supporting the fact that using 2D simulations in education already has

significant motivational benefit, Makransky's result that immersive simulations have significantly more motivational value is promising for the use of VR.

During the recent COVID-19 pandemic, widespread international lockdown led to many classes moving online. One viral video showed a full mathematics lesson being held in VR. Teacher Charles Coomber (2020) posted a video of his lesson, recorded entirely in VR video game "Half Life: Alyx". This game allows for very fine control, and Coomber used whiteboard markers that were already in the game to write and draw diagrams on windows within the game. This shows that when it comes to education, Virtual Reality can be an effective natural choice as a platform.

2.1.4.4 VR in Language Education

There are several studies which explore how virtual reality can be used in language education. Many of them conclude that using virtual reality has a strongly positive effect on the students' academic performance and motivation (Chung, 2012; Solak et al, 2015), whilst others use VR to affirm the effectiveness of gamification (Connolly et al., 2011; Lan et al., 2018).

2.1.4.5 Discussion

When introducing virtual reality, specifically using immersive 3D head-mounted-headsets, other factors are also introduced which impact learning and memory. Immersion is hugely increased within virtual reality, which has been found to also increase memory recall. Memory palaces are mentioned as a similar and relevant comparison, as they are immersive virtual environments.

Presence is when virtual environments begin to feel as if they are real, when they are not. There were contradicting findings here – some say that increased presence improves memory retention, while others showed a worse memory performance with increased presence. There are multiple reasons suggested for this, such as the presences reaching a limit with regards to emotional arousal, after which it starts to negatively impact memory. Whilst this is possible, other studies have concluded that presence isn't always a significant or important factor in memory, suggesting that the way that presence interacts with memory isn't completely understood.

When a virtual environment is compared through either a desktop or a VR headset, VR has led to a better performance, as well as increased motivation. Motivation is an important element to learning, and VR has been shown to lead to more motivation than a computer equivalent. Seen as a computer equivalent already has a significant amount of motivation when compared to other teaching mediums, VR is a promising medium.

2.1.5 Conclusion

This section has extensively explored factors which improve the quality of education, language teaching and memory, and has explored how different technologies affect those factors.

This literature review shows that there is a distinction between formally being taught a language and acquiring the language. Formal education informs and improves acquisition. It is also seen that by challenging learners with language that is one level higher than they are, the learners can still understand, and this is an effective method for acquiring that level.

Many sources agree that curiosity and motivation are key to learning and help improve performance. Motivation especially is an aspect that is said to be greatly increased when using gamification and virtual reality. This suggests that when using these, the quality of learning is increased.

Experiential learning is a learning method where students have experiences, review them, and then experiment and test themselves by trying a similar experience. This is shown to be effective alongside language education, leading to significant improvements in language proficiency. Problem-based learning is an extension of experiential learning, where students are given a problem and are then expected to review and analyse that problem, formulate a hypothesis and then test that hypothesis. This has also been used to greatly improve language skills. If these learning methodologies were effectively used within an immersive and highly motivated environment such as VR, language skills could be greatly improved.

This section shows that actively recalling information and using it provides a more effective memory retention aid than simply reviewing that information. It also demonstrates that there is a distinction between being able to understand a word and to use it. When using strategies that make use of retrieval practice, this will involve using a word and being able to understand it, enabling thorough exposure to each word.

It takes numerous exposures in order to properly learn each word, due to each word having many different aspects to them – the sound, the spelling, the concept etc. These numerous exposures that each word needs could take advantage of the spacing effect, which would also improve the memory retention of the word over time. Spaced repetition can be used to help learn the word through these numerous exposures as well as to improve the memory retention.

As well as spaced repetition, elaboration is an effective way to improve memory retention, by adding extra meaning to the word, in the form of associating the sound, the spelling or a mental image of the word with it. This also helps to expose the learner to multiple aspects of the word and

so more thoroughly learn it. Using VR has been shown to act effectively as an immersive virtual memory palace to improve memory – this can work alongside elaboration as VR models can act as mental images.

Gamification is known to increase learners' motivation, engagement and user activity, all of which improve the quality of learning. This is also used alongside language learning to improve language skills effectively. VR also increases factors that improve the quality of learning, such as immersion, engagement and motivation.

Applied Behaviour Analysis can be used to reinforce positive behaviours, as well as to assess and evaluate individuals. By analysing data, it is possible to gain a better understanding of the learning process and improve the quality of teaching. This can be used to help develop a new teaching method. Richards also shows that it is important to focus on the goal of the teaching method in order to make an effective method.

The findings of this literature review will help inform and improve the quality of the design and implementation of the application.

2.2 Gap Analysis

2.2.1 Existing Tools

2.2.1.1 Vocabulary Tools

This section describes existing vocabulary learning tools and draws lessons that can be used. These were all tested and tried out using a mobile phone.

To properly compare and judge each existing tool, there are a number of criteria that will be checked. These will be given scores, and the reasons will be detailed. These criteria will be:

- Usability. How easy it is to intuitively know how the tool functions. How straightforward it is to use.
- Topics. How broad or well-thought-out the topic range and the language content covered is.
- User Control. How much control the user has with regards to what they're learning and how.
- Customisation Options. How much the user can change and customise to their own preferences.
- Effective Learning Tool. How effective it is for learning and remembering new languages. This will be mainly based on the techniques and algorithms used.

Anki

Anki is a program that is designed for learning vocabulary with flashcards, using the spaced repetition method. It also uses active recall testing, which is very similar to retrieval practice – it means being asked a question and trying to remember the answer.

When using Anki, the user works with 'decks' of flashcards. The user can craft these decks themselves or download a multitude of decks from online sources. When the user select's a deck to practice, Anki will show a flashcard to the user. The user then should think of the answer (the word on the 'other side' of the card). For example, in vocabulary learning, this would be a foreign word and its translation. The user then clicks a 'Show answer' button to reveal the other side. After reading this, the user is prompted to answer how they performed on the flashcard – choosing from 'bad', 'good' and 'great'.

This answer determines how soon that flashcard will be shown again to you. If the user chooses 'bad', the flashcard will be shown to them again promptly, so that they have another chance to correctly answer very soon after seeing the answer. Choosing 'good' will repeat that card in a moderate amount of time and depends on how well the user has answered in the past. Choosing

'great' will repeat that card after a large amount of time, though this is also dependent on how well you've answered that card in the past. These two flashcard screens are shown in Figure 9.

The algorithm that Anki uses is based on an open-source algorithm used by a similar program called Supermemo. This algorithm is called SM-2. Anki makes some modifications, like allowing the user more control over the algorithm's parameters. (Anki, 2020)

There are some lessons that can be drawn from this. Anki's use of spaced repetition has been very successful and is a major part of Anki's appeal. Also, Anki offers a significant amount of customisation, including to its algorithm, and how users learn. Allowing users, a great deal of control and customisation has helped Anki become a well-respected tool. Also, giving the user the ability to judge their own performance gives the user more control, but it also allows for a better learning algorithm, given that it is more fine-tuned to the individual.

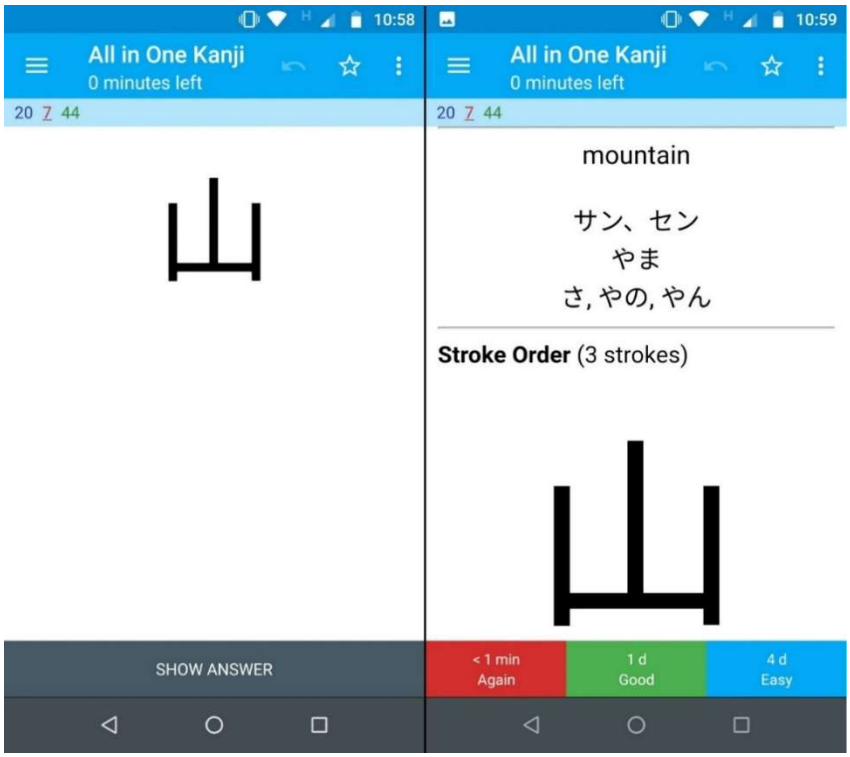


Figure 9. Both sides of the flashcard in Anki

Table 2. Anki Scoring Table

Criteria	Score /10	Comments
Usability	8/10	The UI is basic, but straightforward.
Topics	9/10	The range of Anki decks available online is extremely vast and allows users to choose virtually whatever topics they want. This is also true as users can create their own decks.
User Control	9/10	The user can choose which 'deck' of flashcards to learn from. Users can create their own decks, giving them as much control over their learning. Users can also judge how well they did on their answers, which helps to give them more control over their learning.
Customisation Options	9/10	Users can completely customise the app's algorithm to their preference. They can also change the app's appearance and font choice, storage location and even more advanced options like gesture control. Users can also create and customise their own decks of flashcards.
Effective Learning Tool	9/10	Anki uses Spaced Repetition. It is also possible to have pictures, video and audio on each flashcard to enhance the memory. Users judging their own performance in each flashcard could lead to a bias and slightly inaccurate evaluations. There is a lack of speaking practice, meaning it is effective for learning to recognise and understand words, but not necessarily use them.

Memrise

Memrise is a web and mobile app, which uses spaced repetition with flashcards primarily to help learn vocabulary, along with a few other tricks.

When taking part in one of Memrise's language courses, you're presented with flashcards and prompted to answer, either by typing the answer or selecting the correct answer from multiple choices. Memrise uses spaced repetition, presenting these flashcards at varying time intervals. After correctly answering over a certain amount of times, it deems it 'learned' and the user will need to revisit that flashcard to review them. Reviewing material is a cornerstone of the spaced repetition idea – the user must always revisit information to remember it, otherwise it risks being forgotten.

Each language course is often further broken down into smaller sections, or smaller stacks of flashcards. This is so that each stack is reviewed separately which leads to each stack being more focussed and allows the difficulty of each stack to be grouped and to slowly increase.

There are several other memory tricks that Memrise uses. For each flashcard, the user can create a 'mem' or select from several user-made 'mems'. A mem is an image designed to help the user remember that flashcard, usually made up of an image and some text. This is similar to the Keyword Method, an application of the elaboration method. By creating a link between the two sides of the flashcard, one can enhance the memory.

Instead of simply just being presented with text and being expected to know the answer/translation, Memrise also has video and audio clips that can be used, incorporating more elements of the word (sound, visuals, etc.), which make the learning more effective.

Memrise also uses gamification – there are various game-based elements that used to make Memrise 'fun' and motivate players to continue learning and return to Memrise regularly. After every session, the user is given a point score for how well they performed, based on their correct answers, speed and accuracy, as seen in Figure 10. This point score is the keystone of Memrise's gamification. There is a leaderboard in which the score is displayed alongside those of the people the user 'follows', showing the scores for the week, month, and for all time. This introduces a light competitive atmosphere, as users can compete with others. Another element of gamification is that the user can set a daily goal for each course, so that the user aims to reach a certain point score in that course. Reaching the goal every day builds up a streak, which is celebrated by the app.

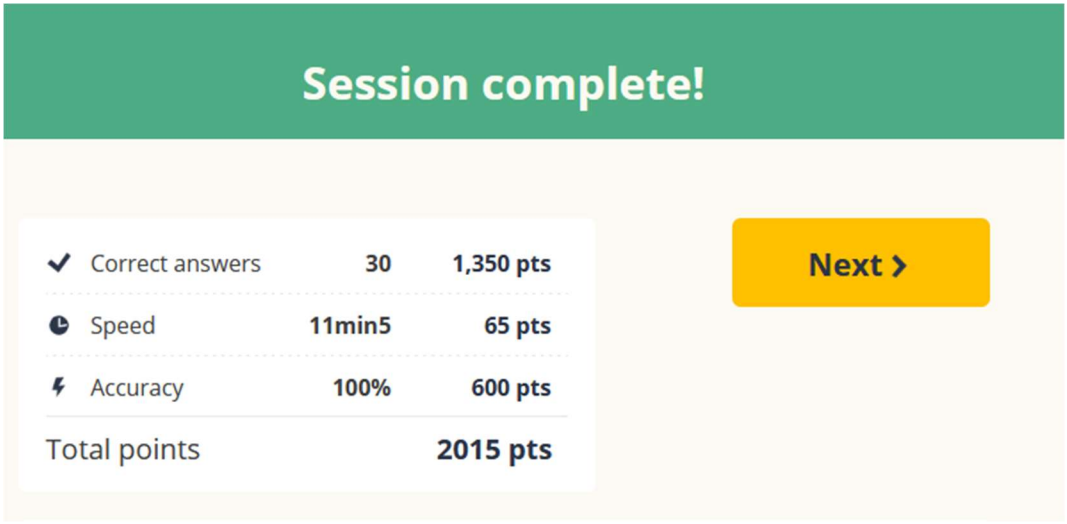


Figure 10. Memrise, screen after completing a session (Memrise, accessed 2019)

Within Memrise, any user can create their own course, in addition to the pre-made language courses made by Memrise. This has led to Memrise hosting courses for a range of topics, from psychology research to Morse code.

Also, there are different modes that Memrise supports – the 'learning new words' and 'review' are intuitive – they introduce words and revisit them. There is also 'quick review', in which the user is given only a few seconds to select the answer to a flashcard from four multiple choices. There is a menu where the user can select which mode to choose from, as seen in Figure 11, however Memrise does have a recommended mode to use by default.

Figure 11. Memrise menu to choose mode (Memrise.com, 2020)

There are a few useful lessons to be taken from Memrise. Memrise's use of a menu for selecting different modes give the user freedom to try different modes out. The app's use of a point system, and a daily goal is an effective method for motivating and gamifying their system.

Table 3. Memrise Scoring Table

Criteria	**Score /10**	**Comments**
Usability	9/10	UI is well-designed but still easily understandable. Everything is well signposted and straightforward. Gamification elements and daily reminders provide motivation for the user.
Topics	9/10	There are many community-created courses spanning a vast array of topics. Memrise-created courses are vast and split into a variety of useful topics. These start with useful conversational and tourist-based topics and then branch into the deeper topics about school and the home.
User Control	9/10	The user can control which activity/style of practice to do, including a strictly timed version, audio only and only video versions. Users can also create their own courses, much like Anki, giving them control over what they learn.
Customisation Options	8/10	There are some customisation options to customise the algorithm and media settings, as well as some limited UI settings. Users can also create and customise their own courses.
Effective Learning Tool	9/10	Memrise uses Spaced Repetition but it has fixed intervals between reviews, which makes it slightly inflexible. Using mems can be used to provide elaboration on flashcards, enhancing the memory. Gamification elements provide motivation and engagement which improves learning. Each flashcard can have text, audio, or videos which enhance the memory. Daily reminders also help encourage spaced repetition. There is a lack of speaking practice, meaning it is effective for learning to recognise and understand words, but not necessarily use them.

Duolingo

Duolingo is also a language learning web and mobile app, which uses some similar elements to Memrise (like spaced repetition and gamification) but implements this using a different structure.

When using Duolingo, the user selects the language they wish to learn, and then they are presented with the pre-made course, which is made up of various 'levels', which are focussed around a certain topic. Each level is made up of several lessons, which use spaced repetition and flashcards. Each level can be completed multiple times, with the difficulty / size of each level increasing. For example, for an arbitrary Level 1 would be made up on 5 lessons, where at Level 3 it would be 15 lessons.

When introducing new words, the words are presented alongside images, as well as the audio clip of each word.

Instead of the same flashcard popping up repeatedly, each lesson presents the user with full sentences, using the vocabulary in different ways and combinations. There is also a variety to how the user can practice these flashcards – they can simply type the answer, construct the answer from pre-defined blocks, or select the answer from multiple choices. They can also be prompted in different ways, such as straightforward text or audio clips. There is also an activity where the user must pair up vocabulary with each other in a collection of 5+ pairs.

Duolingo also uses gamification. After each lesson is completed, the user is awarded a set amount of XP (experience points). The user can set a daily goal in which they aim to gain a certain amount of XP a day. They are also celebrated for gaining a 'streak' for completing their daily goal every day.

The main lessons that can be taken from Duolingo are mainly to do with its structure. The way that it is broken up into topics allows for more concise, focussed learning. Their use of XP and daily goals also is an effective method of gamification. Their variety of question types is also effective, and can allow for a better quality of learning.

Table 4. Duolingo Scoring Table

Criteria	Score /10	Comments
Usability	9/10	Duolingo's UI is well designed and clearly signposted. Gamification elements and daily reminders provide motivation for the user.
Topics	8/10	Users cannot create their own courses in Duolingo, so all the topics and courses are defined by Duolingo. These topics are fairly expansive and cover many varied areas. The order in which the user can complete them is also well-designed. Having completed previous topics, newly unlocked topics should be doable.
User Control	4/10	Users can select which language to learn and select which of the unlocked topics to learn, but largely the user is guided through a curated 'route' of different topics.
Customisation Options	4/10	There are very limited settings available – mainly about toggling speaking or listening exercises and quality of life settings like toggling sound effects.
Effective Learning Tool	9/10	Daily reminders encourage the user to revisit the courses daily, as a form of spaced repetition, and within the courses themselves, each session is built on revisiting words repeatedly – and more so if the user answers incorrectly. This suggests of a spaced repetition system, but this hasn't been publicly disclosed. Audio and cartoon images are included alongside each word and there are numerous ways a user can respond to questions, which encourages the user to properly learn the answer. These both improve the quality of learning. Speaking practice is also possible. Each word is presented within a sentence to show how words are used in context, which is useful. Some of the sentences that are used are strange and don't make sense. This leads to confusion and could lead to learning sentences which would never be used. An example would be "The bear reads a newspaper."

2.2.1.2 VR Language Tools

This section describes existing language learning tools in VR and draws lessons that can be used. These were all tested and tried out using an Oculus Go.

To properly compare and judge each existing tool, there are a number of criteria that will be checked. These will be given scores, and the reasons will be detailed. These criteria will be:

- Usability. How easy it is to intuitively know how the tool functions. How straightforward it is to work.
- Topics. How broad or well-thought-out the topic range and the language content covered is.
- User Control. How much control the user has with regards to what they're learning and how.
- Effective Learning Tool. How effective it is for learning and remembering new languages. This will be mainly based on the techniques and algorithms used.

Customisability was considered as a criterion however at the time of writing, access to the Oculus Go had been lost due to the Covid-19 Pandemic.

Mondly VR

Mondly is a language learning company who are most interested in "introducing language learning to smart technologies". Their services include a web app, a mobile app, and a VR app. Mondly's VR app supports the learning of 29 languages, and also the use of 29 languages as the native language. It is available on all commercially available VR technologies. (Mondly, accessed 2019)

Mondly VR consists of 8 experiences that provide conversation practice. Each experience takes place in different environments that tourists would often find themselves in, for example, checking into a hotel or talking with the taxi driver. They usually take place using conversations, which uses text and audio resources and speech recognition. The 'other person' speaks, and their words are subtitled and read aloud. The subtitles are in the second language, but a translation can be shown in the native language. When the user speaks, they speak in real life, and the program picks this up and uses speech recognition to recognise it. The user must speak one of three phrases provided by the app. There is also a button next to any text to have the audio play.

The VR app is a stationary one; there is no movement required, both in real life or virtually.

Mondly has some features that increase the ease of use of the app – such as the ability to easily see the translation by hovering over the text and the ability to replay the audio on command. Having multiple choices when responding is a useful idea as well.

Table 5. Mondly VR Scoring Table

Criteria	Score /10	Comments
Usability	8/10	UIs are straightforward and understandable. Mondly puts the user in an area that is familiar. The graphics allow for an experience that is close to real life. Speech recognition does not always work, which can be frustrating for the user.
Topics	6/10	Mondly has topics mainly covering topics that would be useful for a tourist or in basic conversation. This is a bit limited.
User Control	3/10	Even with the app promoting communication with others, it doesn't allow for improvisation or active application of the learned language – instead only allowing pre-approved answers. The user has no control. They can select which topic to do from a number of set languages.
Effective Learning Tool	6/10	Does not encourage maintaining the learned words. Allowing the user to see the translation and hear the audio on command is useful. There are multiple options for the user to respond with, providing some variety of phrases that can be learnt. This also is a good memory technique, forcing the user to recall the answer. Largely focussed on the speaking and use of the language.

Busuu VR

Busuu is another online language learning tool that has created a VR app. This app is a targeted app to help learn one language (Spanish) using one language (English). It is also only available on Oculus technologies.

The Busuu VR app is a heavily crafted and guided experience. The app revolves around learning more vocabulary and speaking practice. The user is led around this experience by a guide, who teaches the user vocabulary and encourages the user to speak to various characters. The user is expected to speak a pre-defined phrase, which is picked up by speech recognition. The user can view all the vocabulary and phrases that have been learned in a dictionary. They can also learn new words by picking up items and taking them to the guide.

The app is a physically stationary one – it does not require the user to move in a space, however they can move virtually in the app by teleporting. In order to learn a new word, they can pick up an item and bring it to the guide character.

The movement in Busuu is a useful one, as it allows the freedom to explore and learn words at will, whilst being given guidance and direction. The dictionary is a useful feature that allows the user to review and practice at will.

Table 6. Busuu VR Scoring Table

Criteria	**Score /10**	**Comments**
Usability	8/10	UI is well designed and intuitive. Busuu allows for free movement within the environment. All controls are explained by the guide. Speech recognition does not always work.
Topics	6/10	The topics are fairly basic, and tend towards the home and conversational Spanish.
User Control	6/10	The app allows for user to control which words to learn by taking items to the guide. When it comes to speech, the user must say the scripted phrase that is expected.
Effective Learning Tool	7/10	Fully accompanied by audio clips. Allowing the user to review the dictionary when they want is a useful feature. There is only very limited encouragement for the users to maintain the learned words.

House of Languages

House of Languages is a VR app that teaches 3 different languages using one of 5 languages and is only available on Oculus technologies and SamsungVR.

The app focusses on vocabulary practice only. The app places the user in different locations – such as a living room, bathroom, or café. Within each location, there is a character who asks them to find an item – this item is read aloud, as well as text being shown. There are interactable items placed around the environment, which when looked at display the text of the word and the audio clip of how the word is said. Within each environment, the user finds and clicks all the items which the guide specifies. Then there is often an activity that uses them, acting as a light test. The user can see all of the words they have gone over in the menu. (Fox3D, accessed 2019)

The app is another stationary one – no physical movement is necessary. There are sometimes changes of perspective, but these are more scripted and are not controlled by the user.

There are a few lessons that can be taken from House of Languages. In this app, different rooms help teach different vocabulary which is appropriate for each room – this is like Duolingo's topics and

allows the user to focus on one context at a time. Interacting with the items is intuitive and helps show the user what the item is rather than just presenting the translation.

Table 7. House of Languages Scoring Table

Criteria	Score /10	Comments
Usability	8/10	UI is polished and everything is understandable. Easy to use controls.
Topics	6/10	Different rooms teach different vocabulary, and each word is contextually relevant in each room. There are only 10 rooms to choose from.
User Control	4/10	Besides choosing which language and room to go into, the user has no control over anything that occurs.
Effective Learning Tool	7/10	Cartoony graphics may help make the experience more memorable. The user interacts with items to learn each item, which is a nice technique, rather than a direct mapping between the word in the native language and the word in the foreign language. This way the item is associated with the foreign word directly. The app allows the user to review learned words once a level has been completed. However, it doesn't encourage user to maintain the learned words. There is a light testing activity at the end of each scene. The user is only prompted to learn each item once. There is no help provided to learn via any method; the app just presents the word to the user and leaves them to learn it. The character only ever says one word at a time - the item the user needs to interact with. There is never any sentence that the word is used in. This is simple and could be developed.

2.2.2 Existing Technology

2.2.2.1 Existing Hardware

There are several different varieties of hardware available for Virtual Reality, which cover a large range of power, price and availability. This project will be looking at Head-Mounted Displays (HMD), as research shows these to offer a better immersive 3D experience.

Headsets such as the Oculus Rift, HTC Vive and the Valve Index are examples of the commercially available top-of-the-line HMDs. They require to be connected to a powerful PC, and they offer the best quality VR experience. They also often come alongside sensors so that there can be accurate head-tracking and movement within a 'play area'. This allows for the user to move around physically

and the programs can reflect this accurately. The programs that can run on these headsets can afford to be large and complex. As the headset is using an accompanying computer to do most of the processing, this technology is the most powerful. However, these technologies are also some of the most expensive, ranging from £226 (Playstation VR) to £1000 (Valve Index).

Headsets such as the Oculus Go are more affordable, and don't require a PC. Everything that is needed is on the headset itself. There are accelerometers on the headset, and so head-movements are tracked and reflected in programs, however no movement is tracked and so users are stationary. The programs that can run on these technologies are simpler, due to the lower spec machines that can be fitted into the headset. These technologies are more moderately priced, ranging from £200 (Oculus Go) to £400 (Oculus Quest).

It is also possible to use a smart phone for VR. These (often) have the fewest resources to work with, and so the VR applications that run with these are often toward the simpler end of the spectrum. MobileVR is still usually done using a headset, though these are much cheaper as they act as more of a case that holds the mobile phone in the right way. With most implementations of MobileVR, there is only one way to interact with the scene, and this is by tapping the screen. Head-movement is still tracked using the built-in accelerometers in the phone, however physical movement is not tracked, and the control possibilities are limited. However, this is the cheapest option, with headsets costing as low as £15 (Google Cardboard).

2.2.3 Discussion

There are quite a few existing apps that involve language learning, and all of these are effective and successful, however they all do some things better and differently to others. Anki doesn't involve VR, however it is a popular and effective tool for vocabulary learning. It's spaced repetition system and customisation options are fantastic, though its UI isn't as user friendly.

