THE APPLICATION OF VIRTUAL REALITY TO FOREIGN VOCABULARY LEARNING, MAKING USE OF THE LOCI METHOD.

By Vince McLeod.

2007
A thesis submitted in partial fulfilment of the requirements for the degree of Master of Arts in Psychology at the University of Canterbury
Supervised by Prof. Murray Simmonds
CONTENTS

ABSTRACT

1. INTRODUCTION AND LITERATURE SURVEY

1.1. MNEMONICS

1.1.1. What are mnemonics?

1.1.1.1. The peg-word method and similar mnemonics

1.1.1.2. The keyword method and its use within the loci method

1.1.2. How mnemonics work in general

1.1.2.1. Evidence for the effectiveness of mnemonics in general.

1.1.2.1.1. Range of subjects

1.1.2.1.2. Range of learners

1.1.2.2. Peg-word and similar mnemonics (such as the loci method)

1.1.2.3. The keyword method and its use within the loci method

1.1.2.4. How the loci mnemonic is thought to work

1.1.3. A short primer on memory

1.1.3.1. The study of recall

1.1.3.2. Different stages of remembering

1.1.3.3. Types of memory

1.1.3.3.1. Long-term memory

1.1.3.4. Increasing memory by adding more meaning

1.1.3.5. Failure to retrieve information from long-term memory

1.1.4. Some criticisms of mnemonic learning
1.2. THE LOCI METHOD

1.2.1. The steps of the loci method
   1.2.1.1. The encoding phase
   1.2.1.2. The recall phase (or retrieval phase)

1.2.2. Advantages of the loci method
   1.2.2.1. Evidence for its effectiveness
   1.2.2.2. A synergy with language learning
   1.2.2.3. Ease of organising information
   1.2.2.4. Other effects

1.2.3. Disadvantages of the loci method
   1.2.3.1. Time consumption

1.2.4. Miscellaneous points of interest
   1.2.4.1. Participant- or experimenter-generated keywords and loci?
   1.2.4.2. Other factors that can affect the effectiveness of the loci method

1.3. VIRTUAL REALITY

1.3.1. How it works
   1.3.1.1. A basic description
      1.3.1.1.1. Hardware
      1.3.1.1.2. Software
   1.3.1.2. Immersivity and presence
   1.3.1.3. Prior use of VR in education and the results

1.3.2. How VR could be used to aid foreign language learning
   1.3.2.1. Time saving
   1.3.2.2. Leverage of category learning
   1.3.2.3. Subject specific loci creation
1.3.2.4. Advantages concerning the loci themselves
1.3.2.5. Advantages concerning the social nature of language use
1.3.2.6. Visualisation of data
1.3.2.7. Miscellaneous

2. METHOD AND DESIGN

2.1. Purpose and goals of research
2.2. Participants
2.3. Hardware used in this experiment
2.4. The virtual environment
   2.4.1. The environment itself
   2.4.2. The words to be learned
2.5. Design
   2.5.1. Dependent variables
   2.5.2. Independent variables
2.6. Conditions
   2.6.1. Condition one
   2.6.2. Condition two
   2.6.3. Condition three
   2.6.4. Condition four
2.7. Hypotheses
   2.7.1. VR and loci method facilitation
   2.7.2. Organisation and loci method facilitation
   2.7.3. Colour coding and recall facilitation
   2.7.4. Other hypotheses; other interesting points
2.8. Materials and Apparatus
2.9. Procedure
   2.9.1. Trial levels
   2.9.2. Proper levels

2.10. Testing

2.11. Scoring

3. RESULTS

3.1. Overview of results and miscellaneous notes

3.2. Analysis of variance.
   3.2.1. Body of Swedish words
      3.2.1.1. Descriptive statistics
      3.2.1.2. ANOVA
      3.2.1.3. Virtuality
      3.2.1.4. Organisation
      3.2.1.5. Interaction
      3.2.1.6. Further analysis
   3.2.2. Gender of Swedish words
      3.2.2.1. Descriptive statistics
      3.2.2.2. ANOVA
      3.2.2.3. Virtuality
      3.2.2.4. Organisation
      3.2.2.5. Interaction
      3.2.2.6. Further analysis
   3.2.3. English words
      3.2.3.1. Descriptive statistics
      3.2.3.2. ANOVA
3.2.3.3. Virtuality
3.2.3.4. Organisation
3.2.3.5. Interaction
3.2.3.6. Further analysis

3.2.4. General observations

3.3. Mean recall scores and their correlates with other variables.
   3.3.1. Previous language ability
   3.3.2. Previous simulator experience
   3.3.3. Gender of participants
   3.3.4. Age of participants

3.4. Correlations between the performance measures themselves.
   3.4.1. Gender of words learned
   3.4.2. English and Swedish recall as a correlation

3.5. Demographic statistics

3.6. Miscellaneous points

4. DISCUSSION

4.1. Organisation

4.2. Virtual Reality

4.3. Limitations of the study/ further research.
   4.3.1. The type of words being learned.
   4.3.2. Higher-quality VR.
   4.3.3. Various types of mnemonic learning.
   4.3.4. Duration of experiment.

5. CONCLUSIONS
5.1. Hypothesis 2.7.1
5.2. Hypothesis 2.7.2
5.3. Hypothesis 2.7.3
5.4. Minor hypotheses

6. APPENDICES
6.1. Participant Consent Form
6.2. Participant Demographics
6.3. Participant Information Sheet
6.4. Introductory guide to the Loci Method

7. REFERENCES
ACKNOWLEDGEMENTS

More than any other event in my life, this thesis has made me understand and appreciate the depth of the social network that makes up much of who each of us are. With some understatement I can say that the challenges posed by this work, and the secondary challenges that arose because of it, were many, and difficult. Without the social network that brings me the happiness that I have sometimes been slow to appreciate, this undertaking could only have failed. I began this research as a man who had everything he needed to consider himself uncommonly blessed; some eighteen months later I had lost all of it, amidst a crushing black gloom the strength of which I never believed possible.

Firstly I want to thank Simon and Lizzie, for giving me a haven on that terrible June evening. You did this unthinkingly, unquestioningly, and with deep tenderness and patience. Never in my life have I needed help as much as I did at that time, and you were there for me.

I want to thank, in no order, Lindsay, Philby, Stewie, Alison, Ron, Warrick, Les, Kat and Matse. Lindsay for helping straighten me out. Philby for giving me a window back to reality. Stewie for helping me see that all was not lost, indeed far from it. Alison for not letting me forget that I was someone worthwhile. Ron for stumbling into the same car on the roller coaster, and helping me see how it should be ridden. Warrick for not allowing me to accept despair as a mindset. Les and Kat for a combination of most of these things. And Matse for giving me laughter to chase away the black crows. And to the many others whose kind words and commiserations kept me going throughout.

I also want to thank Murray Simmonds for his bottomless patience and kindliness, and Dean Owen for being the sounding board and inspiration that kicked the whole thing off.