Memrise has a better user interface than the other mobile apps, and it has elements which help enhance the experience – such as gamification and different learning modes. This variety helps to improve the quality of learning, and can be controlled by a menu, which is a good idea. Their spaced repetition algorithm is slightly less advanced than Anki due to fixed, non-customisable intervals, and it also doesn't make use of VR.

Duolingo is perhaps the most different from Anki's flashcard approach. There's still an aspect of answering flashcards, though it has such a variety of ways in which the questions are asked and answered that it changes the experience. The way that Duolingo structures its content helps to focus

the user on certain topics. Its use of gamification elements also helps to improve the learning experience.

These three are effective methods of vocabulary learning, however none of them are virtual reality tools, and none of them do much to show the context around a item, apart from maybe a photo or audio of the item. There are also language learning VR tools, but these tend to focus on other elements of language learning rather than just vocabulary. These do tend to show words and phrases in their natural environment, greatly increasing the context provided per word.

Mondly is an app that focusses mainly on conversational language skills, and so its VR app focusses mainly on listening and reading foreign prompts, and then having the user speak their response. Whilst it is useful that the user can see the translations and hear the audio of what is being said, this app doesn't focus as much on understanding individual words or vocabulary, and more on what to say in certain situations.

Busuu is an app that attempts to teach the user a rounded skillset. It tests the user's speaking and listening and allows for vocabulary practice as well. The user is given freedom to go and explore the environment and choose what item to learn next.

House of Languages is a VR app which does focus mainly focus on vocabulary learning. However, this is very limited. The user can actively interact with items in the context of a room and learn what the word for them is. There is limited testing, which means that this item interaction is the only learning feature.

There is a gap in the market for all the best things of these apps to be combined. There's no good quality vocabulary-focussed language learning VR app. If the good quality aspects of the vocabulary-learning apps Memrise, Anki and Duolingo and taken, and merged with the good aspects of VR language learning from Mondly, Busuu and House of Languages, then there is scope for a VR tool that makes good use of VR, language learning, and specifically vocabulary learning.

Table 8. Vocabulary Learning Tools Comparison Table

Criteria	**Anki**	**Memrise**	**Duolingo**
Usability	8/10	9/10	9/10
Topics	9/10	9/10	8/10
User Control	9/10	9/10	4/10
Customisation Options	9/10	8/10	4/10
Effective Learning Tool	9/10	9/10	9/10

Using Table 8, it can be seen that the quality of each app is quite high, especially when reviewing the Effective Learning Tool, Topics and Usability criteria. Each app does have a downfall. Anki's interface is very basic, whilst Memrise's customisation options are less in-depth as Anki's. Duolingo doesn't allow the user much control over what topics they can learn or over the algorithm.

Table 9. VR Language Tools Comparison Table

Criteria	**Mondly VR**	**Busuu VR**	**House of Languages**
Usability	8/10	8/10	8/10
Topics	6/10	6/10	6/10
User Control	3/10	6/10	4/10
Effective Learning Tool	6/10	7/10	7/10

Using Table 9, it can be seen the quality of existing VR language tools is not quite as high as the existing vocabulary tools currently on the market. However, it is important to note that each of these VR apps are aiming at slightly different goals. Mondly aims towards speech practice and conversational skills, rather than pure memorisation, and Busuu aims for a guided course into conversational Spanish. House of Languages does aim to improve vocabulary skills but has very little customisation and control options.

2.3 Questionnaire

This section describes the questionnaire, its results, and the conclusions that can be drawn from it. It also evaluates the questionnaire. A transcript of the questionnaire is in Appendix A.

2.3.1 Rationale

To provide some insight into public opinion, an online survey was ran (in the form of a short questionnaire) to gauge public interest in using VR as a language-learning tool, as well as gaining some insight into which medium was thought to be the most effective to learn a second language.

2.3.2 Sample

My target population was any adult, though our target population could be focussed just to adults who have started learning a second language. Adults are being targeted only as asking questions of anyone younger would mean that the questionnaire would need to be checked by an ethical committee. It could be more beneficial to target only those who have learned a second language as they will have a more informed opinion on the effectiveness of each medium.

In order to build the sample, a link to the questionnaire was shared to several online groups, allowing participants to voluntarily complete the survey. This is an example of convenience sampling. The groups that were sampled from were a number of university Facebook groups (Computer Science, Virtual Reality Society, Video Games Society, Loughborough Freshers) and also an online language-learning interest group with over 200k members (reddit.com/r/languagelearning).

2.3.3 Results

There were received 201 responses to the questionnaire. The distribution of ages from the participants can be seen in Figure 12. There were 11 responses from people aged under 18 years old. If this age category was given as a response, the questionnaire was ended early. This is because this age group shouldn't be measured for ethical reasons.

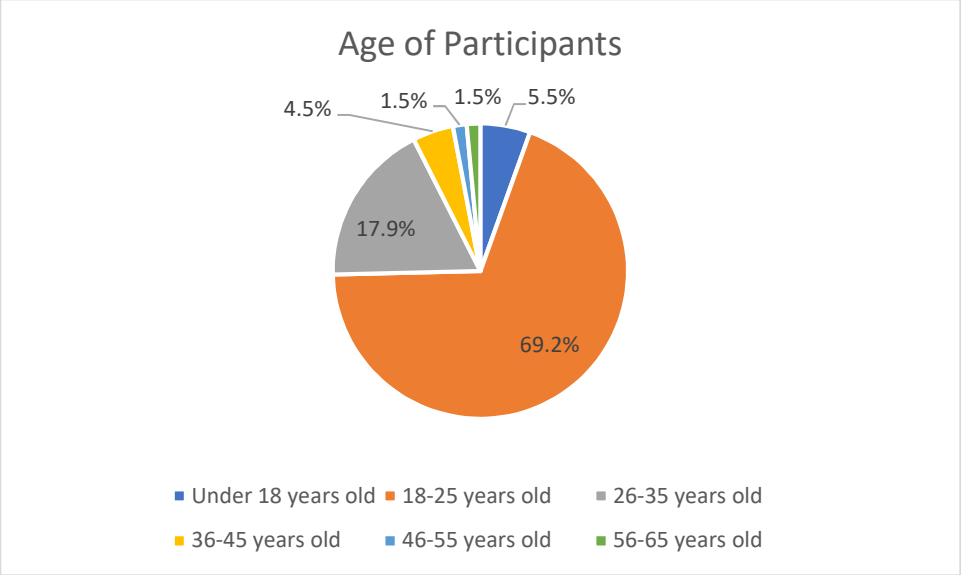


Figure 12. Graph of questionnaire participant age distribution

When asked "Have you ever started learning a second language?", this questionnaire found that 97.4% of the participants had taken part in some form of second language learning. Each participant also evaluated how effective different mediums of language learning were. The results can be seen in Figure 13. Classes and one-on-one sessions gather the most positive responses from our participants, while websites and textbooks gather more moderately positive responses. Classes and textbooks are both traditional learning methods, with years of evidential research and expertise supporting and developing these methods. It is also likely the most common. Websites can also offer good quality learning courses and resources; however, the internet also provides a wide amount of freedom – learners are able to search for topics that they want to learn, and this exploration is also a positive. There is also the possibility that if a website is not effective, the learner will simply find a better website, thanks to the sheer amount of material on the internet. This would lead to the medium of 'websites' being effective merely through the idea that 'there's always something good on the internet, if you look hard enough'.

Mobile apps are moderately effective, whilst forums are the least effective. VR applications also has been rated as more moderately effective, however most responses were 'N/A', signifying that many have not used this medium, and are not familiar with it. This could be for many reasons, for example it could be because VR technology is not very available, or it could be because there is a lack of good, or popular language learning apps on VR.

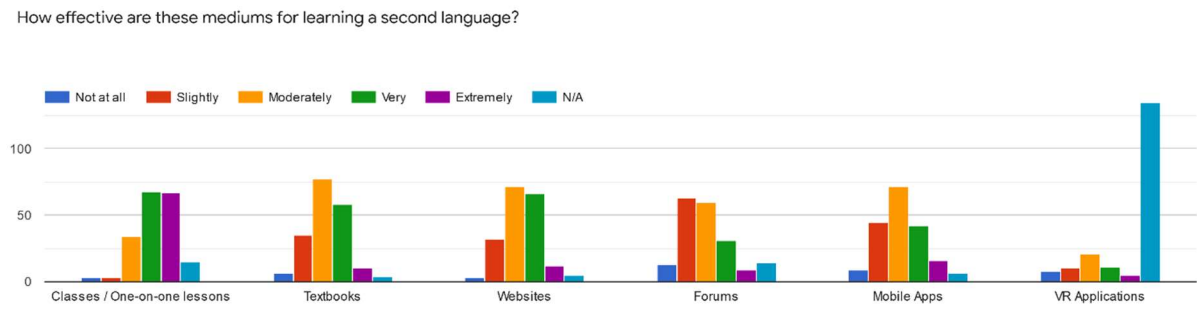


Figure 13. Chart of results to the question: How effective are these mediums for language learning?

The participants were informed that this project was making a VR application for vocabulary memorisation. When asked if they were asked if they were interested in using such an application, 91.6% of respondents said yes. This is shown in Figure 14.

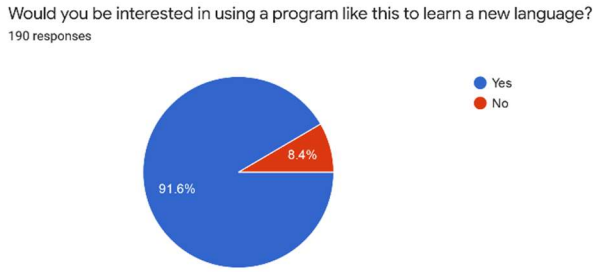


Figure 14. Chart of questionnaire results judging interest in this project's vocabulary learning tool in VR

2.3.4 Conclusion and Evaluation

Given that a very high percentage of my respondents have some experience learning a language, the fact that there is also a very high percentage of my respondents interested in using a VR vocabulary memorisation tool is encouraging and demonstrates that a fair amount of the public are interested in such an application.

The sample was intended to be more of a very general population; however, the sampling method may have skewed this population slightly. By sharing to groups already interested in virtual reality, video game and language learning, the sample population may have been encouraged to be more receptive to the idea, and so some of the responses may have been skewed. This lowers the reliability of the study, as some of the conclusions may not be completely valid.

Some of the questions were worded in such a way that some of the answers may have been ambiguous. The question ‘Have you ever started learning a second language’ is one example, as ‘learning’ has different stages, and some may consider themselves to only ‘start learning’ at a higher stage than others.

2.4 Requirements Specification

This section will detail the requirements of the project.

Minimum:

- Users can learn, practice and self-assess new words in a foreign language.
- Users can interact with items and characters within an environment.
- Users can freely move around the environment.

Core:

- Users can choose from multiple languages.
- When interacting with characters, there will be several methods used to present information to the user.
- When interacting with characters, there will be several methods for the user to respond to the character.

Extended:

- Users can manually change the language to customise to any language for themselves or for other students.
- Users can manually edit their own settings to optimise the experience for themselves, such as customise the size/appearance of subtitles as well as audio settings, along with others.
- The user can build a dictionary of terms they've learned.
- The user can progress from room to room by completing a test by the door.
- Data gathered from the application can be analysed to provide useful insight.

Chapter 3 – Design

This chapter details the design stage, and how each element of the final application was planned and designed. It also details and explains each design decision made and why.

3.1 Learning Process Design

A key component of the project is language learning, and the processes that are designed in this chapter underpin the functionalities and features of the application. Three core elements of a learning process can be extrapolated loosely from the experiential and problem-based learning methods. They are:

- Learning (how the program will introduce words to the user)
- Practicing (how the program will allow the user to practice those words)
- Testing (how the program can be used to assess the user's proficiency)

To create the learning process, a few basic elements were determined first – the basic ways to interact with the program and the virtual environment.

3.1.1 Interaction Methods

There are two key methods in which the user can interact with their virtual environment:

- Items
- Characters

There will be items placed around the environment that the user can interact with. When the user clicks the item, a speech bubble will appear, with the item's word in the foreign language. An audio clip will also be played. This interaction is similar to the item interaction in other language learning VR apps. The user would also be able to hover over the word to show the native language translation. This is also a common element in language learning apps and can be seen in Mondly as an example. This interaction is demonstrated the activity diagram in Figure 15.

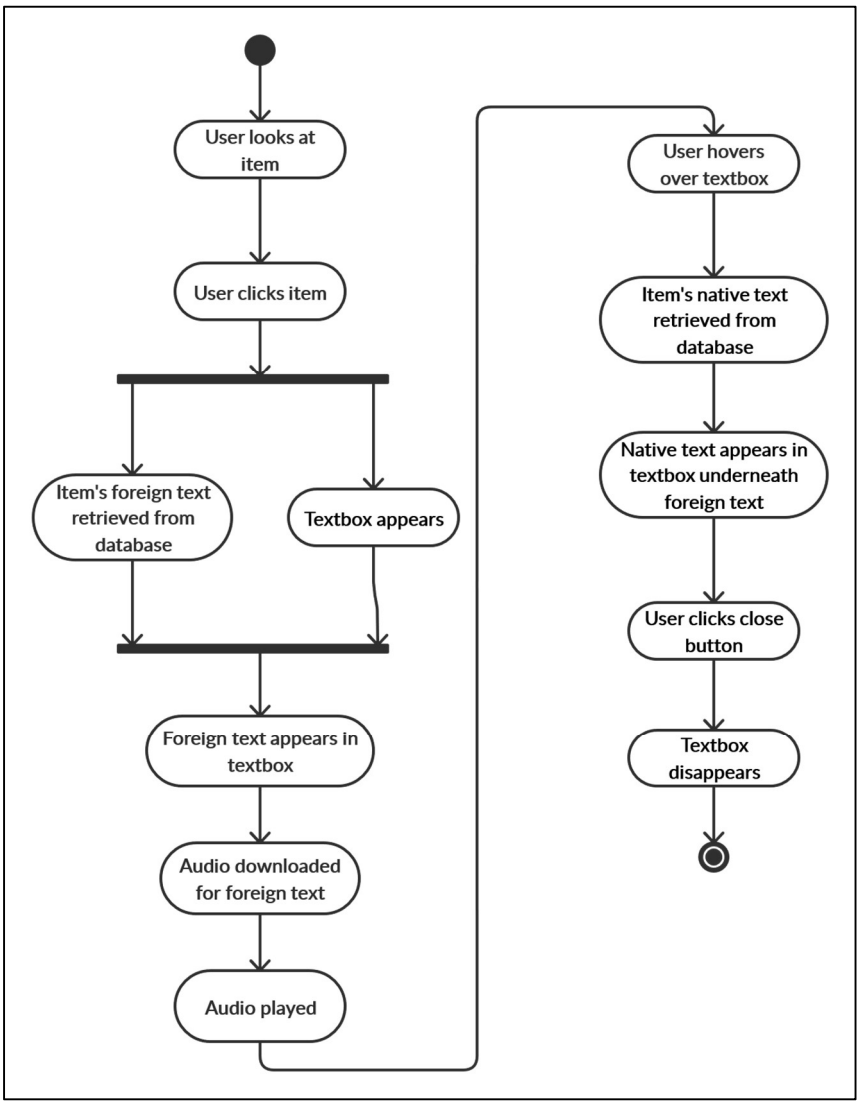


Figure 15. Item Interaction Activity Diagram

There will also be characters in the environment which the user can speak to. These will be predominantly text-based conversations, as seen in Mondly. The character will speak out loud, accompanied by an audio clip, with the second language text displayed in a bubble. The user can hover over this to see a native translation. The user can reply using several methods, which are designed further in section 3.1.2. This basic interaction can be seen in the activity diagram in Figure 16.

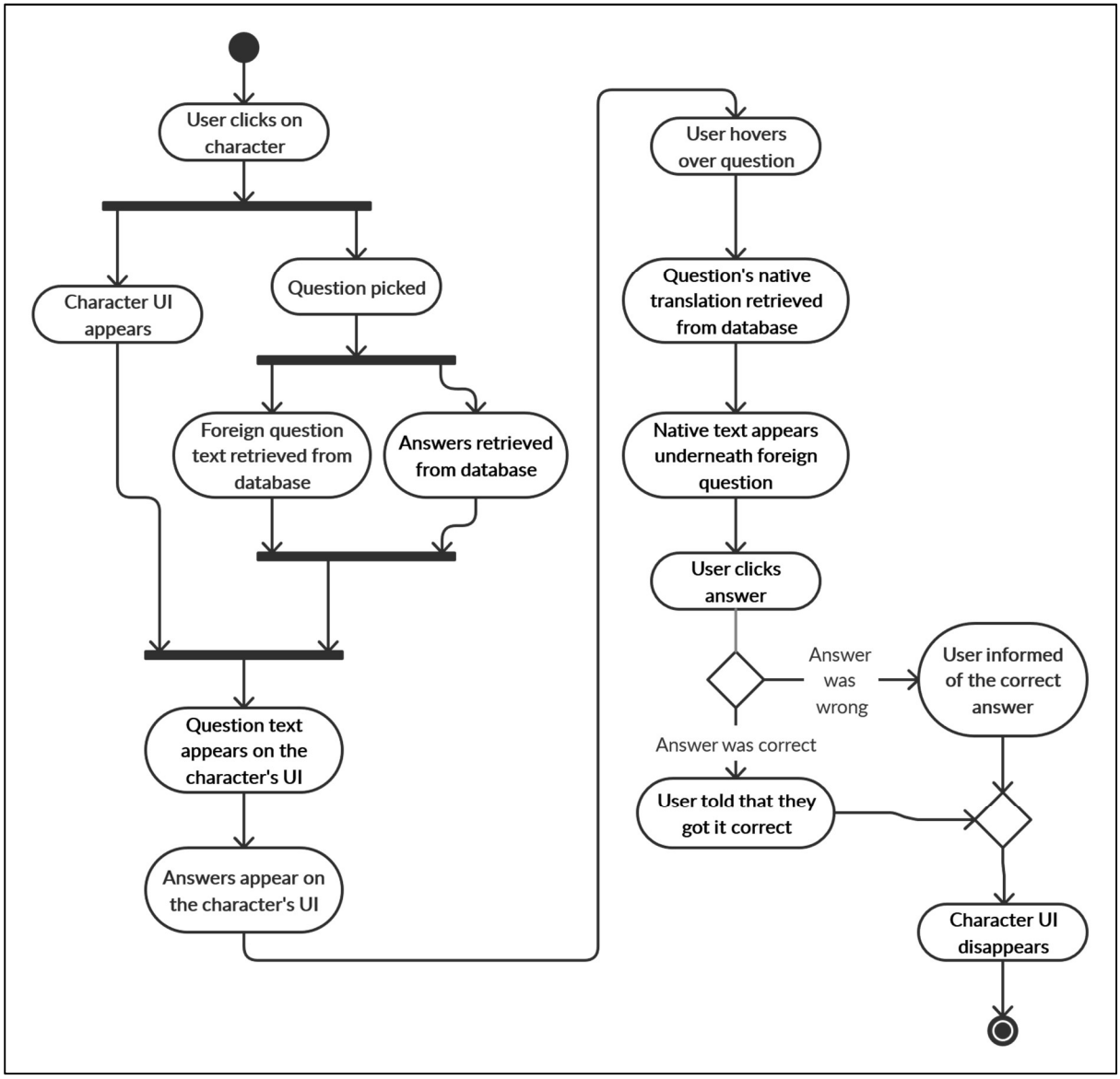


Figure 16. Character Interaction Activity Diagram

There will also be a dictionary in which the user can add words that they have learned. This is an element from Busuu which was useful and provides a concise place where all of the words are stored. Also, allowing the user to put their own words into the dictionary gives the user more control over their own learning. This dictionary will be used to create a deck of flashcards which can be used

for a more traditional spaced repetition experience. By creating their own decks of flashcards, the user thinks about what words to add, which can help memory. Also, the act of using the flashcards is a very common memory strategy, as can be seen in apps like Anki.

The environment will be a space in which there will be many options for items that can be placed. An example would be a room such as a kitchen, as there are many items that can be placed inside to be used as vocabulary words, like all the utensils and utilities. This will maximise how useful each scene is. There can also be a door, which when interacted with, leads to another room. However, this project will likely focus on one or two rooms. The user will also be able to move around the room at will.

3.1.2 Character Conversations

Character conversations provide many opportunities to enforce and enhance the user's learning. There are many ways this interaction can be used.

The character uses full sentences, along a similar line to Duolingo – where instead of prompting the user simply with one word, the user will see and hear a full sentence, which uses the words that are being taught. The sentences that the character uses will be aimed to be simple enough to work out, or just a level above. This level above will challenge the user and help them to acquire the language, as described by Krashen's comprehensible input. The level of the conversations will increase gradually and will be informed by the frequency of correct answers.

The 'conversations' can consist of a simple question and answer, or an interaction where the character and user both only 'say' one thing. However, these conversations do have the potential to be made into longer, more comprehensive conversations.

The words that the conversations will focus on will be different at different times and won't focus on the same words at the same time. This will use the concept of spaced repetition to revisit words after a time interval.

The character's speech will be presented to the user as both audio and text (second language), by default. It will be possible to display the L1 (first language) translation by hovering over the text. Clicking the text again will play the audio clip again.

The way that the user replies to the character is important, and there are a few methods that can be used, where each is backed by a different memory strategy. The app will decide when to use either of these two methods during conversations, depending on the user's ability level.

1) The first is simple multiple choice. The user is presented with multiple options to choose. Hovering over a question can display the translation. When selected, an audio clip will play. This is a common question type, so that it will be easy for the user to understand what they need to do. This also allows the user to recall the correct answer, or use problem solving to work out the correct answer using the items in the room.

2) The second is 'answer blocks'. In this method, different words/letters are broken down and presented in a random order. The user must select the right responses in the right order from these blocks. There can also be false blocks in this selection. This is a method where the user is prompted to remember the answer, rather than just select the right answer, and helps the user to remember, as shown in the idea of 'retrieval practice'. This is a question type which will challenge the user to fully remember the answer. They will need to know the answer as they will be spelling the word out.

There are several modes that could be worth looking into, such as a 'pairing up' exercise like the one seen in Duolingo. It was also considered to implement a form of speech recognition, which is also used by several existing language learning VR apps. However, it is not currently planned to use this feature currently as the technology is not mature enough, and it has a level of accuracy that is low enough that it may frustrate users.

There can also be different 'modes' in which the character can communicate with the user. Each of these offers slightly different challenges to the user and provides a different experience.

1) The basic method that the character will use to communicate will be, as already said, the L2 (second language) text and audio. Hovering over the text will display the L1 (first language) translation. Clicking the text will play the audio clip again. Accompanying the text with the audio will provide the user with more information about the words which helps with language learning.

2) Same as method 1) but no translation offered. A more difficult method that can be used to test the user, it can offer the conversation with no translation.

3) Audio only. A more challenging method is presenting the user with the audio clip alone. The user can replay this audio clip by clicking on an icon.

The app will have a menu where the user can select which type of conversation to have. This could be implemented as individual settings. This allows the user the most customisation control.

3.1.3 Learn

The user will be able to freely explore and move around the environment and interact with items. This freedom allows the user to make decisions on what item to look at next, as well as look at each item to learn the translation. This is done to give the user a stronger sense of control, as well as giving them opportunity to explore themselves.

The user can also talk to the character, which will begin by asking about certain items, or giving light tasks to introduce each item. This will introduce each item to the user, and if they are unaware of what the word means, they can hover to see the translation or go and find the item itself to learn it.

3.1.4 Practice

There are multiple methods to practice words.

Talking with the character will generate questions showing the words used in different contexts, so it can help familiarise the user with its grammar and (depending on the language) the different conjugations. There also multiple methods the user can reply to the character. All these methods use different memory tricks to learn and practice the words.

The character will use a spaced repetition system for which questions it asks. The questions will be cycled and repeatedly reviewed at the best time for the user. This will help maintain and practice the word.

Completing activities will award the user with points. The user can set a daily goal to receive a certain number of points. It is also possible to celebrate the 'streak', or number of consecutive days the user has reached their goal. This gamification can help boost motivation, as well as enhance learning. Similar ideas can be seen in the streak system of Duolingo and Memrise. The streak will encourage more regular sessions, and won't encourage large sessions, but encourages more the idea of 'little and often'. This is better for memory as can be seen from the research on spaced repetition.

Each room will also have a daily quest, or list of activities to complete. This will be optional but will help provide some guidance and direction. It's also another element of gamification.

There will be the deck of flashcards which the user can use to practice the words they've learned. This is a more traditional option to improve language, and doesn't necessarily take advantage of Virtual Reality, but provides another practicing method.

3.1.5 Test

In order to help measure progress, and provide the user with a feeling of progression, it's important to provide method to help test the user and assess them.

The user will be able to use the character's different 'modes' to make the questions more challenging and these can be used to help test the user. These modes can be implemented as a series of settings and can be configured so that the user can turn translations off to challenge themselves without being able to rely on language they can understand easily. The user might also be able to only use audio recordings rather than any text at all, for example.

Each room will have a door, through which it could be possible to move to a new room. However, the door will be locked until a 'riddle' is completed. This riddle will be a few instructions completely in the foreign language, with no option for translation. Example instructions would be 'open the fridge', 'pick up a cup', etc. All the words in the riddle will be available to learn throughout the room, but the user must learn these words. After the user moves from this room, they will be able to return if they wish to learn more. This riddle can provide the central 'problem' of problem-based experiential learning. The user will want to identify and learn the words that they do not know, so that they can pass the riddle and progress.

3.2 Technology Design

Out of the range of existing hardware, working for the cheaper system seems like the best option – this is phones and the Oculus headset-only devices like the Oculus Go and Quest. These are the weaker forms of VR, with access to (usually) the least computational resources, excepting the state-of-the-art phones with increasingly impressive specifications. The key element that influenced this decision was availability. Smartphones are plentiful for MobileVR, and a significant portion of the population has one. Many more people have access to a smartphone than have access to a powerful computer and expensive VR headset. The Oculus headset-only devices are still cheap enough to be considered as the investment of buying such a VR device is significantly less.

However, there are more limited controls using this hardware, and more limited computational resources. Using this hardware also makes it impossible to physically move around a room and for that to be reflected accurately within the app, and so alternate methods of movement must be used. This may impact the quality of the final application, but it was decided that the benefit of availability outweighed the possible negatives.

3.3 Scene Design

In order to plan the environment in which the application will be built around, it was necessary to design the environment. Choosing the best design is important to create an effective application. This section presents these designs and judges them based on the following guidelines:

- There is enough space to move about easily
- There is enough space for items to be placed
- It is not too small
- It reflects a room that could realistically exist
- There is enough wall space for doors, windows and for messages to be shown

3.3.1 Kitchen

A type of room with good potential is the kitchen. There are lots of possible items that can be found in a kitchen, as well as various options for interaction. An example of this is opening the fridge to reveal various food items inside.

3.3.1.1 Kitchen A

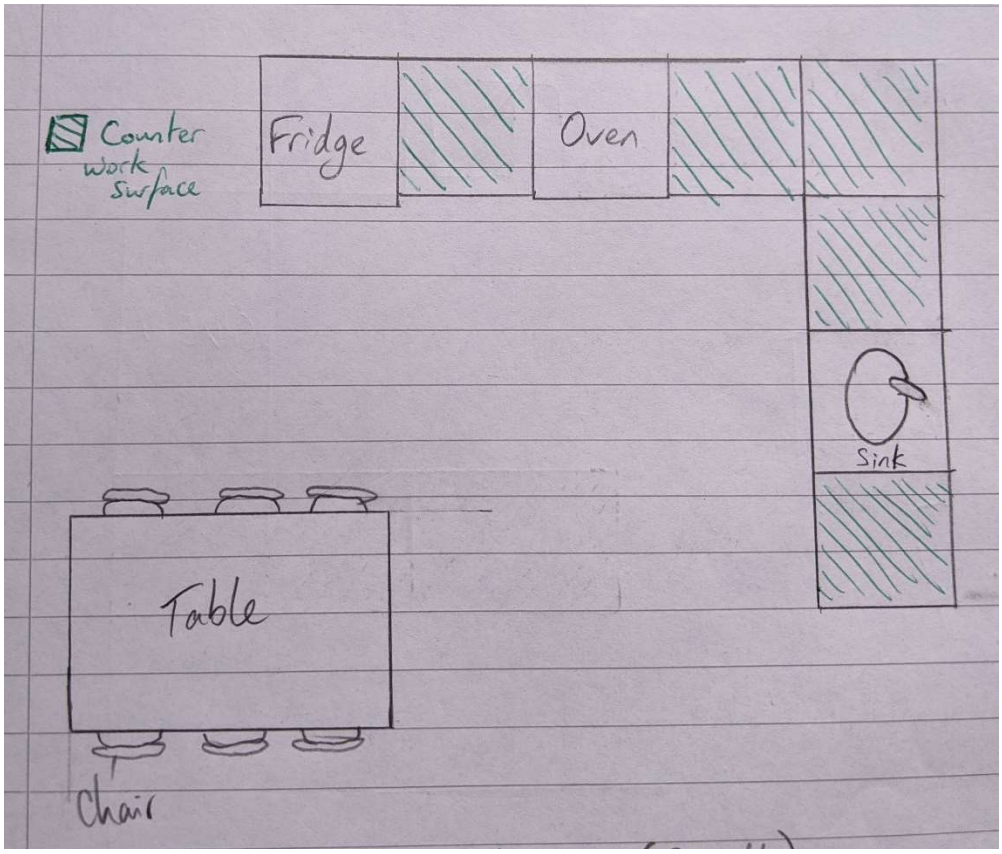


Figure 17. Kitchen A Design

Kitchen A is a design set in a large room, with a large space in the middle. Counters line two of the walls, with a table and chairs occupying the opposite corner. There is a good number of flat surfaces on which items can be placed. There is also a fridge and oven which can be interacted with. There is a good amount of wall space and space for movement.

3.3.1.2 Kitchen B

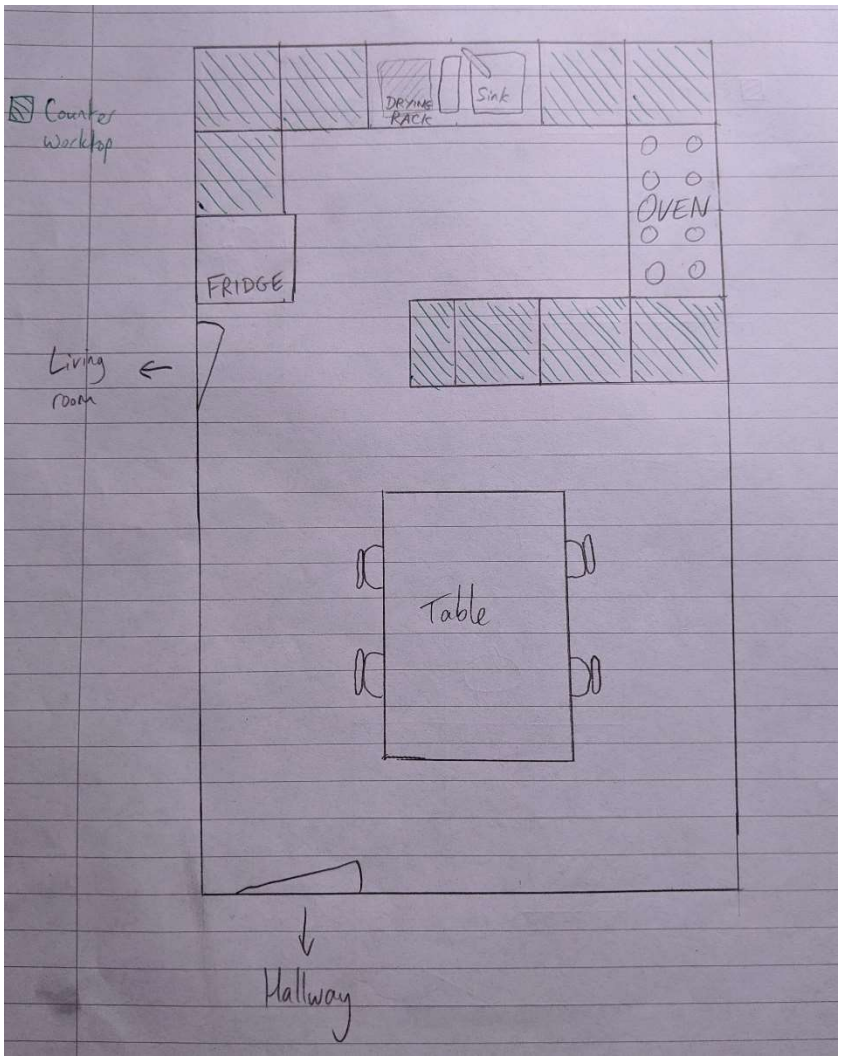


Figure 18. Kitchen B Design

Kitchen B is a design set in a large room, though this is a longer room rather than a square room. On one side of the kitchen is a traditional kitchen area with an oven, sink, fridge and work surfaces. On the other side is a table and chairs. There is a lot of flat surfaces for items to be placed, as well as utilities such as a fridge and a kitchen. There is also a table, and ample wall space for windows and doors. However, whilst it is a big room, there is not too much space to move. There is a row of counter worktop jutting out into the room, separating the 'kitchen' area and the 'table' area, with a

fairly limited space to move between these areas. When working with an already limited movement system in VR, this is not ideal.

3.3.1.3 Kitchen C

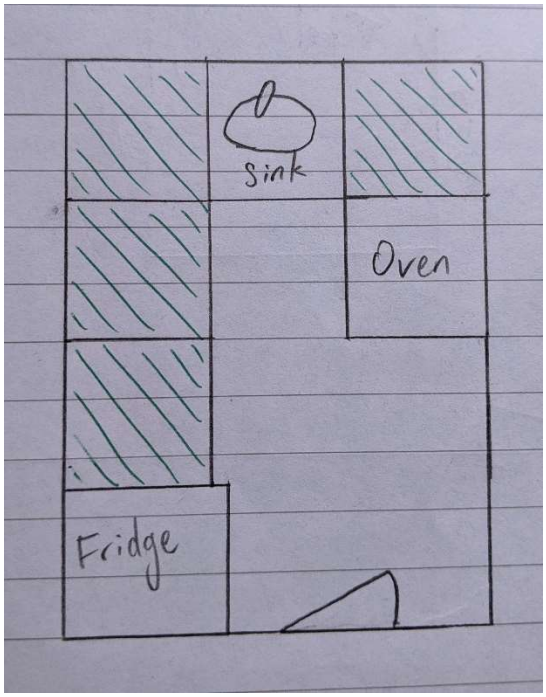


Figure 19. Kitchen C Design

Kitchen C is in a small room. There are the same utilities that are in the other rooms (oven, sink, fridge), with some work surfaces. There is no table in this design. However, the room is very small and would not allow for much movement. This is not necessarily a big problem, as all the items would be close to the user, and movement would not be necessary. However, there are other issues – there is not a significant amount of wall space for a door, window, nor any messages.

3.3.1.4 Discussion

Out of the three kitchen designs, Kitchen A is the best. It is spacious and allows for easy movement; it has plenty of wall space for doors, windows and messages, and it has plenty of floor space too, in the case that there is a need to add more to the room. It has utilities commonly found within a kitchen and has a lot of space for items to be placed.

3.3.2 Office

An office is also a type of room in which there is a lot of possibilities. Many different types of object can be in an office space.

3.3.2.1 Office A

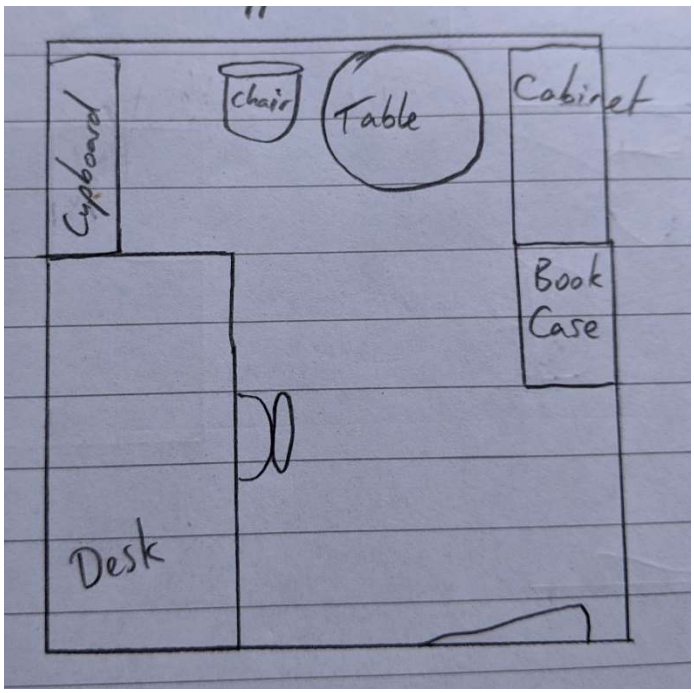


Figure 20. Office A Design

Office A is based on a small office, with a more compact feel. There are many items and pieces of furniture in this room which can be used. Items can be placed on them or they can be items themselves, like a bookcase or desk. However, there is not much space to move, and whilst this isn't a huge issue as if everything is close to the user, there is not much wall space to display messages, or to have a character. It may start to be too small and feel crowded.

3.3.2.2 Office B

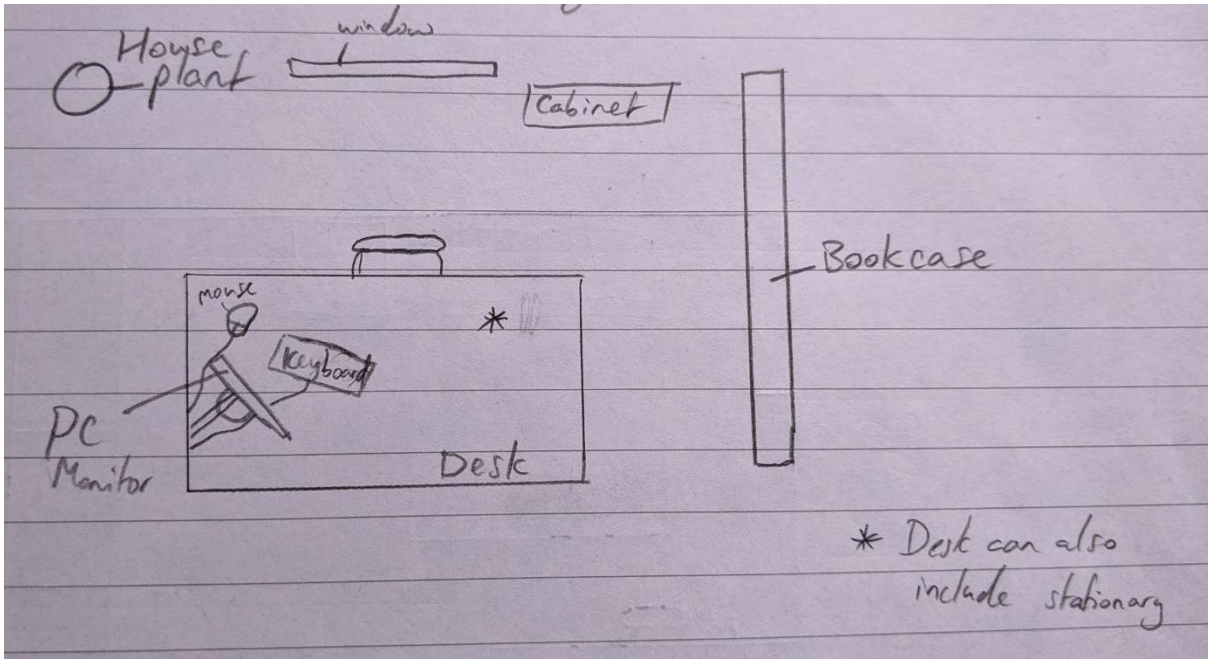


Figure 21. Office B Design

Office B is a much larger office. There is plenty of space to move around and also pieces of furniture that can be used. There is also a lot of wall space and surfaces for items to be placed.

3.3.2.3 Office C

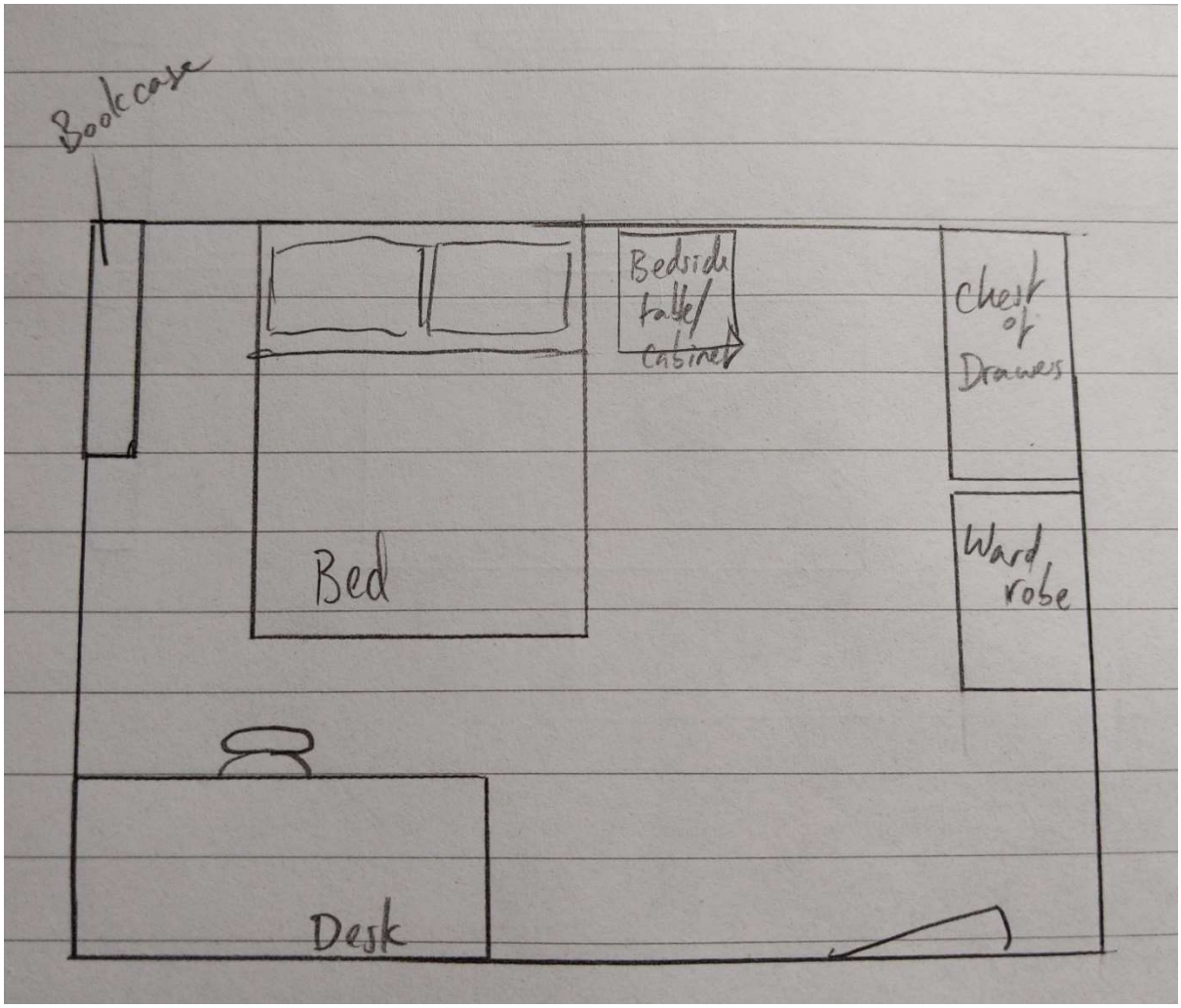


Figure 22. Office C Design

Office C is a take on a home office or a student's bedroom/office. This room is mainly a bedroom with office elements within it. It is realistic, as many students have this setup, but may be too unfocussed as a room with both bedroom and office elements. It is a large room, with many pieces of furniture and wallspace, as well as surfaces. The floor space around the bed and desk may be a bit limiting, which may make movement more difficult for the user.

3.3.2.4 Discussion

The best office design is Office B. This offers the greatest freedom, with a wide area for movement, space for utilities and items, as well as wall and floor space for doors, windows, messages, and any additional items that will be placed.

3.3.3 Discussion/Conclusion

Out of the proposed floor plans for different scenes, the kitchen is the best. The kitchen offers a wider variety of items, as well as a larger space. This is key as it leaves the greatest amount of flexibility, in terms of item variety and placement and interaction opportunities. There are also greater options for where the character, windows and door will be placed. Therefore, Kitchen A was the decided design.

3.4 UI Design

3.4.1 Item Interactable

When an item is interacted with, the plan is to display the foreign word and play the audio file. The method to display the foreign word and optionally the translation would be using a basic textbox, like the one drawn in Figure 23.

This textbox is intuitive and simple, with plenty of space to display the text. There is also a button that can be used to close the dialog, which is designed as a standard red cross. The decision to make this simple was made to help the interfaces be easily understandable, so that the guidance and tutorials on how to work with the app could be kept to a minimum.

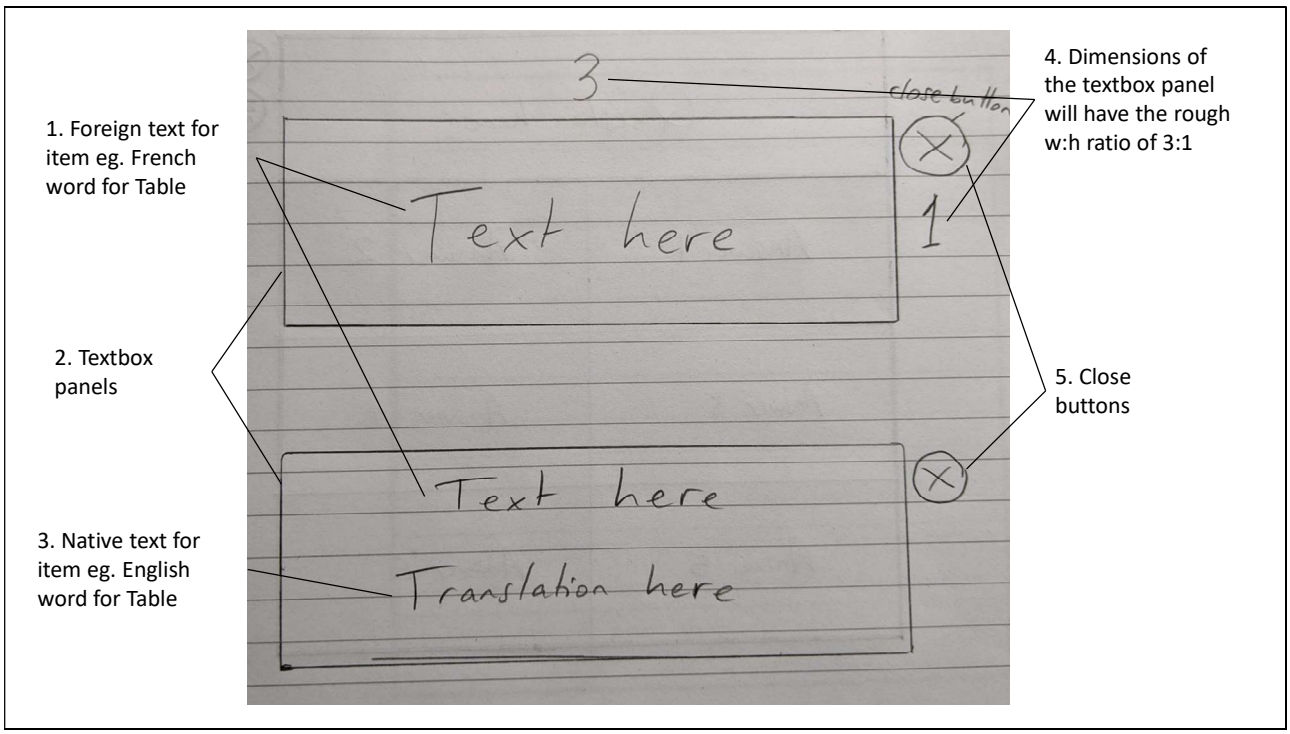


Figure 23. Item Interactable Design

For the foreign word's audio to be playable on demand, the user would simply interact with the item to play the sound.

To show the translation on demand, the user would hover their cursor over the textbox to display it. To display it, the foreign text would be shifted upwards and the translation would be displayed below it, as in the bottom design in Figure 23.

3.4.2 Character Interactable

When the character is interacted with, a dialog will appear with that character's question and possible answers. An audio file will also play the audio of the question. It will be possible for the user to see the translation by hovering over the question text.

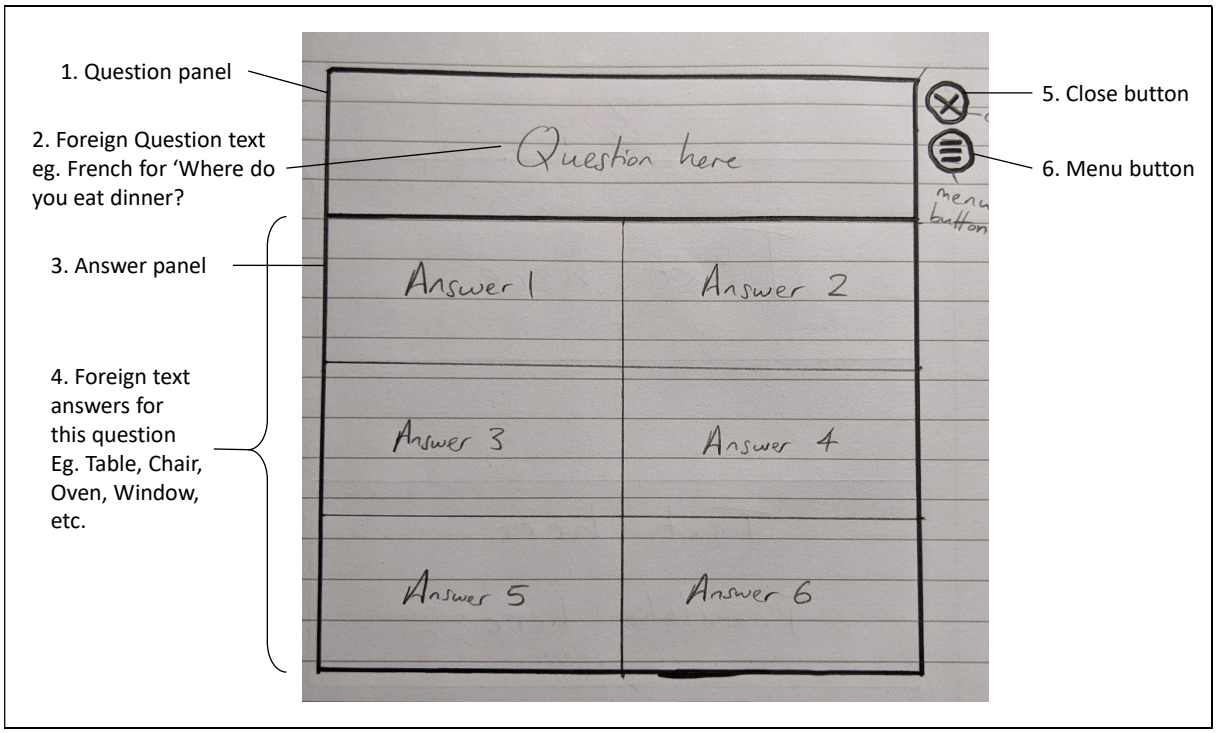


Figure 24. Character Interactable Design

When designing the interface for this, it was important for it to be simple and easily understandable. The final design detailed the question text at the top, with spaces for up to 6 answers below this, as seen in Figure 24. This is a very common way of presenting questions and answers, and the main influence for this design is the layout of many televised quiz questions, as can be seen in Figure 25.

Figure 25. Example of question layout used in quiz shows (shutterstock.com)

3.4.3 Menu Scene

The menu scene will be displayed upon the start-up of the application. The menu can be used to select which room to start in. It should also be possible for the menu scene to be accessed by the user from within a room. As this menu scene is a central location for any rooms to access, it could also be used for additional functionality such as to display statistics, a tutorial or to change some settings.

The room in which the menu is located can be a basic, non-interactable space. It is desirable for the menu to be stationary within a room, as this would mean the user can familiarise themselves with looking around a 3D environment and interacting with static interfaces. This idea also builds on the usual use with existing VR apps.