And, of course, I want to thank my family, because even getting this far was a miracle, and you were the making of it.
ABSTRACT

An experiment was conducted to evaluate the potential for virtual reality-based enhancement of the mnemonic method of loci to yield improvements in performance in the learning of Swedish vocabulary. Groups of 15 participants were given a set of concrete nouns in the Swedish language along with their English translations. They were asked to learn both the gender of the Swedish words and their English translations, using one of four variants of the loci mnemonic. These variants comprised the presence or absence of a virtual environment and the prior grouping of the words into semantically related categories. The manipulation of virtuality did not provide statistically significant results, although the trends were in the direction of enhanced learning in the virtual environment. The manipulation of prior grouping into semantically related categories yielded statistically stronger effects than the manipulation of virtuality. The implications for future research using virtual-reality-based enhancements of the method of loci are discussed.
CHAPTER 1.

INTRODUCTION AND LITERATURE SURVEY

This thesis explores the interaction between foreign language learning, the loci method, and virtual reality (VR). Some of the literature that will be reviewed suggests that the nature of VR and the advantages offered by it could be of considerable use to students in language education. One of the goals of second language education is for the learners to build up a vocabulary of the word-translation associations between the target and the native language (van Bussel, 1993), and certain technologies and techniques can make this process more efficient. In particular, this thesis investigated whether VR could significantly contribute to the efficiency of the loci method of mnemonics.

Bellezza (1981) gives a definition of the term mnemonic: "A mnemonic device is considered to be a strategy for organizing and/or encoding information with the sole purpose of making it more memorable." The basic procedure when learning while using a mnemonic device is to associate the information to be remembered with one or more cognitive cuing structures. These are later recalled by the learner in order to cue the subsequent recall of the required information.

One of the major purposes of this research was to explore whether or not a virtual environment could be constructed that participants could use in order to learn a small number of Swedish language words as effectively as with the regular environments employed in the loci method, with a view to applying the results of this research to foreign language learning (See section 2.7 for hypotheses). If so, then it might be possible to expand this environment to encompass thousands of words, thereby creating a system that removes much of the work for the student
while retaining many of the benefits.

Atkinson & Raugh (1975) noted that many foreign language instructors believe that the most important element of learning how to communicate in a foreign language is not learning the grammar of a language, but rather learning the necessary vocabulary to partake in conversation and read a range of texts. This learning process is almost entirely a question of memorisation (Ott, Blake & Butler, 1976). Some researchers, for example Glenberg (2001), have commented upon the natural link between study of memory processes and study of second language learning, and the need for further research that combines these two fields.

The application of VR to language learning in general has not been widely researched, let alone its application to mnemonic learning or the loci method in particular. It will be argued in this thesis that there are good reasons to study this potential, by conducting some preliminary investigations into the potential benefits of mnemonics and VR to foreign language learning.

Another purpose of this thesis is to explore whether categorisation of the vocabulary of a second or foreign language, specifically organisation of the words into semantic groups before being presented to the participants, has any effect on recall scores.

The third major purpose is to explore whether there is any kind of interaction between VR and increased organisation in facilitating higher recall scores.

The study of mnemonic devices is often seen as something separate from general memory research, as if they were academic curiosities. Roediger III (1980) argued that "the principles underlying the use of mnemonics are simply more efficient variations of normal memory functioning and that clues to normal functioning can be found in the study of mnemonics."
Comments like this highlight the need for research such as the investigation that is reported in this thesis.

To a certain extent, this thesis might be viewed as an exploration in the field of applied linguistics. Second-language learning is in itself a subject of extensive scientific inquiry (Littlewood, 2004).

Research in this area is important because of the need for language educators to know which technologies work most effectively and which of them can be applied to the unique learning demands of their students. The current dearth of research in the area might be due to fashion, rather than a lack of value: Idoine-Shirai (2007) noted that “Mnemonics in foreign language acquisition seems to have fallen out of favour with current interests in research”.

As access to the technology required for computer-based foreign language learning is ever increasing (Zhao, 2005b), the importance of research into the application of such technologies to learning also increases. Personal Computers are not the only technology that provides access to language learning technology: Personal Data Assistants, MP3 players and DVD players, amongst others, also fill this role.
1.1. MNEMONICS

1.1.1. What are mnemonics?

Higbee (1977) defined a mnemonic system or technique as "any system or technique that aids the memory... Mnemonics refers in general to the process or technique of improving the memory, but typically is used to refer more specifically to rather unusual, artificial memory techniques." For a historical overview of the use of mnemonics, see Yates (1966).

The range of techniques that fall under the umbrella classification of 'mnemonic' are broad, but it includes rhymes (for example, "Thirty days hath September"), patterns (such as acoustic, spelling, visual, and others), acronyms, songs, as well as the link, keyword and loci methods (these three methods are discussed below). It should be noted that the phrase used in Higbee's (1977) definition "any system or technique that aids the memory" is perhaps a little too broad: in practice, as such a definition could cover the systematic use of learning drugs, specially constructed environments and the use of sound or music to put the learner in a relaxed state of mind. While these techniques might well aid the memory, no researcher has used the term 'mnemonics' to cover the study of such effects.

Note that in some contexts, the phrase 'mnemonics' is a synonym for, or overlaps with, "mental elaboration", "mediation", or "association" (Ott et al, 1976).

Higbee (1977) noted a common complaint levelled at the use of mnemonics: that many people who are exposed to the idea of mnemonic learning view their use as akin to a circus trick, rather than as an application of psychology. However, he emphasised that "mnemonics do not replace the basic principles of effective learning. Rather, they are based on the principles of learning."

The use of aids to assist the processes of memorisation is no more unusual than the use of aids to
assist anything else.

Higbee also noted that many people fail to see any practical use for such techniques. For this reason, this study will eschew the use of nonsense syllables (which have for many years been the basis of language learning research), and will instead use Swedish language vocabulary as its subject matter. In this way, it is hoped that the possibility that the learner will become bored or unmotivated, because of being unable to see the value in this form of learning intervention, will be lessened.

1.1.2. How mnemonics work in general

Higbee (1977) uses the analogy of a filing system to explain how mnemonic devices work. "Mnemonic systems can be viewed quite literally as mental filing systems. They allow you to store information in your memory in such a way that you will be able to find it and get it back out when you want it." Meaningful organisation is considered vital to more efficient memory (Shimamura, 1984), and with this in mind, mnemonics can be viewed as strategies for organising information.

In many mnemonic strategies, the use of visual imagery is central to the mnemonic's effectiveness, and this is especially true in the case of the loci method, as the learner must use the keyword method (see section 1.1.2.3.). Howard (1995) notes that the memory doesn't quite work like a filing system, as it is actually even more powerful than this - human memory can reorganise and search for information much faster and more efficiently than any filing system. The filing system analogy was also used by Gordon (2003), who pointed out that memory is enhanced when the pieces of information to be learned are stored in the right 'place' in the mind - "a niche in the mind where the information fits easily, can be extracted with little effort, and is clearly connected to other information". In section 1.2. this example is explained using practical
information.