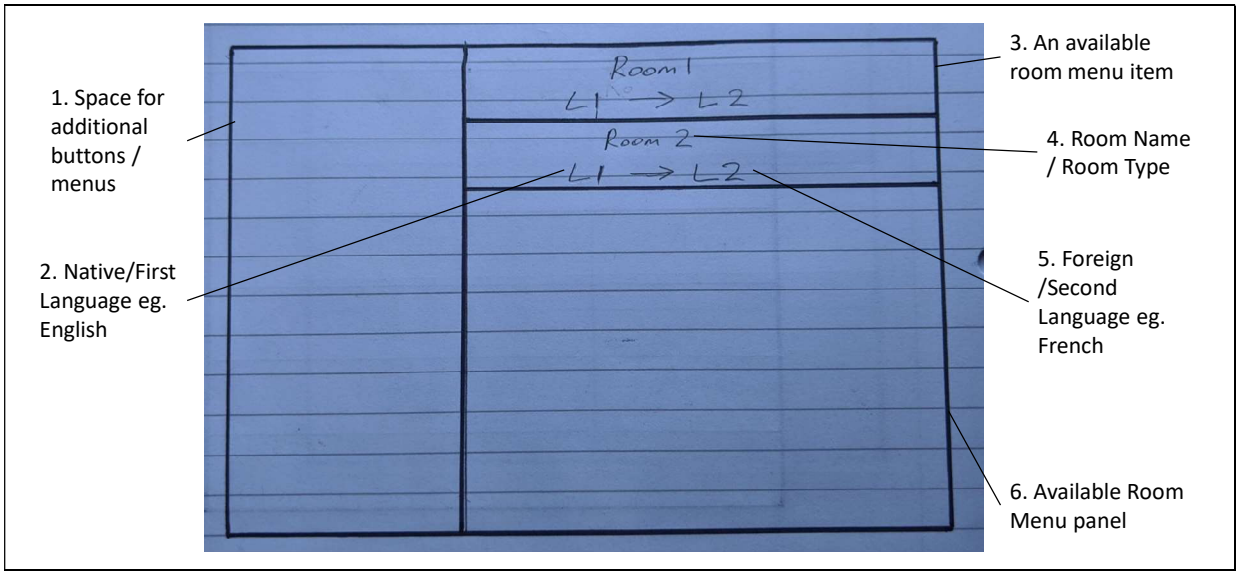


Figure 26. Menu Scene Menu Design

The menu itself will have a list of available rooms on the right-hand side, where each room lists its name and the languages associated with it. A room's name will be used to show the user what is inside that room and give the user an idea of the context of that room. An example of this would be 'Kitchen' or 'Office'. Details of the languages are necessary as they show which language is the 'native' language the user already knows and which foreign language the room is meant to improve. This is shown in Figure 26.

A space on the left of the menu screen is left blank for any additional functionality that can be added. These menus will use many of the basic design principles and won't introduce any new features.

3.4.4 Settings Menu

In order to give the user control over some aspects of the app, it was necessary to implement some settings. These settings would be implemented within an existing interface, for ease of use.

It was decided to implement the settings menu within the existing character interactable because many of the planned settings impact the functionality of a room and so it is useful for the user to be able to control these settings whilst in the room.

The menu screen will use the same dimensions as the character interactable, and have settings listed vertically, as shown in Figure 27. This was the simplest way to implement this, and it is a clear way of showing the settings.

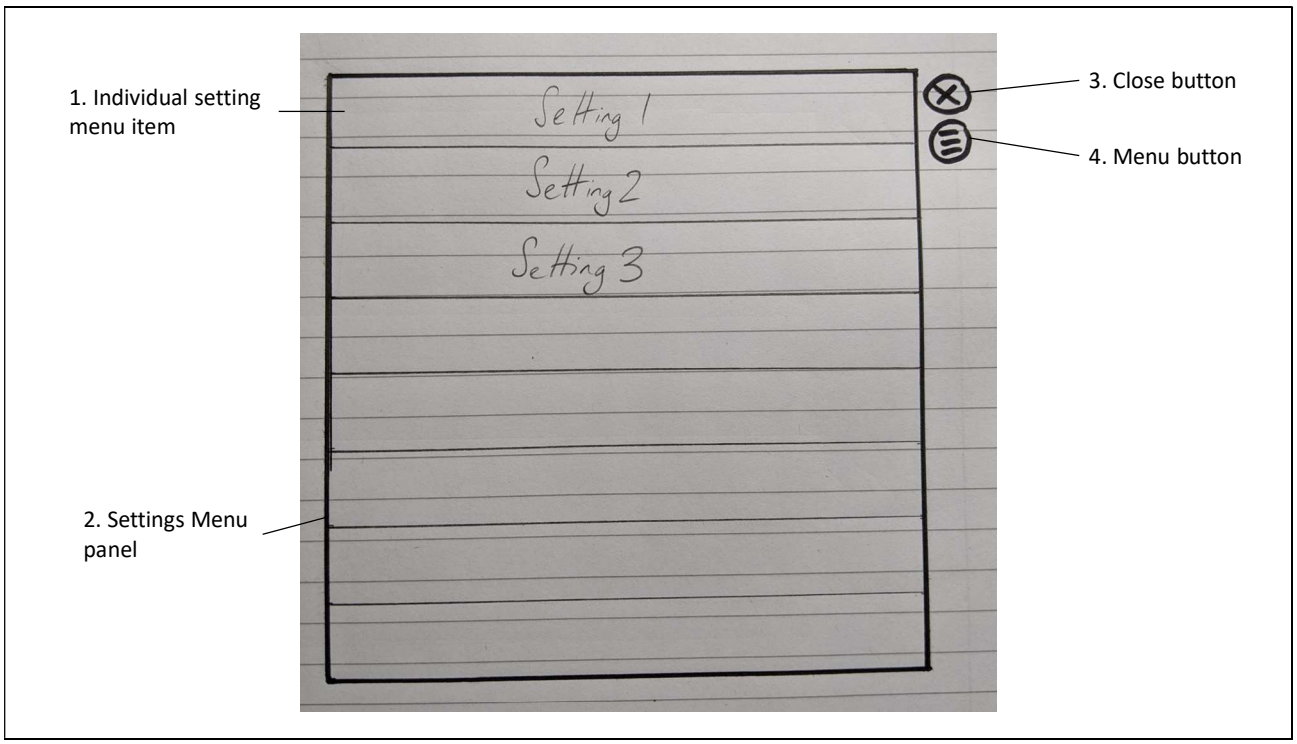


Figure 27. Settings Menu Design

There are two types of settings to include:

1. Settings that help the user's quality of life
2. Settings that change the learning experience

The first type of setting can make the app more comfortable to use, more accessible and be more convenient. Some setting ideas that would fit in this category are:

- Audio Volume
- Screen brightness
- Size / Colour of the fonts to aid readability
- Different colour palettes to help with visual impairments

It is important for this type of setting to be included as increasing the accessibility of the app, as well as providing convenience, both work towards enhancing the user experience. It can help the app become more useful and usable.

Audio volume and screen brightness are settings that also can add significant convenience. If MobileVR is used, when the app is in use, the smartphone device is in a headset, and so it is more difficult to use the phone's volume buttons or to change the screen brightness. Adjusting the volume is also useful for those with sensitive hearing and is generally convenient for use. This also adjusts audio separately from the phone's audio, and so can be adjusted separately from any other audio that is being played on the device. Adjusting the screen brightness can be useful for those with sensitive sight and could be adjusted separately from the device's brightness. The other two setting ideas can both be used to help increase the accessibility of the app.

In the second type of setting, the settings can be used to change how the user interacts and learns with the app. These can be used to challenge the user, and to implement different learning 'modes', which could be used to practice questions and to test the user. Some setting ideas that could fit in this category are:

- Audio On/Off
- Translations On/Off
- Audio Only Mode
- Guided Learning Mode

This type of setting is useful as it augments how the app is interacted with, providing a varied and more effective set of learning 'modes'.

The audio toggle allows the user to completely turn off all audio. This could be used for convenience, if the user cannot play audio at all in their current surrounding, or it could also be used as a

challenge for a learner who is skilled at speaking the language, but unfamiliar with writing or reading it.

The translations toggle can be used to increase the challenge for the user. Without the translations, the user won't be able to see what any question means and will have to understand themselves or must identify the items simply by what item they look like. This can lead to better-quality learning but is also more challenging.

The 'Audio Only' mode represents the idea of removing all text and completing all the activities using only the audio. This closely represents natural communication and provides aspects of 'experiential learning' and the input hypothesis, where the user is challenged by learning and practicing using only audio, and then review questions only using it as well.

The 'Guided Learning' mode represents the idea of 'teaching' the user the words by pointing them in the direction of the answer. This mode is meant to add a more guided approach. Rather than starting to show questions straight away, this mode starts by introducing each item first and indicating which item it is using an arrow.

To control each setting, there will be standard controls, like sliders and checkboxes, where appropriate. There will also be several buttons to press, for example to return to the menu scene. Several of these buttons will have purposes like to close the display, or to open the menu. These will also use standard icons – a red cross to close the display and 'the hamburger' for the menu. These designs can be seen in Figure 28.

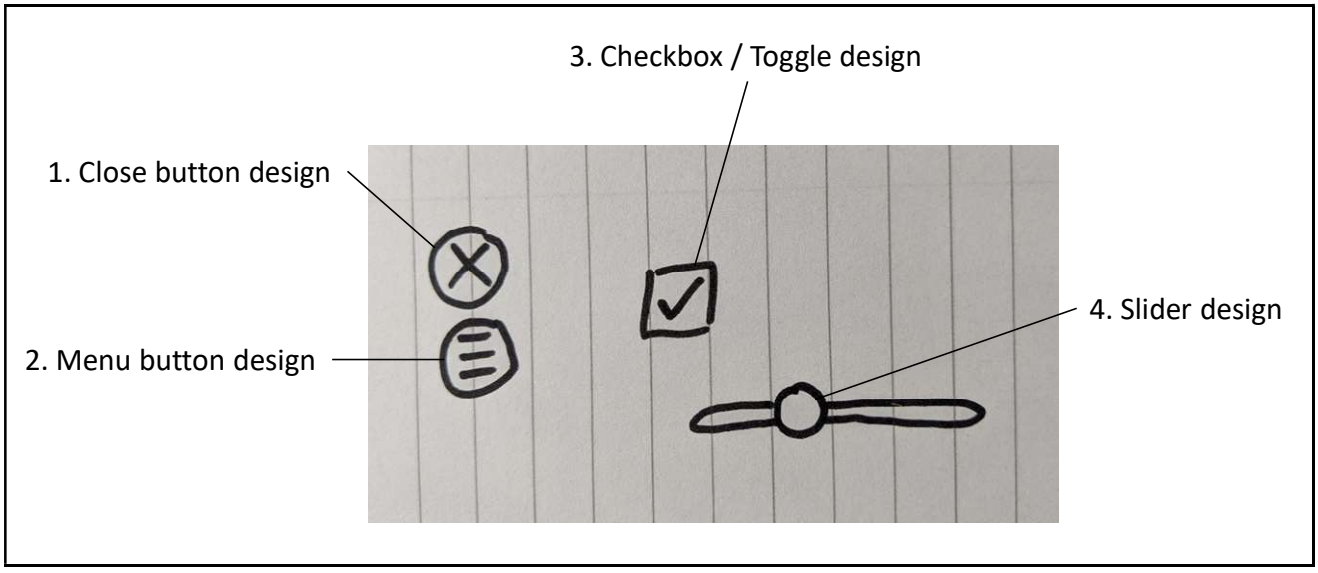


Figure 28. Icon Designs

3.5 Data Design

This section examines the data that the application will need to use and store.

This section separates this data into two categories:

- Data that is necessary for the application to work properly
- Additional Data for Analysis

3.5.1 Necessary Data

The application will need to store all assets that are used to make up the room and everything inside it, as well as all the code needed for the app to function. The application itself will need to be stored on the device.

For the app to support multiple different rooms and languages, it is important to store data about each room. This would involve data about the languages supported in each room and the type of room it is – e.g. 'Kitchen'. This would be used to direct the app in which scene to open.

For the app to function as intended, the application will need to store data about the items that can be interacted with. For each item, the application will need its foreign word, its translation in the native language, and the audio to be played. It will also be necessary to identify which room (and therefore which languages) each item is in. This would be used to help organise data.

The app will also need to store data about each question and its answers. For each question, it is important to have the foreign sentence, its translation, and the audio. It's also important to indicate which room it is intended for, as well as which type of question it is. It could also be useful to store which vocabulary item the question is meant to improve.

For the answers its foreign word will be needed, though it could also be useful to have its native translation and audio. For each question and answer it would also be important to indicate whether the answer was a correct one, and to identify which room the question is intended for. Of course, it is also important to store which conversation the answer is intended for.

3.5.2 Data Analysis

There is a requirement that data generated by this application can be analysed to gain future insight. Completing data analysis is a useful feature and can be used in several ways. It can be used to help the learner at an individual level. This analysis could be shown to the learner to show how they are progressing, and this could be used to motivate the user, and could perhaps inform some gamification features to help do this as well. This analysis could help to highlight any words or

questions the student is struggling with, or to show which settings could be recommended to challenge the user.

Data analysis could also be used to help the teacher, if the app was used in conjunction with a classroom setting. The teacher could view the analysis for their whole cohort, viewing which questions and items give the most trouble, as well as being able to review how individual students are doing. This could also show which settings are preferred overall.

Beyond helping to learn directly through informing the learner or the teacher, data analysis can be used to help the app designer improve the app to help streamline the app's learning process. The app designer could look at the order of interaction and the timescales of the interactions to see if any changes can be made to make this easier. The popularity and use of different settings could be used as well to see which settings need to be changed to make them more useful and usable.

Data analysis can also be used by researchers, as the app could be used as a tool and the data gathered could be used to test and compare different conditions and variables.

There's only limited data that can be generated by the application – mainly data that can be generated when the user interacts with anything. This data can be generated when the user interacts with an item, a character, or when answering a question. It can also be when any settings are changed, when a room is entered, or when the user moves.

To store the data needed for this analysis, this application will store every interaction. To do this it will store the time of the interaction, as well as what room it is happening in, which user's interaction it is and what type of interaction it is. For an interaction with an item, it will record which item was clicked. For an interaction when the user is answering a question, it will store the question's ID and if it was a correct answer. If it is when an interaction where the user changes a setting, it will store which setting was changed and what the setting what changed to. If it's an interaction where the user if entering or exiting the room, it will record which room is involved.

To do this, the application will also need to store some information about each user, so that the interaction can be linked to them.

3.5.3 Entity-Relationship (ER) Diagram Design

When drawing a simple ER diagram for this database, it was realised that many entities store the foreign text (L2 Text), its translation (L1 Text) and audio (L2 Audio), as well as which room that word was intended for. It was decided to bring this out and make a separate text entity that represents this. The ER diagram can be seen in Figure 29.

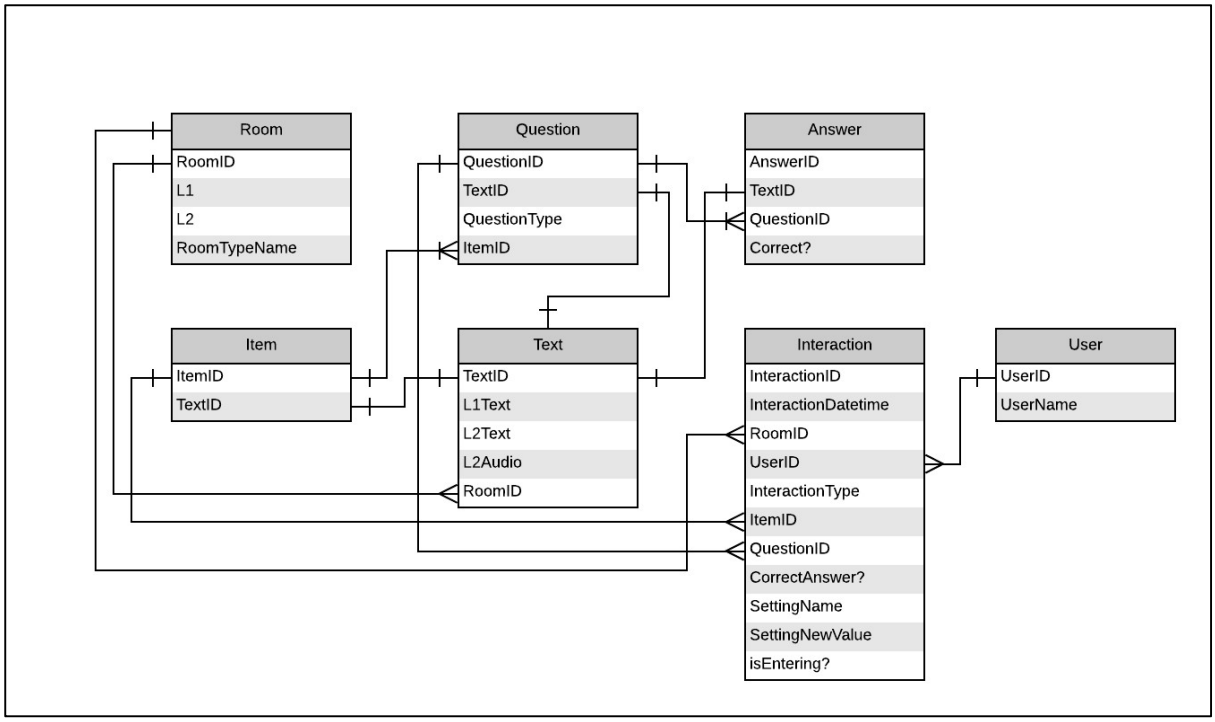


Figure 29. Initial ER Diagram

3.6 Feature Design

This section explores the different functionality and why it is needed and what it should do.

3.6.1 Movement

The user should be able to explore the room freely, which means being able to move around the room. There are several methods for moving around in VR:

1. Using a controller with a joystick
2. Teleporting

Other methods of VR also support physically moving in real space, and this is picked up and translated into the app. Given the potential technologies that are being used, this is not supported.

In the first option, a controller is used, where the joystick controls the movement. This allows for smooth movement. In the second option, the user clicks on the ground to move there. This is an immediate change.

Potentially both methods could be supported, allowing the user to choose their preference in the settings menu. Considering the VR technologies already specified, both options are feasible. When considering movement in VR, motion sickness and lag are elements to consider – though as this does not affect everyone, allowing the user to select their preference gives them the most control.

3.6.2 Character Question Chooser

When the user interacts with the character, the character will choose a question and display this to the user. How this question is chosen can be used to improve the quality of learning, by using techniques like spaced repetition.

When choosing a question, an algorithm can be used to select a question that the user hasn't seen in a while. The interval between a question being asked can be controlled to optimise the affect of spaced repetition and can greatly improve the quality of learning.

3.6.3 Text-to-Speech

In order to have audio simulate foreign text being 'read out loud', there are two options:

1. having each line read by an actor
2. using a text-to-speech computerised alternative

Recording an actor would take time and potentially entail some cost in the form of paying the actor, hiring equipment and any post-processing needed. Storing all voice clips could also potentially lead to a large amount of storage space needed.

Text-to-speech systems usually make use of advanced machine learning and other bleeding-edge research to make the system robust and flexible. Usually these systems are given text and audio is then outputted. This audio could be streamed when it is needed, rather than downloaded in bulk, which could pose issues to storage as audio can take up a lot of storage. Downloading in bulk could also take a good amount of time, dependent on internet connection, whereas streaming requires a smaller bandwidth. This streamed audio could then be saved locally, similar to a cache, so that if the user clicks the same item again, the audio will not need to be downloaded again.

The second option is superior. Text-to-speech offers an easier option with more flexibility in how it is implemented, and potentially provides a more efficient use of storage space. Text-to-speech systems are offered for very little cost, or even free, by large corporations like Google and Microsoft. These systems are very likely to have advanced research and datasets to draw from and use, increasing how well they perform.

3.7 Summary

This chapter began by exploring how the app is intended to be interacted with and how this affects and can contribute to the learning process. This chapter also examines multiple designs for the main environment, before concluding that an open-plan kitchen is the best design.

Different interface designs are then looked at, determining how each method of interaction is reflected visually. A basic plan for how each feature will function is laid out as well.

This chapter also looks at the data that will be stored and why, as well as investigating the idea of data analysis and how this data may be gathered using the designed functionality.

Lastly, more minor functionality is detailed, and some decisions are made on how best to accomplish them.

Chapter 4 – Implementation

This chapter details how the application was developed. It examines the tools used and the decisions made to implement each feature.

4.1 Tools

This section examines each tool used, describes what each tool does and why each tool was chosen.

4.1.1 VR Development Platform

Unity was chosen as the real-time 3D development platform to develop the VR application. Unity can be used for all manner of applications including architecture, film and games (Unity Technologies, 2020), though they also support development for VR applications.

Unity was chosen as it is free and has a large amount of supporting documentation. There is also an extensive 'asset store' where many useful assets could be used for free, to populate the environments without needing to do any amount of modelling. Unity also supports VR for different technologies, such as for smartphones, Oculus Rift and the HTC Vive.

4.1.2 Integrated Development Environment (IDE)

Alongside Unity, an IDE was also needed to write the scripts. Most of these scripts were in C#, which work directly with Unity's tools. Some PHP scripts were also needed for access to the online database.

For the IDE, Microsoft Visual Studio 2019 was chosen. This was the default option offered by Unity, , and it is a flexible, robust IDE. It also offers a variety of add-ons.

4.1.3 Version Control System

It's important to use version control systems in order to backup the application and to save previous versions. This enables me to revert to an earlier version if the application is fundamentally broken, and it also allows comparison between these versions, which can be used to figure out why it went wrong.

Github was chosen as the version control system. Github is free and has a good student program. It also has a useful desktop application to help automate most of the process.

4.1.4 Database Tools

A relational database was chosen as the method of storing information as this is the standard data-storage method and is the most widely used and documented method. It is also good to use for the app itself as it has some clearly defined relationships.

The university offers access to a mySQL database on the sci-project server. By default, this uses the phpMyAdmin set of tools to administer the database. The author has used phpMyAdmin as a management system before. Using this database does mean that the application needs to be connected to the university's network – either directly or using a VPN with correct credentials. This does limit how available the app is – due to these credentials being needed, only university students and staff can easily use it.

It was also chosen to include a second database within the application itself. This would be stored on the device locally. This database would be a lot faster to access and reduce the network latency dramatically. For this database, an SQLite3 database was chosen as this is supported natively within Unity and there is a lot of documentation to support this.

4.1.5 SSH Client

In order to interact with the sci-project database, PHP scripts were used on the server. These could be used to interact with the database, such as to extract information and then to send it back, or to insert new information.

In order to interact with the server and upload these PHP scripts, WinSCP was used, which is a free client allowing me to use SSH to interact with the server remotely.

4.1.6 VR Library (GoogleVR)

Within Unity, there were several different asset packs that support VR for different devices. GoogleVR was chosen, which caters entirely to VR on smartphones. The other option that was considered was the Oculus VR pack, which catered to Oculus technologies such as the Oculus Go, or the Oculus Rift.

It quickly became apparent that it would be difficult to use both packs at the same time. Initially this project aimed to cater the application to any platform, which would involve using both the Oculus and Google packs together. However, for each platform, this would involve configuring the virtual reality environment, and there was a risk of conflicts. This could have been solved by making two separate projects, one for each asset pack, but this would have been a bad use of time.

The end-product was also considered. Using the GoogleVR asset pack, a product would be developed for Android smartphones. The pack was intended to be used with the Google Cardboard or Google Daydream headsets. Using the Oculus pack, a product would be developed for any of the Oculus headsets. Android phones are more available and widespread than Oculus products. Oculus products aren't too expensive for VR headsets, though Google Cardboards are extremely inexpensive. A phone is usually less powerful and potentially has a worse screen than a headset though.

Another valid point was also that when using GoogleVR, the user has limited controls. There is a cursor in the middle of the screen, and the user can tap the screen to click it. This does limit the control and interaction possibilities.

After trying each asset pack out, the GoogleVR pack was easier to use. They both had thorough documentation to help learners, but GoogleVR's event-driven system was simple and effective.

This decision also means that the target hardware for the application has been narrowed down to only mobile phones.

4.2 Setup

This section examines the choices made during the setting up of the application, and how the foundation of the app was implemented.

4.2.1 The Room

A key component of the app is the virtual room that the user is in – this was determined to be the Kitchen scene designed in section 3.3. It's important that the scene is as realistic as possible – this would increase the user's immersion and presence, and this can improve memory recall. It would also help the user to recognise the various items in the room if they were realistic.

Figure 30. Example of the room's window and wall textures

The structure of the room – that is, the walls, floor and ceiling – was made using six planes. These constructed a cube and were adjusted to make plenty of room for all other assets. Applying several texture packs from Unity's asset store to the surfaces helped make them look less like solid blocks of colour and instead look more like walls and floors. These textures can be seen in Figure 30.

A premade door asset was used, however, there were no good prefabricated models for the window, and so one was designed using 3D cubes. Textures were also used to make the frame look wooden. In place of the glass, there were nearly transparent cubes so that the window is still interactable while the user can still see out. An asset was used for the surrounding skybox, which makes the 'view' out of the window change when the user moves, making it appear like the 'outside' and not a still image. The window can also be seen in Figure 30.

The lighting in the room is done using spot lighting. This can be done to emulate ceiling lights.

For the furniture in the room, there was a useful asset pack for realistic kitchen furniture which was used for all necessary items. These assets can be seen in Figure 31 and Figure 32.

Figure 31. The kitchen assets in the Kitchen Scene.

Figure 32. The table assets in the Kitchen Scene

The character asset can be seen in Figure 33 and was chosen for several reasons:

- The asset takes only 10MB to store, which is relatively low compared to other models. This is more advantageous for the phone where storage space is more limited.
- The asset is free.
- The asset looks more realistic when compared to many other characters who are more cartoony. Many models are also intended for use in video game development and this character asset has relatively more normal clothes when compared to soldiers and aliens.

Figure 33. The character asset chosen

The character shouldn't stay completely still, which may appear unrealistic and immersion-breaking, so an asset pack with basic animations was used. This allows the character to stand still, however some movement is involved to appear more realistic, for example shifting his weight from one foot to the other.

Figure 34. Comparison between design and implementation of the room

When the layout of this scene is compared to the design in Chapter 3, as seen in Figure 34, it can be seen that the layout is very similar. There is more room in the actual scene, to allow for the user's movement and any future additions to the room.

4.2.2 Databases

Implementation of the databases began with setting up the online database on the sci-project server. This database will be denoted as the 'web database'. Initially, the database design from Chapter 3 was implemented. Throughout development, there were a few additions and changes to improve the database, such as the addition of additional columns in the interaction table, which allow for more interactions to be recorded in a useful way.