The keyword method is considered to be an explicit, empirist, strategic method of learning, according to Howard (1995). In this context, explicit learning involves declarative knowledge (such as facts about history, and vocabulary items) and is prone to being forgotten. This contrasts with implicit learning, for example learning to ride a bicycle, which can be remembered for years without practice. Empiricist refers to the process of learning by extracting information from the environment, in the case of the investigation presented in this thesis, a sheet of paper or a virtual display (as opposed to learning information by comparing pieces of information already inside the learner's brain from which patterns and connections are drawn). "Strategic", in the context of Howard's conceptualisation, refers to learning according to a deliberately applied scheme.

The memorisation or learning of new information is easier if that information can be linked with memories and categories that the learner is already familiar with (Gordon, 2003).

Although not all mnemonic devices use visual imagery, this thesis is concerned with the loci method in particular, and so attention will be focused on the keyword method and the loci method, the former being a critical element of the latter.

1.1.2.1. Evidence for the effectiveness of mnemonics in general.

1.1.2.1.1. Range of subjects

It has been noted that mnemonics can be used effectively to assist the learning of foreign languages (Higbee, 1988). Thomas and Wang (1996) stated that "the keyword mnemonic is... an encoding strategy widely acknowledged to result in effective memory codes". De Beni (1998) noted that "many experimental subjects have shown how the performance of subjects who use these techniques may be several times better than those who do not". Roediger III (1980) found
that, in an experiment investigating the use of four different mnemonic techniques (imagery, the link method, the peg system, or the method of loci), all four groups using one of these techniques recalled more words from a 20-word list than did a control group. Rummel et al (2003) found that use of mnemonics increased the ability of participants to memorise central textual information when participants were tasked with learning a historical passage on aspects of human intelligence. Jones & Hall (1982) found the keyword method effective for pairing biographical information with the surnames of given people, and for pairing a technical word to its definition.

1.1.2.1.2. Range of learners

Levin et al (1982) reported that the keyword method was effective for teaching children words in their own language, even when those words had been specifically chosen as least suitable for keyword method learning (such as verbs with abstract meanings). Bower (1970) noted that there was much anecdotal evidence that the loci method works, and that in the years leading up to 1970 there were a number of controlled laboratory experiments that reported positive results. In the studies reviewed by Bower, those participants that were making use of the loci method frequently recalled two to seven times as much information as control subjects, an improvement he described as quite exceptional, if not staggering.

The use of mnemonics is not limited to college students or only those of above average intelligence. It has also been noted that a wide range of people have been able to use mnemonics to their advantage, including young children and the elderly (Kliegl, Smith & Baltes 1989; Kliegl, Smith & Baltes 1990; Verhaegen & Kliegl, 2000), and groups such as the intellectually disabled or gifted (Higbee, 1988). A literature review by Thomas and Wang (1996) on the subject of learning by using keyword images again reported that "the recent literature abounds with reports of further successful applications", including providing benefits to children with
emotional disorders, students with learning abilities, and the elderly.

The loci method has been tested, and found to be effective, in various studies, including the memorization of a prose passage regarding archaeological discovery (De Beni, 1988).

Due to the extent of the evidence that suggests mnemonics is more effective than rote learning, this study will not endeavour to compare the two strategies in the context of learning Swedish language vocabulary.

1.1.2.2. Peg-word and similar mnemonics (such as the loci method)

The peg-word method is sometimes known as the hook method (Paivio & Desrochers, 1981). The most common usage of this mnemonic is by memorising, in advance, a list of 10 or so nouns which rhyme with the first 10 numbers, for example one and bun, two and shoe, three and tree, and so on. The learner can then learn up to ten items by imagining a keyword image that links the keyword noun with the piece of information to be remembered. If a learner of a foreign language wanted to learn more words using this system they could simply learn more peg-words (Baddeley, 1998).

1.1.2.3. The keyword method and its use within the loci method

A review of the use of mnemonics in foreign language learning shows that the most commonly used mnemonic is that of the keyword (Higbee, 1988).

The keyword technique, when used to learn foreign vocabulary items, can be regarded as consisting of two distinct phases (Rough & Atkinson, 1975; Pressley 1988z). Firstly, the learner identifies a part of the vocabulary item that they wish to learn that can be associated somehow with their keyword. The basis for this association is usually a shared sound or a shared spelling
between the two words (Atkinson (1975) called this the *acoustic link*. Next, the learner links the keyword and the item's definition by encoding a meaningful association between the two, usually by imagining an interactive visual image, which is then known as a 'keyword image'. This was originally called the *imagery link* by Atkinson (1975), or an *interactive image* by Ott et al. (1973). An example of this, in the context of the Swedish language learning that is the subject of this experiment, would be a keyword image used to learn the word 'Bord' (English: 'table'). The keyword image need not approximate the full sound of the target word, but should approximate the main part of it (Raugh & Atkinson, 1975; Higbee, 1977; Paivio & Desroches, 1981). Recall of the target word is then facilitated by the keyword itself, which acts as a cue for the keyword image, which in turn acts as a cue for the target word. The mention of the foreign word will prompt recall of the keyword, which will in turn enable the learner to recall the keyword image, from which the English word can be determined (Paivio & Desroches, 1981).

According to Atkinson & Raugh (1975), "The keyword method can be described as a chain of two links connecting a foreign word to its English translation through the mediation of a keyword: the foreign word is linked to the keyword by a similarity in sound (acoustic link) and the keyword is linked to the English translation by a mental image (imagery link). One of the reasons why this technique works so well is because it allows the user to relate the material to be learned to their pre-existing knowledge, something that is known to aid learning (Howard 1995).

A literature review by van Hall & Mahn (1997) found that the keyword method enhanced immediate recall in the learning of vocabulary in a range of languages. Pressley (1977) found that use of the keyword method helped children to learn basic Spanish vocabulary items. Similarly, Raugh and Atkinson (1975) found that learners using visual imagery to associate given keywords scored 88 percent in a test of translating Spanish to English, whereas other learners who were given keywords but did not use visual imagery scored around 28 percent. Of particular
importance to this study, Higbee (1977) found that use of the keyword method increased recall both amongst learners tasked with recalling the English word given the foreign word and vice-versa. Atkinson (1975) found that even when learners using the keyword method had only learned to recall the English word upon the sighting of its foreign translation, they were still better at recalling the foreign word when given its English translation then learners who had used the rote method for the same task. Finally, Ott et al (1973) reported that learners using the keyword method to learn German vocabulary items performed consistently better than control groups who used the rote method.

The keyword method, because it links one piece of pre-existing information to a new piece of information that is to be learned, is ideally suited to learning foreign language vocabulary. Most of the literature relating to mnemonic use in the context of foreign language learning has investigated vocabulary learning instead of the other elements of linguistic or communicative competence, such as grammatical rules (for a detailed description of these terms see Littlewood, 2004). This thesis will also limit itself to the learning of vocabulary.

1.1.2.4. How the loci mnemonic is thought to work

In the context of the use of the loci mnemonic, it has been argued that easily recalled material relating to a three-dimensional, physical space that has been previously memorised serves as the 'filing system' (Higbee, 1977). The purpose of this material is to provide good retrieval cues through the series of spatial locations (Shimamura, 1984; Roediger III, 1980). These loci should have been selected with the purpose of being easy to remember. For this reason, many users of the loci method use locations that are familiar to them, such as the rooms, hallways, and other features of the inside of their house. The ability of any learner to recall desired information depends greatly on the quality of the retrieval cues that learner has been given. A familiar three-
dimensional space is an excellent retrieval cue as the learner can visualise the different retrieval
cues at will, and in any desired order. The loci method carries some major advantages over
simpler systems such as the hook method: one such advantage is the learner's ability to imagine
more vivid, richer loci and therefore to be provided with a more distinctly recallable cue than a
learner who is using the hook method would have available (Schoen, 1996).

The task for the learner is then to attach keyword images to the loci. The easy recall of the loci
ought to cue recall of the keyword image, and consequently of the desired vocabulary item
(Massen & Vaterrodt-Pluennecke, 2006) As long as the learner can remember the loci that they
have learned, they ought to have a good chance of recalling the keyword image, which they can
then decode to recall the target information. This particular form of the loci method is,
essentially, the keyword mnemonic with one extra level of complication (i.e., the addition of the
loci).

It is thought that the reason why humans find it advantageous to use a method like the method of
loci is because survival for many animals in their natural environment is dependent on their
ability to remember spatial location and to link information to these spaces. During the course of
human evolution, an early humanoid might have had to remember where a certain fruit was, or
where the tracks from a predator were seen, or which spot in a river was best for fishing. It is
feasible that natural selection would have selected for people who were able to remember this
type of information more easily, and therefore neural pathways might have developed that made
this method of learning more natural for humans.

Some of the effectiveness of the keyword method might be explained by the learner finding it
more interesting and exciting to imagine visual images such as keyword images, and this added
level of interest might increase their ability to pay attention, as many recall failures are thought
to occur not through forgetting but rather through a lapse of attention during the encoding process (Shimamura, 1984; Gordon, 2003).

### 1.1.3. A short primer on memory

#### 1.1.3.1. The study of recall

Psychologists have made use of several different methods to study recall; the method used in this thesis is termed paired-associate recall. In this method pairs of words are given to the learner, who is later presented with one of the pair and asked to recall the other (Higbee, 1977). Learning to associate the names of capital cities with their respective countries is an example of this type of learning. It could be argued that, to a large extent, so too is the learning of foreign language vocabulary (Ott, Butler, Blake & Ball, 1973).

#### 1.1.3.2. Different stages of remembering

Atkinson et al (1987) divide the act of remembering a piece of information into three distinct stages. Their first stage is the encoding stage, the second is the storage stage and the third is the retrieval stage. Gordon (2003) suggests that the inability to recall a piece of information can depend on a failure at any one of these three stages. It is important to note here that some learning strategies operate at differing stages of the remembering process (Harris, 1982). The mnemonic devices detailed in the first section of this thesis are all examples of strategies that work at the encoding stage of remembering. Other strategies, such as mentally going through every letter of the alphabet in the hope of finding a letter that would cue recall of a person's name, are examples of strategies that would be used during the retrieval stage.

#### 1.1.3.3. Types of memory

It is thought that the stages of remembering discussed above do not work the same way in all learning situations (Atkinson et al, 1987). The literature indicates that there is a meaningful
difference between the type of memory used for storing material in the order of seconds, and the type used to remember information for minutes or years. The first type is known as 'short-term memory', the second as 'long-term memory'. Since language learning is obviously dependent on long-term memory, this is the type that we will be most concerned with here.

1.1.3.3.1. Long-term memory

According to Atkinson et al. (1987), long-term memory remembers verbal materials in a system based on the meaning of the items, not on any acoustic or verbal manner. The finding that memory items must be meaningful to the learner in order to be remembered is not surprising. In many situations, the memory items that we wish to learn are meaningful but the connections between them seem random and arbitrary. This can create difficulties when attempting to learn large amounts of information.

1.1.3.4. Increasing memory ability by adding more meaning to the information (such as organisation)

Atkinson et al (1987) note that "Adding meaningful connections is a powerful memory aid." This phenomenon is the basis for all mnemonic learning. An excellent example of an experiment that illustrates this principle was performed by Bower (1972). Bower gave a number of participants a series of word pairs. The pairs were not meaningfully related to each other and their choice was arbitrary. One group of participants were told to learn each pair by thinking of a sentence that used both the words. The sentence, therefore, linked the two words in a meaningful way. The other group acted as a control and were not given special instructions about how to learn the word pairs. The first group produced much higher recall during testing; they recalled about 75 percent of the word pairs, whereas the control group recalled only 32 percent. It is known that if a list of words consists of words from several discrete categories, the words will be more easily
remembered than if the words are unrelated (Bower, Clark, Lesgold & Winzenz, 1969). It was thought by Bower et al (1969) that these phenomena occurred because the hierarchy itself was used by the learners as a retrieval plan for cuing recall.

There are other ways to add meaningful connections to verbal information, although the underlying principles remain the same. Another method is to make use of visual imagery, and it is this method that forms the basis of this experiment (see section 1.1.2.3.). Research has shown that learning in this manner is also far more effective than in any control group that does not make use of visual imagery (Roediger III, 1981).

Furthermore, the more deeply or elaborately the meaning is added to the information, the better the recall will be (Atkinson et al, 1972; Shimamura, 1984). One way to do this is to link the pieces of information together in meaningful ways (Gordon, 2003). This in effect creates a network of information that is stronger than the disparate collection of the memories that were used to create it (Shimamura, 1984). This is a crucial element of how mnemonic systems such as the loci method work: each piece of information is linked to the others so that any one piece might act as a 'cue' for several others. Deese (1959) found that words had a better chance of being recalled if they were presented alongside other words that shared a high level of association, as did Bousfield (1953). Deese (1959) then went on to hypothesise that the reason for this may have been that the words that were presented in association clusters were in effect presented with a higher level of organisation.

1.1.3.5. Failure to retrieve information from long-term memory.

A number of experiments support the suggestion that a learner's ability to successfully retrieve information depends on the quality of the retrieval cues they are given (Atkinson et al, 1972).
One factor affecting the quality of the retrieval cues is the number of items that have been associated with it (Atkinson 1987). The ability of a retrieval cue to assist the recall of information will decrease as the amount of information associated with that particular cue increases. This phenomenon is known as interference (Atkinson et al, 1987). Experiments have shown that if the interference is strong enough, a complete failure of retrieval can result (Anderson 1983).