Implementation of the local SQLite database was straightforward. This database will be denoted as the 'local database'. This database was necessary to implement. Through some unit testing during early development, when the app queried the web database, there would be noticeable lag while the app waited for a response. This is undesirable and so a local database allows for much faster retrieval of data.

There are some differences between the databases. The web database contains information about every room, whilst the local database only stores data about the room it is in. This avoids saturating the local database, as storage space can be limited on smartphone devices.

The local database also stores additional data about each conversation in order to judge when it next needs to be asked. This data is currently only stored locally as this data can be calculated using the interaction logs and this reduces the amount of data that needs to be downloaded to the local database from the web database.

When the user enters a room, all relevant data for that room will be downloaded from the web database all at once, using a PHP script. This is then inserted into the local database. As this data is straight from the web database, there are certain tables missing which would normally be needed for foreign keys. However, these tables are not downloaded to reduce the amount of data the app needs to download, because the data is correct, and the values can be looked up in the web database to see their meaning. The final database architectures can be seen in Appendix B.

4.2.3 Shared Class

Due to several shared features between different objects and scripts, for example database queries, text-to-speech audio and logging interactions, it was important to make sure there was a minimum amount of redundant, duplicated code. This is important because if changes had to be made to this code, the change would have to be made multiple times unnecessarily. This could also lead to some code being updated whilst others are forgotten, which could lead to further issues down the line.

A shared repository of functions is a solution to this. In this shared class, a number of public functions were written that could be used by any other script. This enabled a central place for functions that are used multiple times by different objects. This included functions for:

- Streaming the audio for the text-to-speech, which is used by both the Character and the Item interactable
- Downloading that audio as a pseudo-cache
- Logging interactions. This is used by many objects of the app

In addition to the shared class of functions, an empty GameObject was used to store some variables that could be accessed by any other object. This includes variables that persist throughout the session, such as the user's ID and the room ID, as well as the current settings. This was done as it is more efficient than repeatedly querying the database for this information. It's important to remain a different GameObject as some of the objects that would handle this information, for example the character's interface or the item textbox, are often inactive and are not accessible.

4.3 Features

This section examines each feature that was implemented and explores the implementation decisions made.

4.3.1 Menu Scene

The menu scene provides the first environment users are in when they start the application and allows the user to select which room to enter.

For this room itself, it was decided to use the demo scene from the GoogleVR asset pack, as it provided a simple space that wasn't too plain – with some cubes and other shapes, this space was better than a simple six planes put together.

For the actual menu, a UI appears in front of the user when they first start the app. This menu queries the online database and downloads the available rooms, which it then displays. This can be seen in Figure 36.

Figure 36. Example of the Menu Scene's display

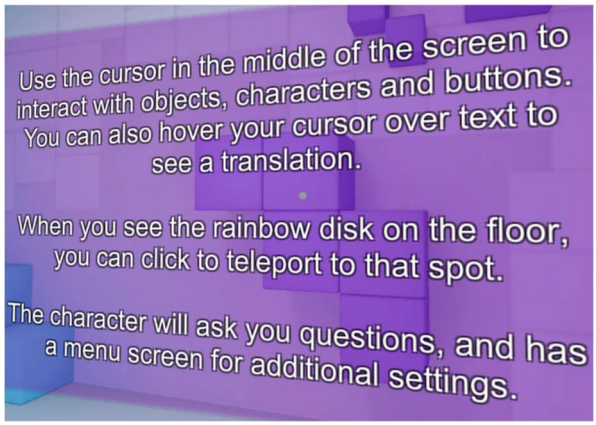


Figure 35. Example of the Menu Scene's tutorial

The menu scene also has a UI canvas appear to the right of the player. This canvas has a tutorial on it, in order to help provide some guidance as to how the app functions. This tutorial can be seen in Figure 35.

When the user selects a room to enter, the app will query the web database again and download all the data needed for the selected room. It then inserts this data into the local database, unless it already exists. Checking it already exists will avoid lots of redundant information. Writing to the local database will provide faster queries.

4.3.2 Item Interactable

When interacting with an item, a textbox should appear with the foreign word on it. There were several ways that this could be implemented within the app:

1. The textbox could appear above the selected item
2. The textbox could appear at a fixed point on the user's screen

Having the textbox appear above the selected item would allow for interaction with the textbox itself. An example of this can be found in Figure 37, which is a screenshot of a quick prototype to test these methods.

The idea behind the item appearing at a fixed point on the screen was that as the user looks around, the textbox follows them at a static point of their vision. This is like the Heads-Up Display (HUD) often seen in games to display information. An example of the 'HUD' method can be found in Figure 38, which is a screenshot of a quick prototype to test these methods.

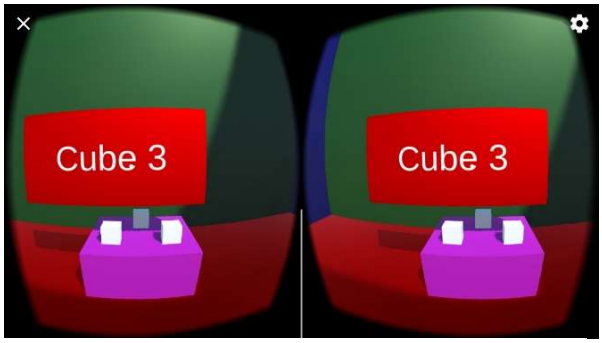


Figure 37. Textbox above the item prototype

Figure 38. Textbox on the HUD prototype

The second option was chosen – placing the textbox above the selected item – for several reasons.

The 'HUD' method does limit the user's ability to interact with the textbox. As the textbox moves with the player's vision, it would be impossible to point the cursor at the box. The user wouldn't be able to use a standard button to close the textbox and a timer could be used to close it instead. However, deciding a default length of time could be difficult, as the length could be unfairly short or too long. The length of this timer could be adjustable through the settings, but the complexity of this issue does impact how hard it would be to implement and would impact the user's ability to intuitively use this application.

There is also a requirement to show the translation of the text on command. If the textbox was above the object, then the user could simply hover the cursor over it, whilst using the 'HUD' method, it would not be as simple. This makes it complex not only to design, but also to use, as the user can't directly interact with the text.

In order to implement this, a textbox object was used with two text objects – one each for the foreign word and its translation. Initially the object is inactive, but when an item is clicked, the object becomes active, moves above the object and then is rotated to look at the user. The foreign word for that item is displayed, with the translation textbox remaining hidden.

When the cursor hovers over the textbox, the two text objects move apart, and the translation text is displayed, as in Figure 39. There is also a close button which can be used to close the textbox.

Figure 39. (Left) After initially clicking on the item (Right) After hovering cursor over textbox

4.3.3 Character Interactable

When the user interacts with the character, the character will ask questions.

To begin with, the character assets could not be directly interacted with, due to them not having any components that could be interacted with. Instead of fiddling with the prefabricated assets, a 3D cylinder was put around the character and turned it invisible. Even when invisible, this cylinder can be interacted with.

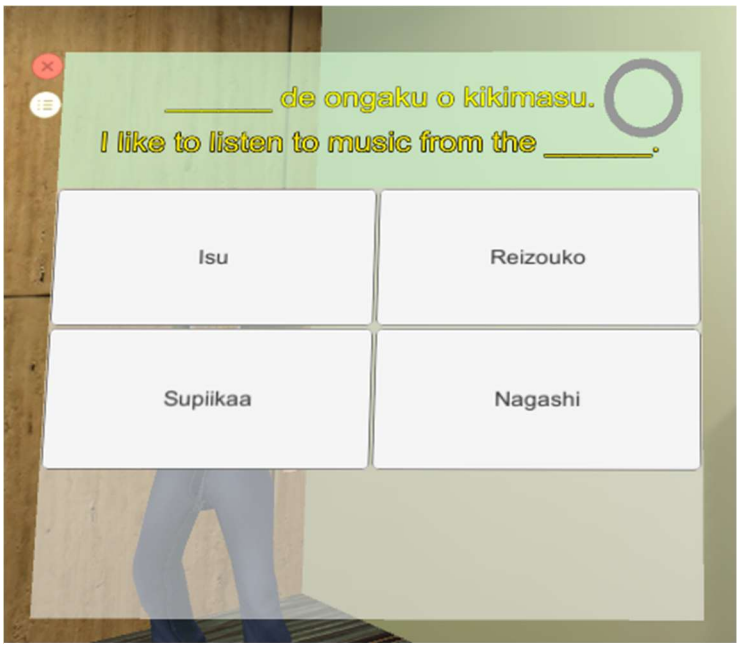


Figure 40. Example of Character Interface

When the character is clicked, this screen becomes active, as seen in Figure 40. The character then identifies which conversation to have, and then finds the appropriate text and translation text. This happens in a very similar way to the item textbox. The audio of the question is also played.

Also, in a similar way, when the cursor hovers over the question, the translation text is displayed, as seen in Figure 40.

There are currently three question types that can be asked, and each one functions slightly differently.

These question types are:

1. Introduction
2. Multiple Choice
3. Answer Blocks

The introduction question type is intended to introduce each item to the user. There are no answers for this type and is merely an example of seeing the word alongside the translation.

The multiple-choice question type is straight forward – a question is asked and then the user must select the correct answer out of the four options. The question in Figure 40 is an example of this question type. For this question type, the four options are taken from the database and placed

randomly, though the correct answer is still tracked. When the user selects one of the options, the program checks whether it is correct and then proceeds.

If the answer was correct, the question is replaced with ‘That is correct’, and a few seconds later the screen will automatically close. If the answer was incorrect, the question will be replaced with ‘That is incorrect’, and a few seconds later will be prompted by the correct answer through ‘The correct answer is X’. The screen will again close after another few seconds.

The Answer Blocks question type contains the building blocks of the correct word – one or two letter combinations that can be combined in the right order to make up the word. An example of this can be seen in Figure 41. For example, in Figure 41, the correct answer is SUPIIKAA, and each component needs to be inputted in the right order.

For this question type, six blocks are extracted from the database and placed on the answer panel at random. When the user clicks one of the answers that does not appear anywhere in the answer, the question is replaced with ‘That is incorrect’. The screen does not close, though, and the user may continue. The original question text is replaced after a few seconds.

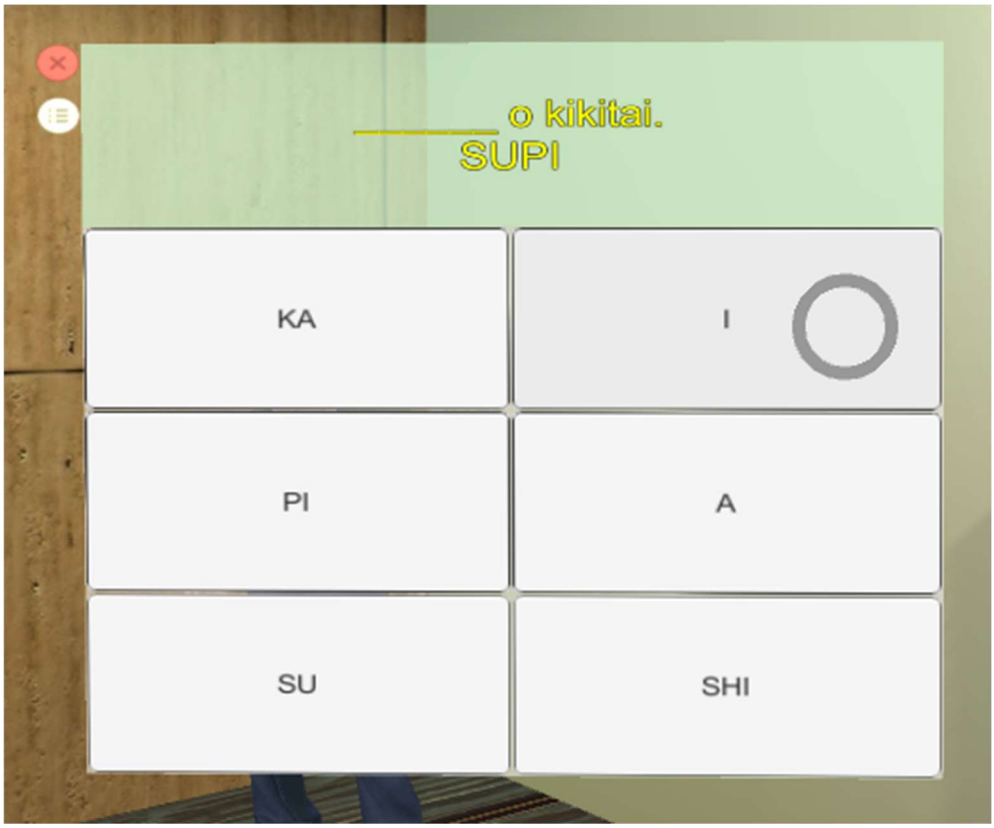


Figure 41. Example of the 'Answer Blocks' question type.

If the user clicks an answer that does appear in the full answer, there are a few things that must be done:

1. As blocks are being used to spell out the answer, first, the selected answer is added to any existing answers have already been received. In the case of Figure 41, if 'I' was selected next, this would be added to 'SUPI' to get the new answer of 'SUPII'. This string is called the 'new answer'.
2. This is compared to the full answer and checks that this new answer is a prefix of it. In other words, check that the full answer starts with the new answer. This is correct for the example.
    a. It is at this point that the user is told that their choice is correct. They are notified for a few seconds by changing the question text to 'That is correct.'
    b. If the new answer is not a prefix of the full answer, the user is notified that it is incorrect.
3. If the new answer has passed all checks to far, then the new answer is checked to see if it is equal to the full answer. If it is, then the user is shown 'That is the correct word' for a few seconds, and then the character's screen is closed.

4.3.4 Movement

There are several methods for moving around in VR:

1. Using a controller with a joystick
2. Teleporting

Seen as this project is targeting mobile phones as its intended hardware, the available controls are limited to a simple cursor. Therefore, it is difficult to use a controller or joystick to implement movement. However, the cursor could be used for smooth movement by clicking and dragging themselves along the floor. This would be difficult to indicate to the user intuitively though and would likely need to be communicated explicitly via a tutorial.

Teleporting is easier to implement, and it would involve just pointing and clicking on the floor and the user will teleport. This would also be easier to indicate to the user – when they look at the floor, a target could appear, indicating where they would move to when they teleport.

It was decided that teleporting would be the chosen method, due to it being easier to implement, intuitively understand and more suited to the hardware.

This was implemented using the Unity asset 'VR Teleporter', which provides some scripts and a few prefabricated models to use, such as the target to place on the floor and the 'master' object which is placed on the user. Using the supplied scripts, it was easy to implement the teleporting.

4.3.5 Text to Speech

When items are selected or the character interacted with, audio is played of the text displayed. Designing a text-to-speech system from scratch was a large undertaking as it would involve large amounts of data gathering and machine learning, and so it was decided that an existing library would be used for this, and an online API seemed ideal. An online API could be queried, and the desired audio would be sent back.

After looking for available technologies, several options were found:

- Microsoft Azure Speech Services
- Amazon Polly as part of Amazon Web Services
- IBM Watson Text to Speech

These options all offered different types of text-to-speech services.

Microsoft Azure was chosen because:

- They had a large selection of available voices, gathered from over 45 languages and locales (Microsoft Azure, 2020b). This would allow each room to have an appropriate voice for its language, making sure the intonation is more accurate.
- It sits on Microsoft's Azure platform and is built on the basis of bleeding-edge research (Microsoft Azure, 2020d).
- Using a free instance of the Azure Speech API, hundreds of thousands of characters could be processed per month. For development purposes this was more than enough (Microsoft Azure, 2020a).
- The Azure Speech module can be used in additional ways, such as to transcribe speech to text and for translation. This would allow for straightforward development of features such as voice recognition and auto-translation into the app in future.

The Azure documentation was used to implement the audio, which helpfully had a tab for how to code it to use within Unity (Microsoft Azure, 2020c). There were several different templates on how to implement it.

When the user interacts with anything that triggered audio, the program first checks to see if the audio has already been downloaded. This is the ideal scenario, as this means that there is no latency or lag from the network connections. If there is no downloaded file, two asynchronous tasks are

started. The first would stream the audio into an audio source component within Unity and then play that audio. The second task would begin downloading the same audio file locally so that if that audio is played in future, it is available and won't have to be re-streamed.

The voice is also changed according to which language and room the user is in – this involves writing a short line of SSML (Speech Synthesis Markup Language) and then send that to Azure's services to get a response. This allows for customisation of the spoken text as well – such as removing certain special characters like underlines, which would otherwise be read out.

4.3.6 Settings / Options

A settings menu was designed, offering the user the ability to change some aspects of the app. To implement this, a second screen was added to the character's interface. Below the close button there is a menu button. When this is pressed, the question and answer screen is made inactive and the menu screen is activated. This displays each setting and a standard way of changing it that the user should be familiar with, such as a toggle switch for on/off settings. This can be seen in Figure 42.

There are two types of settings that were designed:

1. Settings the help the user's quality of life
2. Settings that change the learning experience

In the first type of setting, a slider was implemented that controls the app's audio volume. Any time that an audio source is about to play, the app checks the audio volume setting and adjusts accordingly.

There are a few other setting ideas from the first type that were considered:

- Screen brightness
- Size / Colour of the fonts to aid readability
- Different colour palettes in the case of colour blindness

These ideas would improve the app, though due to time constraints these settings were not implemented. The second type of setting was prioritised as they tend to affect the learning of the app.

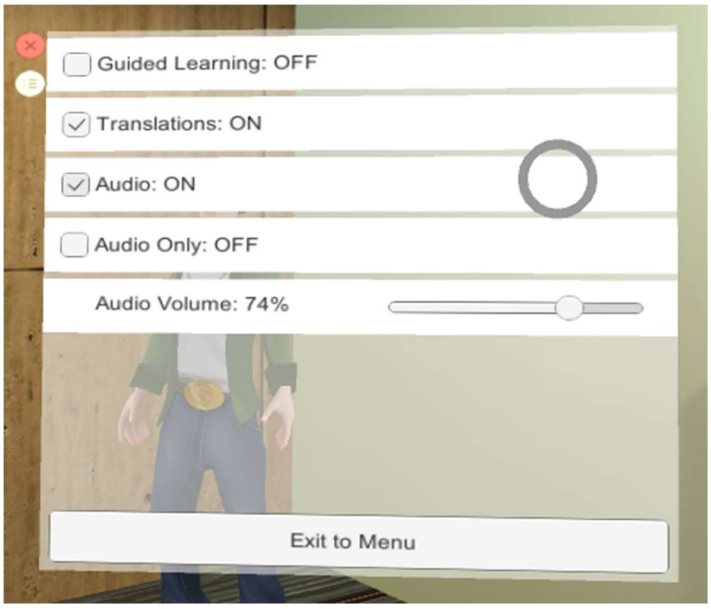


Figure 42. Example of Menu Screen.

In the second type of setting, four settings were implemented:

- Audio Toggle On/Off
- Translations Toggle On/Off
- Audio Only Mode On/Off
- Guided Learning Mode On/Off

These settings can be used to change how the user learns and uses the app. This could be done by removing certain helping features to challenge the user. These can also add features that help the user even more.

The audio toggle allows the user to completely turn off all audio. When an interaction happens, this setting is checked to see whether audio is needed. If the audio setting is OFF, then no audio will play.

The translations toggle is similar. When the cursor begins to hover over a textbox, this setting is checked. If translations are off, then the translation text will not be shown, as in Figure 43. This can be compared to Figure 39 as a comparison for when translations are on.

Figure 43. Example of the Translation setting being OFF

The 'Audio Only' mode represents the idea of removing all text and completing all the activities using only the audio. Whenever an interaction happens, this setting is checked before displaying any text. This setting also negates the audio toggle and turns it on.

The 'Guided Learning' mode represents the idea of 'teaching' the user the words by pointing them in the direction of the answer.

When guided learning is active, the character will ask questions of the 'introduction' question type 1, which starts by going through each item. Most conversations in question type 1 have no answers and are worded to merely show the user what the item is, for example 'This is a table'. When the question being asked is of question type 1, an arrow appears over the characters head and points to the appropriate object, as seen in Figure 44. The user then can interact with the object to see the foreign word and know the word in future. By default, the question is marked as correct.

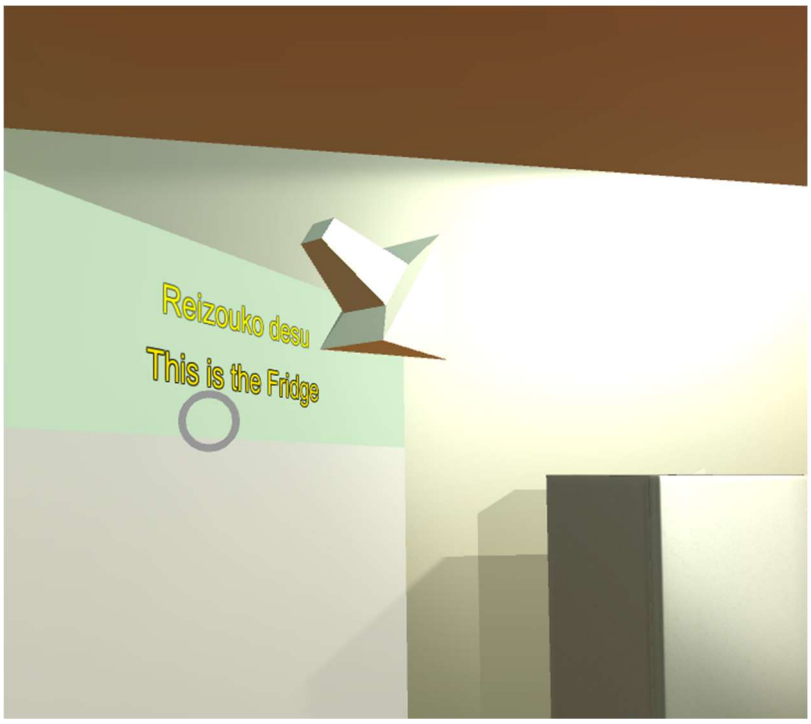


Figure 44. Example of Guided Learning mode

The user closes the character's UI when they are satisfied and can reopen it to continue, as normal.

4.3.7 Conversation Choosing Algorithm

When interacting with the character, the character will ask a question. The question that the character asks should be chosen to give the most benefit to learning. This is controlled using an algorithm that can implement and make use of spaced repetition, which will improve the memory retention of the user.

Several existing applications have written about the tools and algorithms that they use, and many of them use a modified version of the SM-2 algorithm, which is an algorithm originally developed for Super Memo and has since been made open-source (Anki, 2020; Mnemosyne Project, 2020).

The SM-2 algorithm makes some improvements on the Leitner system and figures out exactly when an item is likely to be forgotten, so that it can work out when it is ideal for an item to be reviewed. A review of these two systems by Godwin-Jones (2010) expresses the algorithm as a process that is 'built around the user's actions on reviewing an item'. The user scores themselves on 'how well or poorly the item was recalled' and these scores then impact when the system next schedules that item for review. According to Schimanke et al. (2018), SM-2 achieves the aim of spaced repetition learning, which is to look for the 'longest possible intervals that do not lead to forgetting'.

Based on this praise, and the fact that it is open-source and effective, the SM-2 algorithm was chosen to be implemented. Some modification ideas were taken from Anki's documention (2020). The algorithm is well-documented by Wozniak on the SuperMemo website (1998), and details how it works.

Each item is assigned two additional variables:

- $I(n)$ – interval between repetitions after the nth repetition (in days)
- EF – the 'Easiness Factor', which reflects the easiness of memorising and retaining a given item in memory

For $I(n)$, the following formula is applied:

$$I(1) = 1$$
$$I(2) = 6$$
$$for\ n > 2, \qquad I(n) \ = \ I(n-1) * EF$$

EF can vary between 1.1 for the most difficult items, and 2.5 for the easiest ones. By default, items initially have an easiness factor of 2.5. When a student reviews an item, the new value of EF, denoted as EF', is calculated by a function $EF' = f(EF, q)$. The variable q is the quality of the response, which is a number between 0 and 5.

For $0 \leq q \leq 2$, the response is judged to be incorrect, with different degrees of recall. For example, q = 2 may be 'it was on the tip of my tongue' whilst q = 0 would be 'I have no idea at all'.

For $3 \leq q \leq 5$, the response is judged to be correct, with different degrees of recall as well. As examples, q = 5 would be an instantaneous recall, and q = 3 would have taken some difficulty to recall.

The full equation to calculate EF' is as follows:

$$EF' = EF + (0.1 - (5 - q) * (0.08 + (5 - q) * 0.02))$$

Anki makes several modifications to this algorithm, some of which have been adopted by this application as well. For example, they allow the interval's initial steps to be customisable, rather than rigidly $I(1) = 1$ and $I(2) = 6$. Whilst the project's implementation isn't fully customisable, it does follow Anki's idea that 'it can be necessary to see a new card a number of times before you're able to memorise it', and so this project's implementation only uses $I(1) = 1$, with no other defined intervals. In future, it could be good to follow in Anki's footsteps and make this a fully customisable feature.

In order to implement this algorithm, a few adjustments had to be made to how the character's 'conversations' were handled. A string array was used for the conversation objects, which stored the L1 and L2 texts for the question, as well as the variables needed for the algorithm:

- Interval
- Easiness Factor
- Date that this item was last answered

There are up to 3 different question types that could be asked. These were kept separate, and so when the scene is opened, the app loads in all the conversation objects, separates them into their question types and then concatenates these into one large array. This is done so that all conversations with a question type of 1 will go first, and then all conversation of a question type of 2 and so on.

When the character is activated, a function is run that selects which conversation to ask. This function iterates in order through the conversation array and uses the interval and the 'date it was last answered' for each item, working out whether the conversation is due to be reviewed. Currently, this function will stop as soon as it finds a conversation that can be reviewed and chooses that one.

The user then answers the question. Regardless of whether the answer was correct, the UI will inform the user of how they did, and then close. If the user answers incorrectly, the easiness factor is updated using the value $q = 2$. This is done again following a modification done by Anki. They use a single 'fail' choice, rather than three. This is done because adjusting an item's EF can be done sufficiently by varying the positive q values. The conversation object is then moved to the end of the array, so that it may be reviewed again in the same session. The interval is not updated unless it was answered correctly.

The number of times a user gets a question wrong in one session is tracked for each conversation. When the user answers correctly, this tracked number is used to calculate the value of q. By default, q = 5, and for each incorrect answer, this decreases by 1, until it is equal to the minimum value of 3. When the user answers correctly, the easiness factor and interval are updated using this q value. The 'date that it was last answered' is also updated. By changing the interval, when selecting a new conversation, the correctly answered conversation will be skipped over, as its next review date is in the future.