1.1.4. Some criticisms of mnemonic learning.

Several commentators have mentioned that mnemonic learning actually requires the users to learn more information than they otherwise would if they had merely learned the information without reference to any special system (Higbee, 1977; Schoen 1996). The problem with this criticism is that it misses the point of mnemonic learning, which is not to learn less information but to make learning more effective. By analogy, an accountant who is learning to use a spreadsheet program might find that their progress is slower than if they had just added up the numbers using a calculator. However, when the accountant becomes proficient in the use of the program then he or she will be able to perform the task far faster than even the most skilled calculator operator will.

There is a misconception that learners who make use of the keyword method in paired-associate learning suffer from slower recall because they must first recall the keyword and the visual association before the desired word can be recalled. However, there is no evidence for this in the literature according to Atkinson (1975) and Higbee (1977).

Some students have complained that in the case of the more complicated mnemonic techniques, such as the method of loci, a large amount of effort must be made first before any learning of the target information can begin (Schoen, 1996). As will be argued in section 1.3. below, this is a
drawback of any system using mnemonics that could potentially be somewhat alleviated by the use of a VR system.

1.2. THE LOCI METHOD.

The loci method, otherwise known as the Roman Room Method, Journey Method, or Place Method, is an ancient mnemonic technique whereby the user imagines an extra-dimensional space and links information to it by visual association. The technique has been known since Classical times (hence the name 'Roman Room') and was popular amongst orators and politicians in ancient Greece and Rome (Moe & Debeni, 2005). Higbee (1971) dated use of the technique back to about 500 B.C. De Beni (1998) described the technique as follows: "The loci [method] consists in forming a mental structure represented by an ordered series of places and then separating the material to be memorised into units, imagining each unit as interacting separately with each place." Roediger III (1980) described the method thus: "In the method of loci, one takes a well-learned series of locations, such as a path one travels daily, and in learning a series of items converts each item into an image and deposits the image at some salient location along the path".

1.2.1. THE STEPS OF THE LOCI METHOD

1.2.1.1. The encoding phase.

The encoding phase of the technique involves three major stages. The first is actually a kind of preparatory work, and can be performed at any point before encountering the actual information to be learned (Bellezza 1981). This stage normally involves selecting discrete locations along a familiar pathway or inside a familiar space, and memorizing them to the degree that each could be recalled and visualized from memory with a high degree of detail. The Roman orators memorized their loci in strict serial order, a considerable aid when it came to memorizing speeches, as will be discussed in section 1.1.2.
The second stage consists of creating some form of visual image so that recalling the image acts as a cue for the recall of a certain piece of information (this is an example of use of the keyword method, and the resulting image is known as a 'keyword image'). This image can be created by either the participant or the experimenter (see section 1.2.4.1. for a review of the literature regarding the question of which of these two is more effective). For example, an English-speaking ancient Roman tasked with recalling the name 'Tiberius' might think of a bear wearing a bow tie (*tie* sounds like 'ti', *bear* sounds like 'ber'), so that when he recalled this visual image this would act as a cue for 'Tiberius'.

The third stage involves association of this visual image with one of the pre-selected locations. This is usually made by forming a second visual image, i.e. by imagining some form of interaction between the location and the first visual image. These three stages are repeated as the user wishes (Moe & Debeni, 2005). Usually the locus is imagined first.

1.2.2.1. The recall phase (or retrieval phase).

In the recall phase of the technique, the user retraces the pathway through the loci in their imagination. Using the visual form of each locus as a cue, the user can recall the visual images that were associated with them, which in turn act as cues for the information that the visual images were intended to represent (Moe & Debeni, 2005; Massen & Vaterrodt-Pluennecke, 2006).

1.2.2 ADVANTAGES OF THE LOCI METHOD.

1.2.2.1 Evidence for its effectiveness.

The loci system has been shown to improve word recall ability in all age groups (Rose & Yevasage, 1983). A literature review by Kliegl et al. (1989) revealed much evidence to support
the conclusion that despite the high cognitive load required by the loci method, older people were capable of using it to greatly increase their recall ability.

1.2.2.2 A synergy with language learning.

Any mnemonic strategy based on imagery is generally more effective if the material is presented orally (Debeni & Moe, 2003a). This 'oral presentation effect' is thought to exist because when the material to be learned is presented in written form, the process of having to read it interferes with the process of encoding the information by way of 'overloading' the visual channel of sensory input (Moreno & Mayer, 2002). This interference does not occur when the material is presented audially (Moe & Debeni, 2005).

The oral presentation effect has an obvious synergy with language learning: that the student might be able to hear the vocabulary pronounced at the same time as they encode the words using imagery. Such an advantage would be easy to exploit using the loci method, as the auditory effect would not interfere with the visual. An added reason to study the applicability of VR to the loci method is that adding an audial input to a virtual environment would be very easy to achieve, requiring little programming. Moreno & Mayer (2002) found, in an experiment using VR to teach students about botany, that information narrated to the learners was better recalled than information that is merely read off a screen.

The loci method has been shown to be of assistance when it provides cues to students who wish to learn pieces of information in a particular order (Roediger III, 1981). Although the learning of vocabulary generally does not require a fixed word order, in certain cases an order could be imposed to aid the learner. For example, in the same way that the experiment reported in this thesis uses the colours pink and blue to code for the two different genders in Swedish, more colours could be used in order to create a system for learning the different forms of irregular
nouns. The colour white might code for the infinitive, yellow for the past participle, orange for the gerund, and so on. This could be very useful for learning languages that have large numbers of irregular verbs, such as German or English.

One possible manipulation of mnemonics, based on the findings of Atkinson & Raugh (1975), would be easily facilitated by the use of VR. This would be to provide the participants with the option of requesting a keyword if there is no obvious choice that appeals to them. Having a computer-based system would make it easy to program multiple keywords that could be cycled through sequentially if the learner was not satisfied with the keyword or if they wished for another one for whatever reason.

1.2.2.3. Ease of organising information.

The loci themselves make excellent units to impose organisation upon the material. The study by Bower et al. (1969) that found highly increased recall of words if they were previously organised into a hierarchical structure, shows clear parallels with the loci method. For example, a row of rooms could be imagined for each level of the structure.

The fact that groups of loci, such as rooms, can be made larger or smaller at will, might allow the learner to select an optimum number of words to be learned in each room at that particular level of organisation, as it is known that the more items that the learner tries to remember, the more likely it is that errors will be made (Howe, 1969).

1.2.2.4. Other effects.

Cornoldi (1985) noted that research on the subject of mnemonics ought to look for gains beyond pure recall score in the context of the experiment, as there may be other benefits such as the participants' increased confidence in their memory abilities.
Although the loci method is most commonly used as a memory aid for ordered lists, (Herrmann et al, 1988) it works well for unordered lists too, as noted by de Beni (1998): "The fundamental key to many memory techniques is the use of a structure which organizes the material to be memorized, facilitating retrieval by reference to that structure". This suggests some potential uses for a computer-based simulation of the loci method. Firstly, a computer could readily be programmed to create the necessary structure. Secondly, if the organisation did not suit the learner for some reason, it could easily be restructured with very little time or effort.