4.3.8 Data Analysis

There is a requirement to gather data of how the user uses the application and analyse that data to generate useful information that can further impact the learning experience.

4.3.8.1 Data Gathering

In order to gather the data, the application logs every interaction in the internal and online databases. Currently, the types of interaction that are being logged are:

- Item interaction
- Character interaction
- Settings Changes
- Scene Changes

For all of these types, there are common variables that are logged in all of them:

- UserID
- RoomID
- Date and Time

Upon first opening the application, it will identify the user's UserID by querying the online database. This uses the device's MAC address, which is unique for each device. If the user's MAC address is not in the database, it will make a new UserID and save that.

These variables can identify the user, which room the user is in, and when they were in there. This allows for accurately distinguishing different users' logs from each other and can also be used in SQL queries to identify a certain user's individual session logs.

There is only one additional variable for item interaction, the ItemID of the clicked item. This logs when they click the item.

For character interactions, there are two additional variables that are logged:

- ConversationID
- Correct

These track which conversation was asked, and whether the user gave a correct answer. For the 'answer block' question type, this will appear as multiple correct answers for the same conversation, one for each block. This allows any incorrect block answers to be recorded.

When any settings are changed, the status of all the variables are logged, rather than just the one that was changed. This does not add much strain on the database and allows for an easier way to find the current settings values. When this happens, two additional variables are logged:

- SettingName
- SettingValue

The scene changes when the user enters a room or return to the menu. This is logged using one additional variable – 'isEntering'. This is 1 if the user is entering a room and 0 if the user is exiting the room. This can be used to track when the user enters and leaves the room.

4.3.8.2 Performing Analysis

To perform analysis on the data that is gathered, there are several methods that can be used. SQL is a query language that allows the user to interact directly with the database, and allows them to combine data in useful ways, as well as search for data.

Many tables in the database work using foreign key IDs, which are represented with integers. Whilst this leads to an efficient database, this does make it difficult to read and understand from a human perspective. As an example, Table 10 shows some records from the Conversation table.

Table 10. Example Raw Data from TblConversation

ConversationID	ConversationLevelID	TextID	CharacterID	ItemID
9	1	26	3	10
10	1	27	3	11
11	1	22	3	6

This is difficult to understand and is relatively low-level. SQL queries can be used to join with the relevant tables and extract useful, understandable information. For TblConversation, an example SQL query can be seen in Code Listing 1, and some of the results from this query can be seen in Table 11. As can be seen, this new table is far easier to understand and follow.

Code Listing 1. Example SQL query for TblConversation

```
SELECT ConversationID, TblConversationLevel.LevelName AS
ConversationLevelName, TblWord.L1Text AS L1Text, TblWord.L2Text AS
L2Text, TblItem.ItemName AS ItemName, TblRoom.RoomName AS RoomName
FROM TblConversation
JOIN TblConversationLevel ON TblConversation.WordLevel =
TblConversationLevel.ConversationLevelID
JOIN TblWord ON TblConversation.TextID = TblWord.WordID
JOIN TblItem ON TblConversation.ItemID = TblItem.ItemID
JOIN TblRoom ON TblWord.RoomID = TblRoom.RoomID
ORDER BY ConversationID ASC
```

Table 11. Example Processed Data from TblConversation using SQL

Conv ID	Conversation Level Name	L1Text	L2Text	Item Name	Room Name
9	Introduction	This is a speaker	Supiikaa desu	Speaker	Kitchen
10	Introduction	This is a window	Mado desu	Window	Kitchen
11	Introduction	This is the extractor fan	Kankisen desu	Extractor Fan	Kitchen

Similar SQL queries can be executed for other tables. The most useful of these for data analysis is likely the SQL query for Tblinteraction, which enables easy reading of the table, and can allow a teacher to peruse their student's interactions in full.

SQL can also be used to search the database for more specific details, for example a single user's interaction history with one question. An example query for this can be seen in Code Listing 2, with the results displayed in Table 12. This example examines the interaction history of User 1 on Conversation 22, and also eliminates some foreign keys as well, to make the information more meaningful.

Code Listing 2. Example SQL query to find interaction history for one user and one question

```
SELECT InteractionID,
CONCAT(TblRoom.RoomName, CONVERT(TblRoom.RoomID, char)) AS RoomName,
DateTime, TblWord.L1Text, Correct
FROM TblInteraction
JOIN TblRoom ON TblInteraction.RoomID = TblRoom.RoomID
JOIN TblConversation ON TblInteraction.ConversationID =
TblConversation.ConversationID
JOIN TblWord ON TblConversation.TextID = TblWord.WordID
WHERE UserID = 1
AND TblInteraction.ConversationID = 22
```

Table 12. Sample of Interaction Data for User 1 on Conversation 22

Interaction ID	Room Name	DateTime	L1Text	Correct
683	Kitchen1	2020-04-05 15:43:33	Where do you cook?	1
733	Kitchen1	2020-04-06 17:37:15	Where do you cook?	0
1312	Kitchen1	2020-04-08 13:53:58	Where do you cook?	0
1408	Kitchen1	2020-04-08 14:11:30	Where do you cook?	1

Using data like this, teachers would be able to look at specific user's performance, everyone's performance on a single question, or look at how often certain questions were answered incorrectly. Learners could also view their own data to review their performance.

SQL queries can be executed directly on the database to gain data, or they can be performed in PHP scripts, which can then be used to display the data in tables, charts, or graphs. This would be user-friendly and could be automated.

SQL is a useful tool for data analysis and does help to extract information from the data that the application gathers. In order to analyse the data, it does need the analyst to have knowledge of how to use SQL and so this is not accessible to all. Some pre-defined queries can be written and used, though – for example, a webpage can be written that will query the database for certain information, given some inputs. While this can allow for a certain amount of analysis, this could limit the user to only those pre-defined queries and does not allow them free perusal.

Excel, and other spreadsheet applications, can also be used for additional flexibility. The database can be exported out of phpMyAdmin and into a CSV (comma-separated values) file format. This can be opened in Excel, and then copied into a spreadsheet, which allows the user to use Excel to process the data using a variety of methods.

Table 13. Example Processed Data from TblConversation using Excel's VLOOKUP function

Conversation ID	Word Level	L1 Text	L2 Text	Item Name	Room ID
9	1	This is a speaker	Supiikaa desu	Speaker	1
10	1	This is a window	Mado desu	Window	1
11	1	This is the extractor fan	Kankisen desu	Extractor Fan	1

When the data is in Excel, much of the same operations can be performed on it as with SQL. As outlined above, SQL is useful for reorganising and searching data. Within Excel, the database tables can be reorganised into readable data using the VLOOKUP command. This command will return a value from a table, given the ID. As an example, the command can look at the TextID value in TblConversation and search for that ID in TblWord to return the L1Text. This can be used multiple times to reconstruct data in the same way as the SQL query in Code Listing 1. An example output when this is used can be seen in Table 13.

Within Excel, it is also possible to filter a set of data. This allows the user to sort the data according to each column, as well as hide records with certain values. For example, in Figure 45, the user can uncheck the records that they want to hide. If they only wanted to see records about the Speaker and the Window, they would only select those two options.

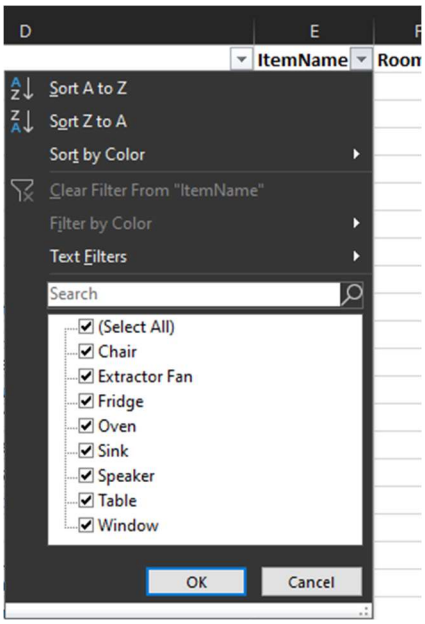


Figure 45. Example of Excel's Filter feature

This filter feature can be used in many ways. If the interaction table has been made readable, this can then be filtered to give details about a specific set of users, or a specific set of questions. This filter feature could be used in a similar way to a SQL query to find all details of a user on a specific question. As an example of how this can be used, Figure 46 shows part of the interaction table when the UserID and ConversationID columns are filtered to 1 and 22 respectively.

There are also Excel functions such as the IF function, which runs a logical test and assigns a pre-defined value for if that test is true or false. This function allows for a more stringent test than filtering, and this can be used in interesting ways.

	A	B	D	E	F	G	H	I	J	K
1	Interaction ID	Interaction Type ID	Conversation ID	L1Text	L2Text	Conversation Type	room ID	datetime	user ID	correct
563	562	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	02/04/2020 14:47:39.0	1	1
573	572	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	02/04/2020 14:53:28.0	1	1
583	582	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	02/04/2020 14:54:47.0	1	0
684	683	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	05/04/2020 15:43:33.0	1	1
734	733	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	06/04/2020 17:37:15.0	1	0
1313	1312	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	08/04/2020 13:53:58.0	1	0
1409	1408	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	08/04/2020 14:11:30.0	1	1
1498	1497	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	08/04/2020 16:21:08.0	1	1
1641	1640	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	08/04/2020 16:30:18.0	1	1
1727	1726	2	22	Where do you cook?	Ryouri o suru no doko desu ka?	2	1	11/04/2020 16:26:27.0	1	1

Figure 46. Example of filtered Interaction table for user 1 and conversation 22 in Excel

Within Excel, there is also a programming language that can be used called Visual Basic for Applications (VBA). This can be used to run basic algorithms across the data. As an example of how

this can be used, interaction data of how User 1 performed on questions when using the default settings and when using the 'audio only' setting was separated using the IF function.

In order to measure the difference, the 'Easiness Factor' (EF) that is generated using the SM-2 spaced repetition algorithm was used, which judges how easy a word is for an individual after some repetitions. Therefore, a VBA function was executed that ran through the interaction data and worked out the Intervals and EFs for each question and user. This code can be seen in Appendix C.

This code began at the beginning with the initial parameter values and went through the interaction data. For each interaction, it updated the parameters. At the end, it outputted the final values, which indicate the easiness factor of each question for User 1 – the results can be seen in Table 14. Using the right-hand 'difference' column and some conditional formatting, it can be seen that rows highlighted in blue are rows where AudioOnly's EF is higher – which indicates that the questions are easier in this mode. Rows highlighted in red indicate using default settings made the question easier.

Table 14. Comparing User 1's Easiness Factors for all questions when using two setting configurations (VBA)

ConversationID	Default Final Easiness Factor	AudioOnly Final Easiness Factor	Difference
17	2.5	2.5	0
18	1.94	1.782	0.158
19	2.184	2.5	-0.316
20	2.5	2.5	0
21	2.5	2.5	0
22	2.176	2.5	-0.324
23	2.5	2.46	0.04
24	2.5	1.71	0.79
25	2.46	2.5	-0.04
26	2.32	2.5	-0.18
27	2.5	2.5	0
28	2.5	2.5	0
29	2.5	2.164	0.336
30	2.5	2.5	0
31	2.5	1.1	1.4

This code can also be used to compare users, by comparing how they perform at the same questions. For example, if the inputted code has data from all users, then data about each user's easiness factors will be generated, which can then be compared. Using data gathered through a few participants, this can be seen in practice in Table 15. This table shows that Users 3 and 4 find Conversation 19 easier than Users 1 and 2.

Table 15. Processed Data showing all user's Easiness Factors with Conversation 19

UserID	ConversationID	Final EF
1	19	2.184
2	19	2.442
3	19	2.5
4	19	2.5

These two applications of the code are a proof of concept, that shows that the data can be analysed this data and different conditions can be measured. Data from two different setting configurations can be analysed to show differences, and data from different users can be analysed to show empirical differences. This can be extrapolated further, as with different VBA code, SQL or Excel functions, a good deal more could potentially be measured and compared and so there is a large range of potential findings.

These tools and systems can be used to do data analysis that can benefit the user's learning. Through giving them data and graphs of their performance and progress, it is possible to motivate them, as well as provide a way for them to review and evaluate their performance, which is a key part of experiential learning. These tools can find different data about a user's performance, for example how they have performed on one question, as well as their record with all questions and items. It can also display which questions the user finds easier than others.

If the user is part of a class, and there is a teacher overseeing a cohort of students, then these data analysis tools can also help them improve. Applied Behaviour Analysis has been used widely by teachers to gather data on how a class is doing and apply that to improving the quality of teaching. These tools allow the teacher to do this as well, as they can view data about a whole cohort. They can review an individual's record to see questions that individual is struggling with or judge how well the individual is doing. This could lead to additional support in the classroom. Alternatively, the teacher can see how the whole cohort performed on each question, to see if some questions are harder than others, or to see if the class performed as expected. This could impact the teaching

environment. There are other possibilities for reviewing individual items, and questions, as well as students.

The analysis can also help the app developer to improve the learning process. If, for instance, they compared data from users using the 'Guided Learning' setting to those who did not, and this showed that using the guided learning did not have an improvement on rates of success, they could conclude that the Guided Learning setting was not working as intended. The setting is meant to improve learning and introduce each item one at a time, explicitly teaching each word to the user. If this had no effect on learning, this could lead to changes in the learning process and improve it. Data Analysis could also help the developer through gathering data about different words, rooms, questions, interactions and settings, and these could be analysed to provide information on how to improve not only the learning process, but the other elements of the app as well.

The data analysis can also be potentially used by researchers. The app could be used as a tool in a research experiment, and this data and the subsequent analysis performed on it, could be a method of measuring a number of variables. VBA can be used to work out the Easiness Factor to compare word difficulty, or different questions' rates of success could be found. Different settings can be compared as well as different languages or different contexts. Some of the analysis performed in this section serves as a proof of concept that different variables can be measured and directly compared.

4.4 Summary

This chapter outlines the tools used throughout the implementation of this project, and why these were chosen. It also details how the fundamental elements of the app were set up and examines each feature of the app. It explores different decisions made and details the final product.

The basic functionality of the app – the menu scene, movement and item and character interactions – were completed with priority, making sure that they worked well. After these worked well, additional features were added – text-to-speech functionality, settings, data gathering and the conversation choosing algorithm were all added and tested sequentially.

Data Analysis was completed after everything else, as some usage data needed to be gathered for analysis to be performed. There are aspects of data analysis that can be improved and built upon, for example a web page could be made to show results from the data analysis to the learners and teachers, or the data could be fed back into the app. However, the analysis section does show that there are tools that make use of the data gathered in interesting and useful ways.

There are many more possibilities that can be added to this project. Once one aspect was finished, there would additional ideas and the project would grow. A deadline was set, to finish development by the start of April, to stop the scope growing uncontrollably. Some features were designed that were not implemented and this is due to time constraints. The development of other features were prioritised and these other features, like having a deck of words to perform a more traditional flashcard practice with or having a final test at the door of each room, were not completed.

Chapter 5 – Testing

This chapter tests each feature to ensure that they function as designed and that there are no straightforward ways to break that function.

Unit and component testing was completed alongside development to make sure that elements of the app worked. This testing was performed using Unity's in-built editor.

5.1 Functionality Testing

This section examines each feature designed in Chapter 3 and runs tests to ensure that they work as intended and that they do not break if used in a reasonable manner. This testing was performed using a mobile phone. This section broke down designed functionality into component parts and tested it. These tests can be found in Appendix D.

5.1.1 Summary

Item interaction works as intended. The user can click on an item and a textbox will appear with the foreign word and will play audio. The user can hover over the textbox to see a translation – all tests passed successfully. There was one comment about the audio, as when the audio is streamed there is some lag. This lag locks the screen, so the user can still move their head, but this movement won't be considered until after the lag. This can break immersion and potentially cause motion sickness. A solution would be to attempt to run this function asynchronously or download the sound file.

Character interaction also works as intended, with all tests passing successfully. There were a few highlighted problems that didn't cause a failed test but could be potential problems in future. There is also lag when audio is played, which can disorient the user. When interacting with the character, the audio that starts playing may be the audio from the question before – this does cut off after a second or two, at which point the correct audio does play. This is due to several processes occurring at the same time and this not necessarily being considered, leading to the audio being played before the correct audio clip has been loaded. To fix this, a proper process diagram could be drawn to make sure that audio isn't set to play before the audio clip is ready.

There is also a potential annoyance if the user wants to play the character's audio again. In the final deliverable, the user must click the character again. This is awkward due to the UI interface in the way and would be improved with a dedicated button. This is due to a lack of planning.

Questions in both question types are tested and passed all their tests successfully. They are displayed correctly and when the user gives an answer the correct actions occur. The algorithm is

tested as well and this was also completed successfully, reviewing items again after an interval and reviewing it more immediately if the question was answered incorrectly. Movement via teleportation also worked flawlessly.

The settings menu was also tested – first to check that the menu worked successfully, but also to check that each individual setting worked as intended. These all completed successfully, but combinations of settings weren't tested to see if they work with each other.

The menu scene was also tested and passed all its tests successfully – available rooms are downloaded successfully and displayed to the user as designed. At the time of the test, there were two rooms available – both lead to the Kitchen Scene, one for Japanese, and one for French.

The text-to-speech functionality also successfully passed its tests, though there was a comment made that some pronunciations are inaccurate for the Japanese room. This could be because the text-to-speech system is using roman characters rather than the proper Japanese characters. Largely, the voice is appropriate, and most pronunciations are fine.

Some tests were also done that test the network and how the app behaves if connection is disrupted. These highlighted some problems and gaps and saw some failed tests. One test was to check that all necessary data is downloaded when connected to the university network directly by being on campus, however this couldn't be tested due to the lockdown. The equivalent test that made use of a VPN was successful.

Some tests were also run to test how the app behaves when connection is lost – these failed. When offline, movement works as normal, but item or character interaction freezes the app. This requires a restart. This is because the request for audio goes unanswered and the app hangs. This could be solved with some code to detect a timeout, or to detect the device's connection status.

Some tests were also done to check that data needed for data analysis is successfully gathered, and these all completed successfully.

5.2 Requirements Testing

This section examines each requirement that was set out in Chapter 2, and explains how each requirement has been met, using the tested functionality above. The tests that were completed can be found in Appendix D, and the appropriate tests are specified throughout this section.

5.2.1 Minimum Requirements

- **Users can learn, practice and self-assess new words in a foreign language.**

As designed in Chapter 3, the user can learn by freely moving around the environment and interacting with the items and the character. To learn by talking to the character, the user answer the normal multiple choice and answer block-style questions and by using the 'guided learning' mode. This is shown in Tests 7-22 and Tests 41-42.

The user can practice by talking with the character. The conversations provided here are cycled and repeated according to an intelligent spaced-repetition algorithm that optimises memory retention. This also shows the words in different contexts and different ways. This follows the design in Chapter 3. This is shown in Tests 7-24.

The user is able to use different settings and different 'modes' to modify the interaction with the character. This makes it more difficult and helps to challenge the user. The user can use this to test themselves. The user can also use the data analysis to view their performance, which they can use to assess themselves. This follows the design in Chapter 3. This is shown in Tests 28-43 and Tests 56-59.

In order to properly gauge if the app is an effective learning tool, the Language Centre at Loughborough University was contacted and asked their opinions. A demo video was created and sent to accompany this request. Among the responses, the work was described as "fantastic", and several elements were praised – most specifically the 'answer blocks' question type and the Translations ON/OFF setting. These were described as "great challenges for the learners".

Some improvements were suggested as well, for example adding elements that practice speaking. Another suggested improvement was to encourage more problem-solving skills by not presenting the correct answer to the user if they get a question wrong. This would encourage the user to explore and find the relevant object themselves. This feedback has been very useful and whilst it does concur that there are effective learning techniques and elements to the app, there are also some noted flaws. A full transcript of these emails can be seen in Appendix E.

- **Users can interact with items and characters within an environment.**

All testing performed on items passed, and the item interaction works as designed. There is some lag when downloading and playing audio, but it is perfectly functional. This is shown in Tests 1-6.

Interacting with the character allows the user to engage in text-based questions that present the question in different ways and allows for different types of questions. The functionality design in chapter 3 is followed, specifically the different ways of presenting each question to the user (foreign text, audio and the native translation), as well as the different question types (multiple choice and answer blocks). There is also some lag here when downloading and playing audio, but it is still perfectly functional. These are shown in Tests 7-14.

- **Users can freely move around the environment.**

Users can freely move around the environment, as shown in Tests 25-27.

5.2.2 Core Requirements

- **Users can choose from multiple languages.**

Multiple rooms and languages are supported. There are multiple rooms with multiple languages, as shown in Test 46.

- **When interacting with characters, there will be several methods used to present information to the user.**

There are multiple ways that questions are presented to the user, as seen in Tests 7-14. These are the foreign text, the audio clip and the native translation. These tests demonstrate these as functional and effective.

- **When interacting with characters, there will be several methods for the user to respond to the character.**

There are several methods that the user can use to respond to the character's questions. These are, as designed in chapter 3, the multiple choice and answer block question type. Tests 15-18 show that multiple choice is an effective and functional method, and Tests 19-22 show the same for the answer block question type.

5.2.3 Extended Requirements

- **Users can manually edit their own settings to optimise the experience for themselves, such as customise the size/appearance of subtitles as well as audio settings, along with others.**

Tests 28-43 demonstrate that the settings menu and all the individual settings themselves are functional and work well.

- **Data gathered from the application can be analysed to provide useful insight.**

Data is gathered as shown in Tests 56-59. This data can be analysed as shown in Chapter 4.

- **The user can build a dictionary of terms they've learned.**
- **The user can progress from room to room by completing a test by the door.**
- **Users can manually change the language to customise to any language for themselves or for other students.**

These requirements were not completed due to a lack of time. As these were extended requirements, they were not high priority. The other extended requirements that have been implemented were priorities for their potential effect on learning.

5.3 Comparison to Existing Apps

In Chapter 2, a scoring system is used to judge how well existing apps perform for certain criteria. In order to properly judge if the final product is an effective and useful tool,

Table 16 shows the applications scores. These could be biased, but reasons are given in an attempt to make this less subjective. A comparison to the other existing apps is shown in Table 17.

Table 16. This Project's Application Scoring Table

Criteria	Score /10	Comments
Usability	7/10	The UI is straightforward, and while it is simple, it is straightforward and uses standard icons. The tutorial in the menu scene allows the user to familiarise themselves with the controls and the methods of interaction, and these become intuitive when the user is familiar with a Google Cardboard and how it is commonly controlled. However, it is basic and some elements are not clearly signposted. Some users would not know to talk to the character without prompting, for example.
Topics	3/10	Currently, the app has an extremely limited number of topics available, limited to a single kitchen, covered in two languages. This covers 8 items, each with at least 3 different questions related to them.
User Control	7/10	The user can select which room to access and interact with the items that they want to interact with. Users can change settings to adapt how they learn, and they can control how challenging their learning experience is. Users can also turn on the different modes to control what their experience is like. The questions are chosen automatically, and unfortunately the users have no control over this.
Customisation Options	6/10	There are customisation options available to the user. Volume can be changed to their preference, and the other settings can be used to customise the experience as well. Unfortunately, the user doesn't have the ability to control everything, such as the spaced repetition algorithm, and so these customisation options are limited.
Effective Learning Tool	7/10	The app makes use of a spaced repetition algorithm, which helps to build memory retention. The app uses techniques from experiential and problem-based learning, as well as showing different aspects of each word. The questions asked by the character are posed in different ways, and the user must respond in different ways, which means the user must become familiar with the word in order to correctly respond. This app has been praised by language teachers, though improvement have been made to improve upon it.

Table 17. Comparison of this project's application against existing apps

Criteria	Anki	Memrise	Duolingo	Mondly	Busuu	House of Languages	This application
Usability	8/10	9/10	9/10	8/10	8/10	8/10	7/10
Topics	9/10	9/10	8/10	6/10	6/10	6/10	3/10
User Control	9/10	9/10	4/10	3/10	6/10	4/10	7/10
Customisation Options	9/10	8/10	4/10				6/10
Effective Learning Tool	9/10	9/10	9/10	6/10	7/10	7/10	7/10

These scores, when compared with existing apps, do show that the app is flawed in some ways and shows that these other apps may still be better. Whilst these scores do suggest that the app is subpar, it doesn't consider that the app is attempting to do something different. It attempts to bring together many good aspects from other apps and bring them together to make a really good app in VR. Instead, this may have been too much to do within the scope of this project, and so the good aspects of each of the existing apps – Anki's customisation, Memrise's different modes, Duolingo's different question types, etc – may have been dulled because too much was being attempted. Each of these good aspects can be put together in an effective way, however in this app, they may not have been implemented with as much detail and care as needed.

There is potential to improve these scores. It scores low in the Topics criterion due to a limited amount of content. With more content, this would improve, and if the ability for users to make their own content, this would improve even more. With additional customisation and control settings, those scores would increase too. Usability could be improved with more guidance, or with slightly more stylised interfaces. Improving the effectiveness of the learning is doable by adapting some of the learning processes and giving the user motivation to return regularly – currently there is no incentive for the user to regularly use the app.

5.4 Conclusion

This chapter systematically went through the designed functionality that was proposed in Chapter 3 and broke them down into small elements that were manageable and measurable to test. It then tested these using the intended final device, a smartphone with a Google Cardboard headset.

Overall, the vast majority of these tests passed, indicating that the application fulfils the intended and designed functionality.

Whilst these tests didn't fail at performing their intended functionality, there were some tests which did not function exactly as planned. Several of these introduced lag, which is a minor problem that affects the user and could frustrate them or, at worst, induce motion sickness. This is because the lag would freeze the app's output for a few moments, which would no longer react to head movements. Other tests would work as intended but introduce some other unwanted behaviour. Test 11 shows that the wrong audio would play momentarily, before cutting to the correct one. This test passes as the correct audio does always play, however this test does suggest that there was a failure to properly plan the process, and this issue is a result of that. Test 14 shows that some functionality could be made better with superior design, where a dedicated button for requesting to hear the audio again would be easier and more convenient than the current, awkward method of clicking the character whilst avoiding clicking any of the answer buttons.

There were some tests that failed, and these were largely issues with error handling. The application will freeze if any interaction is attempted without a connection to the internet and to the university network, which is an issue that could have been solved with some limited error handling.

There were some elements that were designed in Chapter 3, but not implemented, such as a standalone flashcard deck, or the user's ability to manually edit and modify a room's layout or language easily. These were not implemented due to a lack of time and were left out of the testing as a result.

This chapter then examined each requirement as set out in Chapter 2, and details how each requirement was met or not.

The minimum requirements are met with ease. These were requirements that were the highest priority. The basic functionality of the app, being movement and interacting with items and characters, is met and all tests pass. One requirement is that the user should be able to learn, practice and assess themselves using the app. A variety of methods proposed in Chapter 3 translates this to into functionality. The tests that assess this functionality all pass as well.

The core requirements were also passed. These requirements focussed on making the character's interactions more varied, presenting the question to the user in several ways, and accepting answers in different ways. These requirements were met, given the successful tests on the different question types and on the different elements of a question – text, audio and translation. There was also a core requirement about supporting more than one language, and this was met. In the final deliverable, there are two available rooms, one that supports Japanese and one that supports French.

The extended requirements were essentially lowest priority, as they were intended to only be implemented after all other requirements. As a result, not all of these requirements were met due to time constraints. However, a couple of these requirements were implemented and have successfully passed their tests. There is a complete settings menu in the application that can be used to modify how different interactions work, and data is gathered through interactions, which can then be analysed for insight.

In summary, all of the implemented functionality works successfully to a high standard, with one or two issues regarding error handling and process planning. All of the minimum and core requirements are met, with several of the extended requirements being met as well. This indicates that the app meets the core vision and design of how it would work and includes several additions that greatly improve the flexibility and impact the app can have.

Some of the other uncompleted requirements could have been useful to implement. For example, the extended requirement of crafting a tool that allows users the ability to manually edit the language within a room would have worked well alongside the core requirement of supporting multiple languages. This functionality would allow the app to have a much larger scope, because in addition to an app that is effective at teaching vocabulary, it would be customisable to many languages and so would be much more useful and flexible.

Chapter 6 – Conclusion and Future Work

This chapter draws conclusions based on the final product and the testing of it and evaluates this final product and the project as a whole. It examines any issues or limitations of these and proposes future work that can be done to solve these and improve upon what has already been done.

6.1 Application

Throughout this project, a VR application has been designed and implemented to enhance and improve vocabulary learning. The application's implementation relied upon the design, which relied on the requirements, which in turn drew heavily from research and existing language learning tools.

This project completed a literature review, which heavily investigated techniques and research that could be used to effectively improve language learning and vocabulary memorisation. This review also explored ways computers and specifically virtual reality could be used to improve the quality of learning as well. The project also conducted a gap analysis where existing tools with similar ideas were analysed and useful lessons were drawn that could inform my design.

This research was instrumental in defining a set of achievable and effective requirements. The requirements, gap analysis and literature review helped to form a foundation which helped to inform decisions made throughout the rest of the project. To design the app, the requirements were taken and, using the research, plans were drawn up for the functionality and visuals of the app.

These plans for how the app was going to look and act were used alongside a set of tools that would be used to implement the app. A number of libraries and existing tools were used by the app to implement the more advanced features like Microsoft Azure's text-to-speech library and GoogleVR. Other than this, each functionality was made using Unity and C# scripts.

The final product was tested against the designs set out for it and the requirements that defined it. Most of the tests passed, except for a few that failed. The issues that were identified were a lack of error handling when there was network issues and noticeable network latency when streaming audio from the internet. There was also some issues that arose through several processes interacting with the same audio file, sometimes in the wrong order leading to incorrect behaviour.

After the completion of this process, the final application is one that fulfils all of the defined minimum and core requirements, showing that it is a product that makes use of effective vocabulary memorisation techniques. The product also met some of its extended requirements, which extended the scope of the product to be more flexible and more widely effective, by adding more learning

possibilities using a variety of settings that the user can control, and by gathering data that ca be analysed for widespread insight.

The final product is an app that supports multiple rooms and multiple languages. The user can explore these rooms by teleporting around and can interact with various items to see the foreign name and the corresponding audio clip. They can also see the translations on demand. The user can also interact with the character through text-based questions and answers, which are cycled and reviewed in an intelligent, controlled way by the SM-2 algorithm. These questions are also presented using the foreign language, the corresponding audio and an optional translation, and there are multiple question types. There are a number of settings that can be changed to improve the quality of life, but also to configure how the room is interacted with, allowing the user to make a room more challenging, to assess themselves.

Each interaction is also tracked, which gathers data which can be used in a number of useful ways. This can be analysed to show the learner their progress and their track record, which can be used to motivate them and for self-assessment. This analysis can be used by teachers to review their cohort's progress, and to identify areas which the student's struggle with, or by the app developer, to streamline and improve the learning process.

This analysis has the potential to be used by researchers as a way of comparing different conditions and variables, which allows the final application to also have the potential to be used as a tool in future research.

This final product uses many techniques shown to improve memory retention. Virtual reality and computer-based learning are both meant to increase the quality of learning, as well as the engagement and motivation of its users. The app introduces vocabulary to the users in different ways. Instead of just showing one aspect of the word (foreign word and translation), the app also presents the audio of the item along with a 3D realistic model of the item itself. These different aspects of the word help to give it more meaning, which helps the word to be learned more easily. The app also uses spaced repetition, which reviews questions and words in an intelligent manner to help memory retention.

The user can explore freely, without a great deal of guidance and this helps to provide a user-centric experience that the user can reflect on or review using data analysis. This helps to provide experiential learning, and the questions posed by the character can be investigated by the user, drawing techniques from problem-based learning.

Using various combinations of settings, the user can challenge themselves by removing audio or translations, which could help the user to learn different aspects of the language. Following the comprehensible input, this is also likely to help the user acquire the language more effectively.

Using the expert opinions of language teachers, it is shown that there are some good aspects of the app, with some area for improvement. Several question types and settings were described as posing some good challenges, but some improvements were also offered, such as adding aspects that practice speaking and problem-solving skills.

It is also possible to compare the application that was made against the other existing apps that have been made. When judged against the same criteria, the resulting scores judge the app to be flawed in several ways and is contrasted to other apps to show that they may be better. However, this project's app attempts to take the best elements from each app, and may not have the ability to do all of these to a high extent, given the time available. It is possible to improve these scores by improving on the highlighted issues and limitations.

6.1.1 Issues

There were some issues that were identified during the testing. Effective error handling could be used to identify when there are problems with the network and avoid anything going wrong. Another option would be to have some code identify when a connection has timed out, and to abandon the connection request. More thorough process planning could be used to improve the issues with lag and audio conflicts.

Some of the unimplemented requirements and designs added value which is now missing from the project. For example, having a method for users and teachers to create and change their own rooms would have created a huge amount of flexibility, where any number of different rooms could have been available, and many more languages could have been supported. This could also have allowed the app to be more useful in a classroom or education setting, where a teacher could create or customise a room to align with their curriculum.

Gamification was a piece of designed functionality that wasn't implemented. Research showed that it was an effective method of introducing engagement and motivation, and existing apps implemented this using daily goals, point scores and leader boards. One of the proposed designs for this project's application was to generate point scores through interaction and this would be a method to encourage regular sessions. These would enhance memory due to the spacing effect, as well as providing motivation to the learner, which has been shown to greatly enhance the quality of learning.

The final product implemented a spaced repetition algorithm, which enhances the memory retention of items when they are reviewed multiple times after an interval. This is an effective method to achieve better vocabulary memorisation, however unfortunately there is very little to encourage users to repeatedly return to a room, which could counter the positive impact of the spaced repetition algorithm.

Throughout the project, there was a few instances of requirement creep and scope jump, where new ideas were generated, and more desired functionality was added to the project. This could be positive, as it shows that the project has potential and there isn't a shortage of future work to add to the application to improve it. However, these scope jumps could have impacted the amount of work done on the core functionalities, distracting from work that had a higher priority. This was managed towards the end by a hard deadline. For this deadline, it was decided that all development work would stop, and the result would be the final product, even if there were issues. This did lead to several, simple bugs making their way into the final product that could have been avoided otherwise. This was a decision that could have been handled differently. As it stands, scope jumps were meaning a good deal of time was being spent on the product that otherwise could have been used for polishing the product so that it was bug-free, rather than trying to implement new features such as using different voices for different languages. Stopping these scope jumps from continuing was a good idea, but likewise some planning could have eliminated unnecessary development and prioritised the correct work.

6.1.2 Limitations

There were some areas that could have been polished more, in a way which would improve the quality of learning. There were some issues and bugs identified by testing the final product, as well as some features that could have been developed more. Lag and network latency can frustrate the user but could also induce motion sickness as the headset display does not change when the user moves their head for a moment or two. Audio issues could annoy users or make them mute the audio entirely. Audio is an important element of language learning and removing this because the audio implementation is flawed is a failure to deliver a high-quality product.

Some features would also have benefitted from being polished and worked on more – for example, there is currently a good collection of settings, items and questions, however these are limited, and a wider variety could have led to a better final product.

The final product doesn't support any speech practice. At the moment, reading and listening skills are practiced by showing and playing the foreign item names and the foreign questions. The user practices using the languages when answering the questions, though this is fairly limited, and it has

been shown that being able to use a language is distinct from being able to understand one. The application could improve if there was more opportunity to use the language, either through speech recognition or more thorough writing exercises.

In the same vein of thought, the app has limited functionality. Beyond interacting with items and regularly answering the character's questions, there isn't much else for the user to do. More could be added by fulfilling the unmet requirements and unimplemented designs.

Allowing users greater control over the app could also be beneficial. At the moment, the settings are useful and can be used to good effect, however, the choice is limited, and a more varied selection could be used to even greater effect. Also, the spaced repetition algorithm currently has fixed intervals for when the user first uses the app – after one attempt, the interval is always set to 1 day. Several existing apps allow users to control this interval and other parameters of the algorithm to fully customise the algorithm to the user's preference. This level of control cannot currently be achieved with the app.

Using the app generates data that can be analysed for a number of uses. Unfortunately, this analysis must be manually extracted from the database and manually analysed using various tools. This limits the amount of use that can be gained as the current set-up requires a fair amount of effort to be put in to gain results.

Currently the app also has some limitations that impact how useful the app can be. Currently, due to the libraries that the app uses, it is limited to Android phones only. This does limit how accessible the app can be. The app can be ported to work for other platforms, but it is unknown how difficult this would be. Additionally, a connection to the University network is needed for the app to work. This means that an additional VPN may be needed if the app is used off-campus. It also limits the users of the app to those with access to the network, i.e. students or staff of the University. This could be managed by hosting the web database on a public server that would be more universally available.

Currently the app holds and manages data that could be passed to Microsoft, through its Azure Speech API, Google, through its GoogleVR Unity package, or Loughborough University, through the use of their sci-project server. It is also possible data is passed through other asset packs that are used within this project. Whilst no personal information is handled by the app, a policy of data privacy may need to be considered before the app is widely usable, as some individuals may be uncomfortable with data being handled by these third parties.

6.1.3 Future Work

Future work that can be done involves the work that fixes the issues and solves the limitations of the final product as already detailed. This involves polishing the features that already exist, fixing the bugs and annoyances, working to improve the scores given in

Table 16 and adding more content and more features.

Future work that can be done to build upon what has already been made also involves work that improves on the product and makes it more effective and useful. This includes adding more control to the users. For example, this could involve adding more settings and giving more customisation options in terms of the algorithm, but it could also involve making an editing tool which can be used to make new rooms and modify existing ones. These modifications could include being able to manually change the language, move items, add items, or add questions to the character.

The text-to-speech library the app uses for audio can also be used for translation. Either through this library or another translation API, auto-translation could be a possible improvement, where the user can select the language that they would like to learn, and that room could be created automatically.

The data analysis that can be currently completed is completely manual. The data is good enough that analysis can be carried out, however creating a tool that provides automatic analysis would be greatly beneficial and would greatly increase the user-friendliness of this data. If some framework or API was made to query the database directly and provide analysis as well as raw data, this tool could be very flexible and customisable, allowing users a huge amount of control.

6.2 Project

The project aimed to create an application in virtual reality to assist with second language learning, particularly with vocabulary learning. This aim was broken down into several objectives.

The first objective was to study existing research and case studies to explore what was already written about this area, draw relevant lessons from this existing work and to identify any gaps within the material. A literature review was conducted that investigated research in this area and identified useful lessons that could be taken. A gap analysis was also completed to examine existing tools that were similar. This analysis identified that there was a gap in the market for a language learning VR app that focussed on vocabulary memorisation.

The second objective had several parts. The first part was to identify any existing tools, features and functionality that could be analysed and used in this project. This was achieved by the gap analysis as well. The analysis examined existing vocabulary learning apps, and existing language learning VR apps. This analysis evaluated each one and identified important aspects of each that could be effective in this project.

The second part of the second objective was to determine if there was a public desire for this tool to be made. A questionnaire was conducted which surveyed the public about the use of different

platforms for language learning and this identified that there was a great deal of positive interest in a VR language learning app, as well as one specifically geared toward vocabulary memorisation.

The last two parts of this objective are to identify what hardware and software to use, and this was completed when the implementation was started.

The third objective was to create the language learning app itself. To do this, requirements were generated based on the research and gap analysis, and these were expanded into detailed designs on how the app should function and look. The app was then implemented, using these designs and implementing them. Decisions were made on how certain aspects should be implemented and these were presented clearly in Chapter 4.

The fourth objective was to evaluate the project. One part of this was to judge how successful this project has been, and this is completed in this chapter, making use of the rigorous testing that was completed on the app itself. This objective also is to evaluate what could have been achieved with a longer timeframe, how to improve the product further and what can be done in future to improve the field. This is also completed in this chapter.

The aim of creating a language learning tool in virtual reality for vocabulary memorisation has been completed successfully, and the objectives that made up this aim were also met. This shows that the project did manage to complete its purpose to a high degree.

Largely, development of the app itself went smoothly. The Autumn term was spent learning how Unity worked and this preparation really aided the development. There was a few bugs and difficulty in development, but these were solved relatively painlessly.

For the January deliverable, a first draft of the Introduction, Requirements and Design Chapters were written – and as this was before implementation had started, this meant that a solid design that was based on research had already been made, which greatly helped, as while implementing, attention would not need to be spent working on the design.

As previously mentioned, there were features and elements of the app that were added and the scope of the app did grow. Several designed features could not be implemented, such as the editing tool, and there were several features that were included and could have been implemented more. With more time, the app could have become very good, however currently it is limited due to having not much content and not much motivation to return regularly.

Work was distributed well over the year. There was not a last-minute rush at the end, and with a month to go until the deadline, development of the app had finished and a first draft of three

chapters was completed. This helped allow plenty of time to write and then polish the report to a high standard.

6.2.1 Issues

A Gantt chart was created at the start of the project, aiming to plan how much time to spend on each element. Whilst the general order that these elements were planned to be worked on was the same as the order that they were worked on, this Gantt chart wasn't followed very closely, and the time spent on each element was different to the time planned. Throughout the project, more short-term plans were made and used to plan how certain chapters / functionalities were to be approached. This helped organise the workload and prompt the correct element to be worked on, however the lack of a more long-term plan in place didn't help. Using a long-term plan like the Gantt chart would have helped keep the prioritised work on track and guided the project to be more time efficient.

Several requirements and designed elements could not be completed as there was not enough time to implement them. This is a shame and could suggest that time was spent inefficiently in some places.

There were several time constraints – mainly exams and various courseworks, where attention and focus were not on the project. During these times, it was planned to split time effectively between the project and the other work, however this did not go ahead. Instead, during the periods where there was an exam or coursework, all of my attention was spent on that, and afterwards I had to reacquaint myself with the project.

Also, another potential issue with how the project was worked on was work distribution. Whilst work was consistent throughout the year, which was good, there was an imbalance between how much work was spent working in any one day. Whilst this should have been a consistent number of hours a day, the amount of time spent on the project varied hugely day to day, from only 3 hours to the more excessive 12 hours. This perhaps should have been managed more efficiently

6.2.2 Limitations

The literature review that was completed was thorough and explored many areas. Some of the areas reviewed, for example, second-language acquisition, was explored in great detail, but not used very much, and this could suggest that some elements of the review are investigated to great depth unnecessarily. Likewise, some areas that were used, such as gamification, were left without a great deal of exploration. There was some imbalance in the literature review, where certain time and

space could have been used more effectively to represent the literature that would impact the project more.

There were also some parts of the literature review that explained each field and the findings of the studies but did not do any in-depth evaluation or discussion. If this had been completed, certain points could have been found that could have helped the project, however this instead was not done. In this way, these parts of the literature review serve more as background research than actuals parts of the review.

The questionnaire that was run to survey whether the public want a VR tool for vocabulary memorisation was shared in two main groups – on an online language-learning community, and on the main university student group. Arguably this doesn't represent the general population, and instead is focussed toward people who already have experience or an interest in language learning. In the student group, many are international students who may have already had to learn English, and many UK students have had to learn a second language through the education system. Furthermore, many students who completed the survey would have been at least somewhat interested in the material.

This different sample population would mean the results – overwhelmingly positive – could only represent that of the population that has already had experience or interest in language learning. This does not mean the results are worthless, but it does potentially change whether this app is desired by the general population or specifically by language learners.

The testing that was completed in Chapter 5 assesses if the final product acts as it was designed to, however this can be flawed in that it doesn't explicitly test the finer details, where the app might break. The testing completed does ensure that the final product is the app that was designed, but the testing cannot conclude whether the app is good-quality or well-made and does not break under closer scrutiny. Likewise, the testing cannot properly determine whether the app can teach a user new words or improve their learning. Testing the app's quality of learning would require a more robust experiment to measure a participant's language performance over time as the app was used.

6.2.3 Future Work

Given a longer timeframe, the application could have been developed further, adding more functionality to the app. This would give the users more options and more control, and potentially expand the scope of the app, and make it a product where the user can create their own rooms. Existing tools, like Memrise, allows its community to make their own sets of flashcards, and this has allowed a broad range of material to be studied. Allowing users to create their own rooms could

allow the same – rooms for language memorisation but also rooms for familiarisation, where users can use the room to learn new terms or jargon for more specific items like ‘server cabinet’, ‘Ethernet’ and ‘Switch’, for example.

There are a few ways the existing work can be improved. The app itself can be improved by fixing the bugs and issues it already has, and new features can be added and built on top of it.

The project can be improved by redrafting the report to better balance some elements. Issues with the questionnaire, literature review and testing expose some flaws with the project and how some aspects of the app and its implementation process are not presented in a balanced way. Adding more detail to the data analysis section of the implementation would provide a clearer method of analysing results, and potential work could be done to automate this and make this more user-friendly.

The analysis that can be done on the data gathered by the app can be used as a tool to compare and measure different variables. This suggests that this app and the analysis can potentially be used as a tool within further research. This tool could be used to compare the effectiveness of different settings, the easiness of different sets of words or languages, or perhaps it could be used to evaluate whether the easiness factor in the SM-2 algorithm is an accurate measure of a word’s difficulty.

6.3 Summary

Throughout this project, numerous lessons were drawn from the body of existing research and work done on language learning and VR, and these lessons helped to design and implement the final product. This product is perfectly functional and can be used alongside classroom learning or standalone to help teach vocabulary and aid in vocabulary memorisation. There are some issues and bugs in the final implementation, and some features that could not be implemented in time, however the project achieved its aim of building a functional tool that has the potential to be built upon and improved further, as well as be used as a measuring and analysis tool in future research. This project was a success that managed to build this tool while bringing research, case studies and public opinion into the final design, using several powerful libraries and algorithms to ensure a quality implementation and learning experience.

References

Adamidi, F. *et al.* (2017) 'Problem-Based Learning in Language Instruction: A Collaboration and Language Learning Skills Framework in a CSCL Environment', in *Communications in Computer and Information Science*. Springer Verlag, pp. 133–146. doi: 10.1007/978-3-319-62743-4_12.

Anki (2020) *FAQs - Anki Manual*, *Anki*. Available at: https://docs.ankiweb.net/#/faqs?id=what-spaced-repetition-algorithm-does-anki-use (Accessed: 5 April 2020).

Ardoin, S. P., Wagner, L. and Bangs, K. E. (2016) 'Applied Behavior Analysis: A Foundation for Response to Intervention', in *Handbook of Response to Intervention*. Boston, MA: Springer US, pp. 29–42. doi: 10.1007/978-1-4899-7568-3_3.

Atkinson, R. C. and Shiffrin, R. M. (1968) 'Human Memory: A Proposed System and its Control Processes', in *Psychology of Learning and Motivation - Advances in Research and Theory*, pp. 89–195. doi: 10.1016/S0079-7421(08)60422-3.

Bailey, J. *et al.* (2012) 'Presence and Memory: Immersive Virtual Reality Effects on Cued Recall', in *Proceedings of the International Society for Presence Research Annual Conference*. Philadelphia, Pennsylvania.

Cambridge Dictionary (no date) *VOCABULARY | meaning in the Cambridge English Dictionary*, *Cambridge Dictionary*. Available at: https://dictionary.cambridge.org/dictionary/english/vocabulary (Accessed: 22 December 2019).

Cho, Y. (2018) 'How Spatial Presence in VR Affects Memory Retention and Motivation on Second Language Learning: A Comparison of Desktop and Immersive VR-Based Learning', *Theses - ALL*. Available at: https://surface.syr.edu/thesis/204 (Accessed: 22 December 2019).

Chung, L. (2012) 'Incorporating 3D-Virtual Reality into Language Learning', *International Journal of Digital Content Technology and its Applications*, 6(6), pp. 249–255. doi: 10.4156/jdcta.vol6.issue6.29.

Coomber, C. (2020) *Let's Play: Angle Vocabulary Review (in Half Life: Alyx) - Supplementary, Vertical and Complementary - YouTube*, *Youtube*. Available at: https://www.youtube.com/watch?v=_T2-9MwA5JI&feature=emb_logo (Accessed: 7 April 2020).

Cooper, J. O. (1982) 'Applied behavior analysis in education', *Theory Into Practice*, 21(2), pp. 114–118. doi: 10.1080/00405848209542992.

Craik, F. I. M. and Lockhart, R. S. (1972) 'Levels of processing: A framework for memory research',

Journal of Verbal Learning and Verbal Behavior, 11(6), pp. 671–684. doi: 10.1016/S0022-5371(72)80001-X.

Fox3D (no date) *Fox3D VR*, *Fox3D*. Available at: http://fox3d.com/vr (Accessed: 30 December 2019).

Galoyan, T., Shah, M. and Foster, A. (2017) 'Facilitating Language Use and Communication in ESL/EFL Classrooms through Game-Based Learning', in *The IAFOR International Conference on Language Learning –Hawaii 2017 Official Conference Proceedings*. Hawaii, p. 13. Available at: www.iafor.org (Accessed: 21 December 2019).

Glanzer, M. and Cunitz, A. R. (1966) 'Two storage mechanisms in free recall', *Journal of Verbal Learning and Verbal Behavior*, 5(4), pp. 351–360. doi: 10.1016/S0022-5371(66)80044-0.

Godwin-Jones, R. (2010) 'EMERGING TECHNOLOGIES FROM MEMORY PALACES TO SPACING ALGORITHMS: APPROACHES TO SECOND-LANGUAGE VOCABULARY LEARNING', *Language, Learning and Technology*, 14(2), pp. 4–11. Available at: http://llt.msu.edu/vol14num2/emerging.pdf (Accessed: 5 April 2020).

Gwern.net (2019) *Spaced Repetition for Efficient Learning - Gwern.net*, *Gwern.net*. Available at: https://www.gwern.net/Spaced-repetition# (Accessed: 27 December 2019).

Hmelo-Silver, C. E. (2004) 'Problem-Based Learning: What and How Do Students Learn?', *Educational Psychology Review*, 16(3), pp. 235–266. doi: 10.1023/B:EDPR.0000034022.16470.f3.

KOLB, D. (2000) 'The Process of Experiential Learning', in *Strategic Learning in a Knowledge Economy*. Elsevier, pp. 313–331. doi: 10.1016/B978-0-7506-7223-8.50017-4.

Kolb, D. A. *et al.* (2000) *Experiential Learning Theory: Previous Research and New Directions*.

Kornell, N. (2009) 'Optimising learning using flashcards: Spacing is more effective than cramming', *Applied Cognitive Psychology*, 23(9), pp. 1297–1317. doi: 10.1002/acp.1537.

Kornell, N. and Bjork, R. A. (2008) 'Optimising self-regulated study: The benefits—and costs—of dropping flashcards', *Memory*, 16(2), pp. 125–136. doi: 10.1080/09658210701763899.

Krahnke, K. J. and Krashen, S. D. (1983) 'Principles and Practice in Second Language Acquisition', *TESOL Quarterly*, 17(2), p. 300. doi: 10.2307/3586656.

Krokos, E., Plaisant, C. and Varshney, A. (2019) 'Virtual memory palaces: immersion aids recall', *Virtual Reality*. Springer London, 23(1), pp. 1–15. doi: 10.1007/s10055-018-0346-3.

Lee, K. M. (2004) 'Presence, Explicated', *Communication Theory*, 14(1), pp. 27–50. doi:

10.1111/j.1468-2885.2004.tb00302.x.

Lessard-Clouston, M. (2013) *Teaching vocabulary*. Tesol International Association.

Levy, M. (1997) *Computer-assisted language learning: Context and conceptualization*. Oxford University Press. Available at: https://books.google.com/books?hl=en&lr=&id=RRGgrjteVjUC&oi=fnd&pg=PR8&dq=computer+assisted+language+learning&ots=8r8jG2qGo9&sig=rRgRy_K15NkEzBSmQAJTnocMMvE (Accessed: 27 December 2019).

Levy, M. and Stockwell, G. (2008) *CALL dimensions: Options and issues in computer-assisted language learning*. Routledge. Available at: https://books.google.com/books?hl=en&lr=&id=PsgRcUHoocEC&oi=fnd&pg=PP1&dq=computer+assisted+language+learning&ots=79Zr1Qp8Ud&sig=t7Mw7oImtV0ZrTdv0qu6rkBET5I (Accessed: 27 December 2019).

Makransky, G. and Lilleholt, L. (2018) 'A structural equation modeling investigation of the emotional value of immersive virtual reality in education', *Educational Technology Research and Development*. Springer New York LLC, 66(5), pp. 1141–1164. doi: 10.1007/s11423-018-9581-2.

Mania, K. and Chalmers, A. (2001) 'The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach', *CyberPsychology & Behavior*, 4(2), pp. 247–264. doi: 10.1089/109493101300117938.

McLeod, S. (2007) *Levels of Processing | Simply Psychology*, *SimplyPsychology*. Available at: https://www.simplypsychology.org/levelsofprocessing.html (Accessed: 19 December 2019).

McLeod, S. (2008) *Serial Position Effect | Simply Psychology*, *SimplyPsychology*. Available at: https://www.simplypsychology.org/primacy-recency.html (Accessed: 19 December 2019).