Massen & Vatterodt-Pluennecke (2006) found that learners using the loci method suffered less from proactive interference than learners using the link method or rote rehearsal did.

1.2.3 DISADVANTAGES OF THE LOCI METHOD.

1.2.3.1. Time consumption.

The major problem with the loci method for beginners is that preparation and use of the loci is time-intensive. The user must themselves decide which set of locations is the most appropriate (from a list of perhaps dozens of familiar places), then decide which specific loci will be used for encoding, then memorise these loci with a high degree of fidelity.

If one should wish to learn a foreign language, starting from no knowledge and passing through to fluency, constructing a set of loci that contains one location for each of the several thousand words that a speaker is required to learn is a difficult and time-consuming task. This may serve to discourage students from using what appears to be an extremely powerful mnemonic technique.
1.2.4. MISCELLANEOUS POINTS OF INTEREST

1.2.4.1. Participant or experimenter-generated keywords and Loci?

There is some evidence to suggest that subject-generated loci are generally more effective than experimenter-generated loci (Wang, Thomas, & Ouellette, 1992, Thomas & Wang, 1996). The basis for this claim is that experimenter-provided keywords might not harmonise fully with the subject's mode of learning.

Atkinson & Baugh (1975) suspected that supplying the keywords is helpful for the beginner, but that this becomes less useful as the learner gains familiarity with the language and the keyword method.

Briggs, Hawkins & Crovitz (1970) found that self-production of loci is not essential, and participants are in fact able to effectively use a modified form of the loci method without having prior exposure to the loci that were used. Moreover, certain evidence suggests that for keyword mnemonics to be effective, both keywords and the corresponding interactive images should be supplied by the experimenter (Troutt-Ervin, 1990). The rationale for this is that the participant, "thinking on the fly", might not be able to select keywords that are as effective as those the experimenter might have discovered over a much longer time, through experience, practice and research. In addition, the participant would be able to concentrate more fully on encoding the material rather than on thinking up the keyword images for themselves. Pressley & Levin (1980) found evidence that for vocabulary items with easy-to-generate or obvious keyword images, it matters little if the keywords are participant or experimenter-generated.

The author of this thesis considers the question of the superiority of participant or experimenter-supplied cue images/words to be an undecided issue, a view shared by Atkinson & Baugh (1975). It should be noted that in practice, experimenter-generated loci are more commonly used
due to the ability to control variables such as the type of pathway chosen (Moe & Debeni, 2005),
and as a way of controlling for factors such as individual differences in participant creativity.

In this study, the participant retained most of the control in creating the visual images that link
the English with the Swedish words - the only restrictions were placed by the map outlining the
loci themselves. In conditions one and two, the loci were presented to the participant in the form
of a map. In conditions three and four, the experimenter created the loci and presented these to
the participants by means of the VR system. In both of these conditions, the participants
themselves created the keyword image.

1.2.4.2. Other factors that can increase the effectiveness of the loci method.
The greater the discriminability of the visual cues, the lower the chances that some of the images
will interfere with each other in the user's memory store, and therefore the higher the chances
that the images will be able to be recalled without error. VR would offer an advantage here, since
it may be possible for a computer to generate a virtual environment with a much higher degree of
detail than could an average person who is just using their imagination.

1.3. VIRTUAL REALITY
1.3.1. HOW IT WORKS.
1.3.1.1. A basic description.
Burdea and Coiffet (2003) offer the following definition of VR, it being: “a high-end user-
computer interface that involves real-time simulation and interactions through multiple sensorial
channels”. Kalawsky (1993) used the definition "Virtual environments are synthetic sensory
experiences that communicate physical and abstract components to a human operator or
participant". Pares & Pares (2006) emphasised the importance of making clear distinctions
between terms that had otherwise been used with little care in the literature, in particular VE, VR,
telepresence and cyberspace. They considered the term VE to be equivalent to the model of a simulation, whereas the term VR was equivalent to the act of simulation itself. "VR is the real time experience a user can have of a VE". This thesis followed these guidelines.

The basic VR system has as its core a computer that is ultimately responsible for the quality and the believability of the virtual world (Vince, 2004). This computer will generate graphics in real-time and present them to the user in such a way that the images they see appear quasi-real (Vince, 2004). The 3D graphics of a VR system respond to the user's actions, and in this way allow the user to experience a sense of presence in that environment (Reid, 2005). Technically speaking, the use of the term 'VR' ought to cover signals that stimulate the other senses of the participants (Durlach & Mavor, 1995), but in this experiment the virtual experience is limited to visual interaction only. Bricken & Byrne (1993) considered a VE primarily to be a place where a learning experience can occur.

Psotka (1995) called for a program of research to discover the instructional conditions that best suit VR learning, and it is in this spirit that this thesis proceeds.

1.3.1.1.1. Hardware
The demands made by Virtual Reality upon computer hardware are considerable. A computer tasked with generating a high-quality VE will need very large physical memory, a powerful graphics chip, a high-bandwidth mass storage device and a high-speed set of input and output peripherals (Durlach & Mavor, 1995). One of the most important requirements for creating a realistic VE is having a high frame rate. A frame rate is "The number of still images that must be presented per second to provide the illusion of continuous motion." (Durlach & Mavor, 1995). The value of the frame rate depends chiefly on the amount of data needed to be processed by the graphics chip/chips (in general, the more realistic-looking the VE, the more data involved), and
the power of the graphics chip itself.

There are generally considered to be two different types of VR (Kalawsky, 1993): the immersive kind, and the non-immersive, desktop-based kind, although to some extent there is a continuum of differing levels of immersivity (Jacobson, 1994). Fully immersive VR is sometimes known as inclusive VR (Jacobson, 1994). An immersive Virtual Reality system involves a three dimensional environment which is presented to the user by means of a head mounted display that can use stereoscopic vision to create the illusion of three-dimensional space. Features such as auditory and tactile sensations are common but optional. The user is able to navigate through this environment using a variety of different interfaces, such as a mouse or a wand, or by walking (Vince, 2004, Durlach & Mavor, 1995).

A non-immersive system displays the environment on some form of small screen. Some researchers, e.g. Moreno & Mayer (2002), regard this as a low-immersivity learning environment, as opposed to "non-immersive", by which they mean something like text presented on paper. This experiment used such a method, with the environment displayed on the 17 inch monitor of a desktop computer. It is unclear if the advantages (if any) of using a more immersive method would outweigh the disadvantages of having to learn to use the equipment.

This experiment used a non-immersive system, in front of which the participants were seated on a comfortable office chair with two arms. The computing demands were low enough that a regular PC belonging to the experimenter was sufficient to run the experiment. For more information about the hardware specifications, see section 2.3.