McLeod, S. (2017a) *Atkinson and Shiffrin | Multi Store Model of Memory | Simply Psychology*, *SimplyPsychology*. Available at: https://www.simplypsychology.org/multi-store.html (Accessed: 19 December 2019).

McLeod, S. (2017b) *Kolb's Learning Styles and Experiential Learning Cycle | Simply Psychology*, *SimplyPsychology*. Available at: https://www.simplypsychology.org/learning-kolb.html (Accessed: 23 December 2019).

McLeod, S. (2018) *Questionnaire | Simply Psychology*, *SimplyPsychology*. Available at: https://www.simplypsychology.org/questionnaires.html (Accessed: 31 December 2019).

Memrise (no date) *Dashboard - Memrise*, *Memrise*. Available at: https://www.memrise.com/home/ (Accessed: 5 January 2020).

Meyer, L. L. and Schmitt, N. (2002) 'Vocabulary in Language Teaching', *TESOL Quarterly*, 36(2), p. 235. doi: 10.2307/3588334.

Microsoft Azure (2020a) *Cognitive Speech Services pricing | Microsoft Azure*, *Microsoft Azure*. Available at: https://azure.microsoft.com/en-gb/pricing/details/cognitive-services/speech-services/ (Accessed: 3 April 2020).

Microsoft Azure (2020b) *Language support - Speech service - Azure Cognitive Services | Microsoft Docs*, *Microsoft Azure*. Available at: https://docs.microsoft.com/en-gb/azure/cognitive-services/speech-service/language-support#speech-to-text (Accessed: 3 April 2020).

Microsoft Azure (2020c) *Quickstart: Synthesize speech to a speaker - Speech service - Azure Cognitive Services | Microsoft Docs*, *Microsoft Azure*. Available at: https://docs.microsoft.com/en-gb/azure/cognitive-services/speech-service/quickstarts/text-to-speech?tabs=unity%2Clinux%2Cjre&pivots=programming-language-csharp (Accessed: 3 April 2020).

Microsoft Azure (2020d) *Speech to Text | Microsoft Azure*, *Microsoft Azure*. Available at: https://azure.microsoft.com/en-gb/services/cognitive-services/speech-to-text/#overview (Accessed: 3 April 2020).

Mnemosyne Project (2020) *Memory research | The Mnemosyne Project*, *Mnemosyne Project*. Available at: https://mnemosyne-proj.org/help/memory-research (Accessed: 5 April 2020).

Mondly (no date) *Press Releases Archives - Mondly Blog*, *Mondly*. Available at: https://www.mondly.com/blog/press-releases/ (Accessed: 30 December 2019).

Nation, I. S. P. (1990) *Teaching and learning vocabulary*. Heinle & Heinle Publishers.

Osipov, I. V. *et al.* (2015) 'Study of Gamification Effectiveness in Online e-Learning Systems', *International Journal of Advanced Computer Science and Applications*. The Science and Information Organization, 6(2), pp. 71–77. doi: 10.14569/IJACSA.2015.060211.

Perkins Coie (2018) *VR/AR investment focus worldwide 2018*, *Perkins Coie*. Available at: https://www.statista.com/statistics/829729/investments-focus-vr-augmented-reality-worldwide/ (Accessed: 4 January 2020).

Raugh, M. R., Schupbach, R. D. and Atkinson, R. C. (1977) 'Teaching a large Russian language vocabulary by the mnemonic keyword method', *Instructional Science*. Kluwer Academic Publishers,

6(3), pp. 199–221. doi: 10.1007/BF00120656.

Richards, J. C. and Rodgers, T. S. (Theodore S. (1986) *Approaches and methods in language teaching : a description and analysis*. Cambridge University Press.

Royal Literary Fund (no date) *What is a literature review?*, *Royal Literary Fund*. Available at: https://www.rlf.org.uk/resources/what-is-a-literature-review/ (Accessed: 31 December 2019).

Rustamovna, R. V. (2019) 'Effective Teaching Strategies in Teaching Foreign Language', *International Journal of Research*, 6(6), pp. 60–65. Available at: https://journals.pen2print.org/index.php/ijr/ (Accessed: 21 December 2019).

Schimanke, F., Mertens, R. and Huck, B. S. (2018) 'Player Types in Mobile Learning Games – Playing Patterns and Motivation', in *2018 IEEE International Symposium on Multimedia (ISM)*. IEEE, pp. 247–252. doi: 10.1109/ISM.2018.00035.

Schmitt, N. (2008) 'Review article: Instructed second language vocabulary learning', *Language Teaching Research*, 12(3), pp. 329–363. doi: 10.1177/1362168808089921.

Schutz, R. E. (2019) *Stephen Krashen's Theory of Second Language Acquisition*. Available at: https://www.sk.com.br/sk-krash-english.html (Accessed: 19 December 2019).

Settles, B. and Meeder, B. (2016) 'A Trainable Spaced Repetition Model for Language Learning', in *54th Annual Meeting of the Association for Computational Linguistics*. Berlin: Association for Computational Linguistics, pp. 1848–1858. Available at: https://www.duolingo.com (Accessed: 10 April 2020).

Solak, E. and Cakir, R. (2015) 'Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning', *The Journal of Educators Online*, 12(2), pp. 50–72. doi: 10.9743/JEO.2015.2.5.

SuperData Research; Statista estimates (2019) *Consumer VR software market size worldwide 2016-2022*, *Statista*. Available at: https://www.statista.com/statistics/550474/virtual-reality-software-market-size-worldwide/ (Accessed: 4 January 2020).

SuperData Research (2019) *Consumer virtual reality market size worldwide 2016-2022*, *SuperData Research*. Available at: https://www.statista.com/statistics/528779/virtual-reality-market-size-worldwide/ (Accessed: 4 January 2020).

Tanaka, K. and Ellis, R. (2003) 'Study-abroad, language proficiency, and learner beliefs about language learning', *JALT Journal*, 25(1), pp. 63–85. Available at: http://www.jalt-

publications.org/archive/jj/2003a/art3.pdf.

Turner, J. (2014) *Using Statistics in Small-Scale Language Education Research: Focus on Non ... - Jean L. Turner - Google Books*. Routledge. Available at: https://books.google.co.uk/books?hl=en&lr=&id=TafmAgAAQBAJ&oi=fnd&pg=PP1&dq=data+analysis+in+language+education&ots=zAS6ocD8qJ&sig=TZGADGyAQzx6UX0UBQy8JRfU-V4&redir_esc=y#v=onepage&q=data analysis in language education&f=false (Accessed: 10 April 2020).

Turner, J. (no date) *Multi-Store Model of Memory | Psychology | tutor2u*. Available at: https://www.tutor2u.net/psychology/reference/multi-store-model-of-memory (Accessed: 19 December 2019).

Unity Technologies (2020) *Unity Real-Time Development Platform | 3D, 2D VR & AR Visualizations*, *Unity Technologies*. Available at: https://unity.com/ (Accessed: 1 April 2020).

Vinet, L. and Zhedanov, A. (2011) 'A "missing" family of classical orthogonal polynomials', *Journal of Physics A: Mathematical and Theoretical*. IEEE, 44(8), p. 085201. doi: 10.1088/1751-8113/44/8/085201.

Weinstein, Y., McDermott, K. B. and Roediger, H. L. (2010) 'A comparison of study strategies for passages: Rereading, answering questions, and generating questions.', *Journal of Experimental Psychology: Applied*, 16(3), pp. 308–316. doi: 10.1037/a0020992.

Wilkins, D. (1972) *Linguistics in language teaching*. London: Edward Arnold.

Wozniak, P. A. (1998) *Application of a computer to improve the results obtained in working with the SuperMemo method*, *SuperMemo*. Available at: https://www.supermemo.com/en/archives1990-2015/english/ol/sm2 (Accessed: 5 April 2020).

Appendices

Appendix A – Questionnaire Transcript

How old are you? *

- Under 18 years old (If chosen, ends questionnaire straight away)
- 18-25
- 26-35
- 36-45
- 46-55
- 56-65
- 66-75
- 76 years or older

Have you ever started learning a second language? *

- Yes
- No

How effective are these mediums for learning a second language? (Marking out of 5 (1 = Not at all, 5 = Extremely), with a N/A option as well) *

- Classes / One-on-one lessons
- Textbooks
- Websites
- Forums
- Mobile Apps
- VR Applications

Do you think VR would be a good medium to learn a language? *

- Yes
- No

If the user selected yes to 'Have you ever learned a second language?'

- What language(s) have you learned?
- What mediums have you used? *
    - Classes / One-on-one lessons
    - Textbooks
    - Websites
    - Forums
    - Mobile Apps
    - VR Applications
    - Other

My final year project is to create a VR application for vocabulary memorisation.

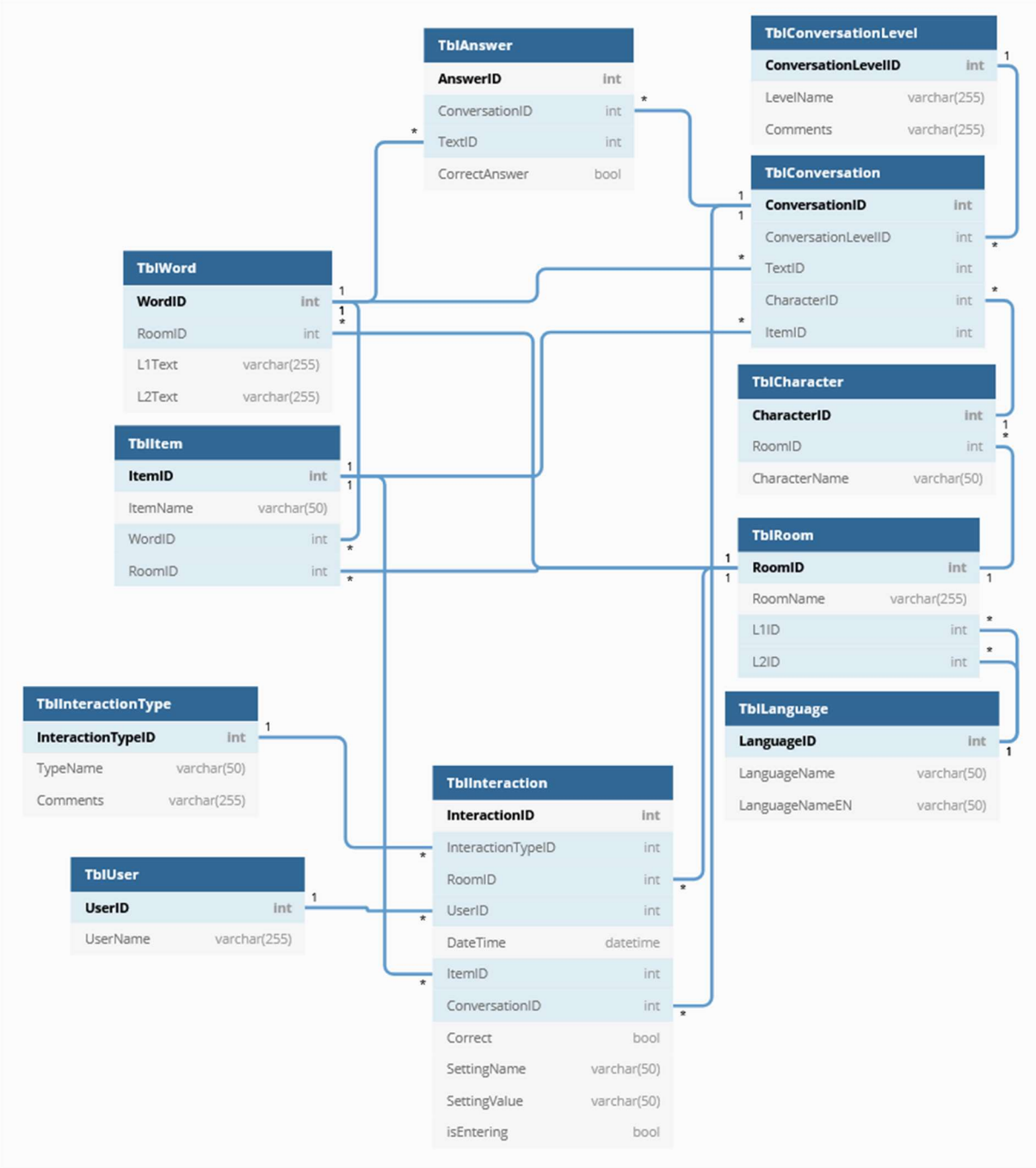
Would you be interested in using this program to learn a new language? *

- Yes
- No

Do you have any ideas to add to this?

Appendix B – Database Architectures

Web Database Architecture



Local Database Architecture

Appendix C – Data Analysis VBA Code

```
Sub test()
  On Error GoTo Error

  Dim cell As range
  Dim row1 As range

  Dim intIDList As Object
  Set intIDList = CreateObject("System.Collections.ArrayList")

  Dim userIDList As Object
  Set userIDList = CreateObject("System.Collections.ArrayList")

  Dim conIDList As Object
  Set conIDList = CreateObject("System.Collections.ArrayList")

  Dim correctList As Object
  Set correctList = CreateObject("System.Collections.ArrayList")

  Dim EFList As Object
  Set EFList = CreateObject("System.Collections.ArrayList")

  Dim intervalList As Object
  Set intervalList = CreateObject("System.Collections.ArrayList")

  Dim indexList As Object
  Set indexList = CreateObject("System.Collections.ArrayList")

  Dim qList As Object
  Set qList = CreateObject("System.Collections.ArrayList")

  Dim index As String

  Dim counter As Integer

  Dim q As Integer
  Dim newInterval As Double
```

```
  Dim newEF As Double

  Worksheets("user-con-alg-audioOnly").Activate

  Debug.Print ActiveSheet.UsedRange.Rows.Count

  'for every row
  For Each cell In ActiveSheet.UsedRange.Rows

    'is it not FALSE
    If cell.Cells(1, 1).Value <> "False" Then

      'have we seen the pairing before?
      index = CStr(cell.Cells(1, 2)) & "," & CStr(cell.Cells(1, 3))
      If indexList.Contains(index) = True Then
        ' we have seen this before on row:
        counter = indexList.IndexOf(index, 0)
        q = qList.Item(counter)

      If (cell.Cells(1, 4) = 1) Then
        'if correct answer

          EFList.Add 2.5
          cell.Cells(1, 5).Value = 2.5

          If (intervalList.Item(counter) = 0) Then
            intervalList.Item(counter) = 1
            cell.Cells(1, 6).Value = 1
          Else
            newInterval = intervalList.Item(counter) *
EFList.Item(counter)
            intervalList.Item(counter) = newInterval
            cell.Cells(1, 6).Value = newInterval
          End If

          q = qList.Item(counter)
```

```
          newEF = EFList.Item(counter) + (0.1 - (5 - q) * (0.08 + (5
- q) * 0.002))

          If (newEF > 2.5) Then
            newEF = 2.5
          End If
          If (newEF < 1.1) Then
            newEF = 1.1
          End If

          EFList.Item(counter) = newEF
          cell.Cells(1, 5).Value = newEF

          qList.Item(counter) = 5

        Else
        'if incorrect answer
          q = q - 1
          If q < 3 Then
            q = 3
          End If

          qList.Item(counter) = q


          newEF = EFList.Item(counter) + (0.1 - (5 - 2) * (0.08 + (5
- 2) * 0.002))

          If (newEF > 2.5) Then
            newEF = 2.5
          End If
          If (newEF < 1.1) Then
            newEF = 1.1
          End If

          EFList.Item(counter) = newEF
          cell.Cells(1, 5).Value = newEF
```

```
  End If

Else
  ' new combo
  indexList.Add index

  'set up all lists
  intIDList.Add CStr(cell.Cells(1, 1))
  userIDList.Add CStr(cell.Cells(1, 2))
  conIDList.Add CStr(cell.Cells(1, 3))
  correctList.Add CStr(cell.Cells(1, 4))
  If (cell.Cells(1, 4) = 1) Then
  'if correct answer

    EFList.Add 2.5
    cell.Cells(1, 5).Value = 2.5

    intervalList.Add 1
    cell.Cells(1, 6).Value = 1

    q = 5
    qList.Add q

  Else
  'if incorrect answer
    newEF = 2.5 + (0.1 - (5 - 2) * (0.08 + (5 - 2) * 0.002))

    If (newEF > 2.5) Then
      newEF = 2.5
    End If
    If (newEF < 1.1) Then
      newEF = 1.1
    End If

    EFList.Add newEF
    cell.Cells(1, 5).Value = newEF
```

```
          intervalList.Add 0
          cell.Cells(1, 6).Value = 0

          q = 4
          qList.Add q
        End If

      End If

    End If

  Next cell

  Debug.Print indexList.Count


  Dim cnt As Long
  For cnt = 0 To indexList.Count

  Debug.Print "Index: " + CStr(indexList.Item(cnt)) + " Ending EF: "
+ CStr(EFList.Item(cnt)) + " Ending Interval: " +
CStr(intervalList.Item(cnt))

  Next
  GoTo Ending

Error:
Debug.Print "Error " & Err.Description

Ending:
End Sub
```

Appendix D – Testing Tables

Item Interaction

Test No.	Name	Expected Outcome	Pass/Fail	Comments
1	Interactable Item Cursor Recognition	The user's cursor will form a circle when it hovers over an interactable object	Pass	
2	User Clicking Input on Item	When the user clicks an interactable item, a textbox will show up above the object	Pass	
3	Foreign Item Name	The textbox will have the foreign word displayed on it	Pass	
4	Item Audio	An audio file will play at the same time as the textbox appearing, and will be the foreign word being read out	Pass	There is some lag before this plays. This lag locks the screen, so the user can still move their head, but this movement won't be taken into account until after the lag. This can break immersion and potentially cause motion sickness.
5	Item Translation	When the user's cursor hovers over the textbox, the translation is shown.	Pass	
6	Item Close Button	When the close button is pressed, the textbox disappears.	Pass	

Character Interaction

Test No.	Name	Expected Outcome	Pass/Fail	Comments
7	Interactable Character Cursor Recognition	The user's cursor will become a circle when it hovers over the character	Pass	
8	User Clicking Input on Character	When the user clicks the character a UI appears	Pass	
9	Foreign Character Question	When the UI appears a question is displayed in the foreign language	Pass	
10	Character Audio	When the UI appears an audio clip plays at the same time	Pass	There is a slight lag between the UI appearing and the audio being played. This lag is similar to the lag in Test 4.
11	Character Audio Correctness	The audio that is played is the appropriate audio where the question is read out.	Pass	When interacting with the character, the audio that starts playing may be the audio from the question before – this does cut off after a second or two, at which point the correct audio does play. This is due to several processes occurring at the same time and this not necessarily being considered, leading to the audio being played before the correct audio clip has been loaded. To fix this, a proper process diagram could be drawn to make sure that audio isn't set to play before the audio clip is ready.

12	Character Answers	When the UI appears, relevant answers will be displayed	Pass	
13	Question Translation	When the user hovers over the question, the translation is displayed	Pass	
14	Character Audio on Demand	The user can interact to play the audio again	Pass	The user must click the character again. This is awkward due to the UI interface in the way and would be improved with a dedicated button. This is due to a lack of planning.

Character Conversations

Test No.	Name	Expected Outcome	Pass/Fail	Comments
Multiple Choice				
15	Multiple Choice Answer Count	There will be four options when a multiple choice question is displayed	Pass	
16	Multiple Choice Correct Prompt	When the user selects an answer, they will be told if they got it correct or not	Pass	
17	Multiple Choice User Corrected	If the user selected the incorrect answer, they will be told the correct answer	Pass	
18	Multiple Choice UI Close	After answering the question, the UI will disappear	Pass	
Answer Blocks				
19	Answer Block Answer Count	There will be six answers for an 'answer block' question.	Pass	
20	Answer Block Correct Prompt	When the user selects an answer, they will be told if they got it correct or not.	Pass	
21	Answer Block Word Construction	If the user answers correctly, the text being constructed will be displayed	Pass	
22	Answer Block UI Close	If the user has constructed the full answer, this will be displayed, and the UI will close.	Pass	

Character Conversation Algorithm

Test No.	Name	Expected Outcome	Pass/Fail	Comments
23	Spaced Repetition Intervals	Questions will be revisited at varying time intervals according to a spaced repetition algorithm	Pass	
24	Incorrect Answer Review	If a question is answered incorrectly, it will be revisited soon after to review and confirm that the user has learned it.	Pass	

Movement

Test No.	Name	Expected Outcome	Pass/Fail	Comments
25	Teleport Target Recognition	The user will be able to see the teleportation target when they look at the floor	Pass	
26	Teleport Click	The user will be able to teleport to the target when they click	Pass	
27	Teleport Spamming	The user will teleport multiple times if the teleport target is clicked multiple times in quick succession.	Pass	

Settings Menu

Test No.	Name	Expected Outcome	Pass/Fail	Comments
28	Menu Button	When the user clicks the menu button, the UI will change into the settings menu screen	Pass	
29	Menu Toggle Input	When the user selects a toggle, it will change from a tick to an empty box and the text will change from ON to OFF or vice versa	Pass	
30	Menu Slider Input	When the user changes the slider, the slider will change to reflect that, and the text will display the percentage it is at.	Pass	
31	Item Audio Setting OFF	When the audio is OFF, no audio will be played when clicking on items	Pass	
32	Character Audio Setting OFF	When the audio is OFF, no audio will be played when clicking on the character	Pass	
33	Item Translation Setting OFF	When the translations are OFF, no translations will be displayed when hovering over an item's textbox	Pass	
34	Question Translation Setting OFF	When the translations are OFF, no translations will be displayed when hovering over questions	Pass	
35	Audio Setting Persistence	Audio settings will be saved and retained upon re-entering the room.	Pass	
36	Audio Volume Setting	Changing the audio volume slider leads to an appropriate change in the volume of the audio played.	Pass	
37	No Text when using Audio Only Setting on Items	When Audio Only is ON, there will be no text displayed when clicking on items	Pass	

38	Audio when using Audio Only Setting on Items	When Audio Only is ON, audio will play when clicking on items	Pass	
39	No Question Text when using Audio Only Setting on Characters	When Audio Only is ON, there will be no question text displayed when clicking on the character	Pass	
40	Audio when using Audio Only Setting on Characters	When Audio Only is ON, audio will play when clicking on the character	Pass	
41	Guided Learning Sentences	When Guided Learning is ON, talking to the character starts by have sentences as the question text, with no answers.	Pass	
42	Guided Learning Arrow	When Guided Learning is ON, talking to the character starts with an arrow above the character pointing at the right answer.	Pass	
43	Exit to Menu Scene Button	When the user presses the 'exit to menu' button, the scene will change to the menu scene.	Pass	

Menu Scene

Test No.	Name	Expected Outcome	Pass/Fail	Comments
44	Menu Scene Startup	Upon startup of the application, a room is displayed to the user	Pass	
45	Menu Scene Panels	When the room appears, there are two UI panels	Pass	
46	Menu Scene Available Room List	On one UI panel is a list of available rooms	Pass	At the time of the test, there are two rooms available – both lead to the Kitchen Scene, one for Japanese, and one for French.
47	Room Menu Item	When the user clicks on one of these rooms, it opens the selected room	Pass	
48	Menu Scene Tutorial Panel	On the other panel is a list of instructions and guiding sentences.	Pass	

Text-to-Speech

Test No.	Name	Expected Outcome	Pass/Fail	Comments
49	Audio Human Voice	When audio is played, the audio sounds like an understandable human voice	Pass	
50	Audio correct phrase	When audio is played, the appropriate word or phrase is read	Pass	
51	Audio correct pronunciation	When audio is played, the voice and pronunciation is recognisable as being typical of the appropriate language	Pass	Some pronunciations are inaccurate for the Japanese room. This could be because the text-to-speech system is using roman characters rather than the proper Japanese characters. Largely, the voice is appropriate and most pronunciations are fine.

Network

Test No.	Name	Expected Outcome	Pass/Fail	Comments
52	Network Downloads Onsite	When connected to the university network directly by being on site, the app downloads all data without issue. This was measured on a mobile phone using a laptop to measure the logs being generated.	N/A	This could not be tested before the lockdown and so the outcome is unknown.

53	Network Downloads VPN	When connected to the university network via a VPN, the app downloads all data without issue. This was measured on a mobile phone using a computer to measure the logs being generated.	Pass	
54	VPN disconnects	When using the app, if the VPN disconnects, the app continues functioning as normal and notifies the user of any issues.	Fail	When disconnected, movement works as normal, but item or character interaction freezes the app. This requires a restart. This is because the app is attempting to update the web database and failing. This could be solved with some code to detect a timeout, at which point it gives up.
55	Internet Connection Lost	When using the app, if the internet connection fails, via WiFi or data, the app continues functioning as normal and notifies the user of any issues.	Fail	When offline, movement works as normal, but item or character interaction freezes the app. This requires a restart. This is because the request for audio goes unanswered and the app hangs. This could be solved with some code to detect a timeout, or to detect the device's connection status.

Data Analysis

Test No.	Name	Expected Outcome	Pass/Fail	Comments
56	Data Gathering after item interactions	After every interaction with an item, data is gathered about which item was clicked.	Pass	
57	Data Gathering after character interactions	After every interaction with a character, data is gathered about which conversation was answered and if it was a correct answer.	Pass	
58	Data Gathering after settings interactions	After every interaction with the settings, data is gathered indicating which setting was changed and its new value.	Pass	
59	Data Gathering after entering/exiting a room	After leaving or entering a room, data is gathered about which room was involved and whether it was entering or leaving.	Pass	

Appendix E – Language Centre Emails

Cristina Hidalgo-Orihuela

Dear Peter,

My name is Cristina and I am one of the Spanish teachers at the Language Centre.

First of all, I want to congratulate you for your final year project. It is a very good one. Well done!

I have mostly enjoyed two of the app options: moving around without translations and putting the syllables in the right order to form the words. Great challenges for the learners!

Your work is fantastic, but if I could suggest any improvements, they would be:

1. when the answer is wrong, instead of giving the correct answer, encourage the learner to move around the room until he/she finds the appropriate word.
2. for practising pronunciation, it would be useful to include an activity in which the learner speaks and the character moves towards the object (if correct). Thus, the app would allow to practice all four language skills (listening, reading, writing, speaking).

I hope this helps. Best of luck!

Kind regards,

Cristina

Cristina Hidalgo Orihuela FHEA
University Teacher (Spanish)
Language Centre
Herbert Manzoni K.1.23
Loughborough University
LE11 3TU
UK

Telephone: 01509 222974
Email: c.hidalgo-orihuela@lboro.ac.uk