1.3.1.1.2. Software

There is a wide range of software that can be used to display the contents of the virtual world on
the hardware. This experiment used the cheapest currently available solution. The software known as the *Cortona browser* can read scripts written in Virtual Reality Modelling Language (VRML) and display them through various internet browsers. In this experiment, the Cortona plug-in was used through the popular open source *Mozilla Firefox* browser (see www.mozilla.org) to display the virtual worlds. One of the features of this plug-in was the ability to set the display to full-screen mode, which the participants could navigate using a mouse.

The author learned the VRML programming language as part of the background research for this thesis and applied this knowledge to write the VRML scripts that displayed the virtual environments in conditions three and four (see 2.6.3 and 2.6.4.).

1.3.1.2. Immersivity and presence

The value of greater levels of immersivity to learning tasks is unclear. Moreno & Mayer (2002), in a study of students tasked with learning about botany through a multimedia game found no significant difference in data retention or transfer based on the media through which the information was presented, although students that used a Head Mounted Display (HMD) reported feeling a higher level of presence. Pslotka (1995), in an introduction to VR and its possibilities to education and training, considered immersion to be the key benefit of VR.

It has been theorised (Moreno & Mayer, 2002) that a greater sense of presence ought to lead to a higher sense of interest in the learning environment and therefore a higher level of motivation, which should in turn promote learning. However, in that same paper, Moreno & Mayer (2002) found no significant effect of immersivity on learning. In any case, Baumann & Sayette (2006) noted that much of the research involving VR uses a low-immersivity setup that is based on a standard computer monitor or a larger projection screen.
1.3.1.3. Prior use of VR in education and the results.

VR has been successfully used in a range of previous educational tasks (Tichon et al, 2004), of which only a few will be mentioned here. Some of the most common uses are in practical training, such as flight simulators (Bricken & Byrne, 1993), road crossing training in children (Thomson et al, 2005), police training to make potentially life-threatening decisions (Hamit, 1993), giving students of psychology the experience of what it is like to suffer from certain mental disorders, such as schizophrenia (Tichon et al, 2004).

Similarly, VR is often used in clinical trials and treatment (Kalawsky, 1993).

An extensive literature survey by Zhao (2005) reported that students who used 'digital multimedia technologies' (of which VR is one) acquired foreign language competence in a range of areas more effectively than learners using traditional methods did.

There appears to be no previous literature regarding the application of VR to the loci method of mnemonics. This is perhaps not surprising, given the relatively small amount of literature devoted both to VR and to mnemonic learning.

1.3.2. HOW VR COULD BE USED TO AID FOREIGN VOCABULARY LEARNING.

Most of the advantages to be described below are pertinent to computer-aided learning rather than to VR. However, since VR is the medium through which this research will investigate potential benefits, these advantages are to be understood in the context of VR.

Durlach & Mavor (1995), in their review of some of the potential uses of VR, consider its application to education to be of critical concern. It is in this spirit that this thesis attempts to explore some potential applications.
1.3.2.1. Time saving.

It has been shown that arranging words in a distinctive visual pattern results in better recall performance than does presenting the same word lists in a pattern that is always the same (Bellezza, 1986). A major problem with applying this potential advantage is the time required to generate distinctive visual patterns. A student wishing to do this would ordinarily have to generate a number of distinctive patterns by applying their own imagination. From the student's perspective, this activity is inefficient and could be performed better and faster by a computer.

It is estimated that the number of words a student must know in order to be able to read at a level equivalent to first year university study is 3,000 (van Hell & Mahn, 1997). Creating a map, for example on paper, for an environment with sufficient space to hold keyword images for this many words would be extremely time consuming. A computer, however, could be programmed to generate an appropriate map in seconds. In effect, the computer would carry much of the memory load that would otherwise be the task of the human learner.

1.3.2.2. Leverage of category learning.

Higbee (1977) emphasised the importance of organisation in learning new material. The better the information in question is organised at time of encoding, the easier it will be to retrieve. Bower et al. (1969) found that learners who were presented information in logical categories were able to memorise it far faster than others who saw the same information without categorisation.

It is easy to construct virtual environments in ways that students could take advantage of categorised learning. For example, a virtual world intended for the use of learning Spanish vocabulary might generate some of the locations with the colour blue strongly present, and some
others with the colour red, with the intent that the students would (for example) associate words of the masculine gender with the blue-coloured locations and words of the feminine gender with red-coloured ones. The colour of the locus that was associated with the visual image coding for the word to be remembered would reveal the gender of the word in question. This phenomenon was included for experimental investigation in this research by requiring participants to recall the gender of Swedish-language words (see hypotheses and results). This advantage has been exploited previously by researchers using other mnemonic techniques, such as the hook method (Paivio & Desrochers, 1981).

Although a user might be able to generate a set of loci that took advantage of the obvious example of noun genders, a virtual environment created by experts in the target language could be organised according to features of the language of which a beginning language learner would probably not be aware. For example, a program designed to assist learning Swedish might create a set of five distinctly coloured rooms for each of the five different plural noun endings. A language learning system that made use of a computer-generated virtual environment to order the vocabulary items would be able to impose a large amount of organisation and structure on the material for the student's benefit. Another advantage with a computer-based system is that the users would be able to reorganise the vocabulary items if they default array was not to their liking. The fact that digital information is recorded in bits allows for unprecedented ease in re-organising information to suit the needs of the learner (Zhao, 2005b).

1.3.2.3. Subject specific loci creation.

The use of VR to create loci for vocabulary learning has a range of potential benefits that will not be examined in this experiment. For example, a user wandering the virtual environment to learn vocabulary words in French could interact with virtual avatars that represent historical figures such as Camus and de Gaulle, they could click on items within the virtual environment to call up
interesting information about the culture and language, and the virtual environment itself could be constructed with the aim of simulating modern France. The combination of these factors would probably increase the student's motivation to learn by immersing them in a 'French' environment, as well as capitalising on any advantage gained by having a more immersive environment. Such factors were mentioned by Durlach & Mavor (1995) as some potential uses of VR in education. Hamit (1993) mentioned that one of the greatest advantages in VR-based education was the ability to simulate real-life situations in conditions that would be expected in real-life.

1.3.2.4. Advantages concerning the loci themselves.

For maximum effectiveness of a set of loci, each locus should be well illuminated, at similar distances from each other and of comparable size (Moe & Debeni, 2005). A beginning user, limited to imagining an environment familiar to them, may have difficulty selecting a set of loci that fulfils these criteria, or might be unaware that such criteria are important. A computer could be easily programmed to handle these tasks; in fact, programming a virtual environment where the loci were at a similar distance from each other and of similar size might well be much simpler than doing otherwise.

There is some concern that with repeated use of the loci method, the learners might come to suffer from proactive interference caused by the repeated use of the individual loci (Messer & Vaterrodt-Plunnecke, 2006). A computer-based system that generated large numbers of loci would overcome this limitation, as learners would not need to use the same loci more than once.

It has been mentioned above that an important element of the loci is that they are easy to remember. Although a computer-generated virtual world will not initially be familiar to the user, it could be created in a manner that would be easy to remember for other reasons. Higbee (1977)
mentioned certain principles of memory improvement, such as meaningfulness, organisation, association, visualisation and attention and interest. Although the principle of organisation will be investigated in this thesis, a virtual world could be constructed in a way that could take advantage of these other principles.

Previous research into the effectiveness of the loci method has suggested that learners can use the method much more effectively when they can actually visit the loci they are using, and from there look around and see how they connect with the others (Schoen, 1996). If a learner wished to learn, say, 3,000 words, then a VR-based set of loci would be far more practical than walking around in reality.

1.3.2.5. Advantages concerning the social nature of language use.

One major potential use of VR is in the use of networks, as opposed to simply a human-computer interface (Durlach & Mavor, 1995). A VR network would allow real-time communication between teachers and students via the Internet, as well as between learners and native speakers (Yang & Chen, 2007).

If learners can use keyword images to learn the vocabulary items, then they could use mnemonics and keywords that were generated by other learners and shared online. This kind of advantage would be predicted by the results that were found by Campos et al. (2004), who reported that (at least for high-vividness words) keyword images supplied by peers were more likely to be recalled than keywords that were generated by the participants themselves. This is relevant to the finding by Ellis & Beaton (1993): keywords must share a reasonably high degree of acoustic similarity with the words to be learned to serve as effective cues. It is easy to imagine that a language learner using the keyword method might not think of the most effective word immediately, but would be able to use it in the construction of a keyword image if it were
presented to them. A Virtual Environment could be constructed that suggested keywords, and automatically changed these based on which other keywords other users found most effective.

Even better would be an improved version of the Conversim(TM) system developed by *Interactive Drama Incorporated* (Harless, Zier & Duncan, 2005). This system allowed students to engage fictional native speakers in extensive dialogues using natural spoken language. The fictitious native speaker was portrayed to the learner on a desktop monitor. The student was able to ask the native speaker a series of questions about his culture, interests and military intelligence, and the program would respond according to a series of responses that existed in the program database. This system was described as a 'virtual immersion experience'. A future version of a similar program might involve a head-mounted display and a larger virtual world through which a learner could navigate and therein converse with a range of various 'native speakers' about a number of different subjects.

**1.3.2.6. Visualisation of data.**

Erickson (1993) claimed that visualisation was one of the most promising future strategies for making more effective use of data. The purpose of advanced levels of visualisation is to present data in a way that makes more sense to the human viewer. This has some likely interaction with the categorisation effect, because the viewer, if they had the ability to see the data more clearly, would presumably wish to do so in more logical categories. Not only is VR capable of displaying massive, multidimensional sets of data but it can do so with high levels of user interaction (Kalawsky, 1993).

**1.3.2.7. Miscellaneous.**

A computer-based learning system would be able to take advantage of the spacing effect. A useful description of this effect is given by Delaney & Knowles (2005): "The spacing effect is
the finding that memory performance is better when study repetitions are separated by other events or items (spaced items) than when repetitions immediately follow one another (massed items)"; see also van Bussel, (1993). A feature could quite easily be written that presented words to the learner in an order that separated study repetitions in a manner consistent with these findings. This is essentially another way that VR could be used to reduce the effect of interference between the learned items. Related to this is the idea that there might be an empirically optimum number of vocabulary items in each 'study unit' (van Bussel, 1993).

In the context of learning lists of foreign language vocabulary, the greater the 'imageability' of a word (the degree to which the word elicits a clear mental image) the more likely it is to be recalled (Ellis & Beaton, 1993). The rate of improvement of modern computer graphics would allow for increasingly greater image detail.

A factor that should not be underestimated is that the enjoyment that participants in VR tend to derive, and that this enjoyment might motivate them to continue entering the learning environment when other motivating factors have lost their power. Bricken & Byrne (1993) reported that, in an opinion survey of students who had used VR technology to create and interact with their own virtual worlds, 95% of them said they would rather go into a virtual world than see a virtual world on a computer screen, and 96% preferred this to watching television.

Tichon et al. (2004) remarked that one of the obstacles to providing real-life training is often the cost. In the context of foreign language learning, the costs are obvious, especially if the learner needs to travel a large distance to immerse themselves in a native-speaking environment. While a classroom might provide some element of simulation, it would still not be very close to actual (or virtual) immersion. Related to the cost aspect is the problem of wear and tear on places like museums and archaeological sites, which, as Hamit (1993) pointed out, could be avoided if these
areas were explored virtually instead.

A computer-based system with access to the Internet would be able to take advantage of all the existing information that relates to foreign language learning, which would also include websites and audio and video streams in the target language (Zhao, 2005b).

Having a prepared set of loci that are displayable on a monitor might also remove one of the psychological barriers that prevents learners from experimenting with the loci method, namely the daunting size of the loci the learners must commit to memory (Schoen, 1996).

**Summary**

It is clear that several potential benefits to the use of VR in simulating the loci method for use in foreign language learning exist. A necessary initial step will be to determine if a smaller-scale use of VR, like the one in this experiment, is at least equal to non-VR methods in producing recall scores of foreign vocabulary items.
CHAPTER 2.

METHOD AND DESIGN

2.1. Purpose and Goals of Research

The learning of foreign languages is a common academic pursuit and its importance need not be emphasised. It would be expected that any significant improvement in this area could be of benefit to large numbers of people. Therefore, one of the aims of this research was to reflect real-world conditions of foreign language learning. Cornoldi (1988) remarked that "Memory tasks scarcely reflect memory requirements of everyday real life." The idea behind this research was to discover information that could be applied to actual language learning environments. Cornoldi (1988) suspected that one reason for the declining interest in mnemonic research was the suspicion that mnemonics could not be used in most situations of daily life.

The primary purpose of this research was to evaluate, in the context of vocabulary learning, the effectiveness of a computer-generated virtual environment as a substitute for a participant-generated one by making use of the loci method of mnemonics. We explored whether or not a VR-based system can create an environment that is useful for the user who wishes to use this environment as an associative instrument for learning the vocabulary items (See hypotheses in section 3.7.). If so, then it should be possible to develop a system in which the user could scale up the environment so that it contained a separate locus for each of thousands of vocabulary items.
2.2. Participants

60 participants, 20 males and 40 females, were recruited from the University of Canterbury. Ages ranged between 16 and 51 years, with a median age of 22 and a mean age of 24.03. The ages for the males ranged between 18 and 38, with a median age of 25 and a mean age of 25.05, whereas the ages for the females ranged between 16 and 51, with a median age of 21.5 and a mean of 23.53. The participants were randomly assigned to each of four conditions, with 15 in each condition. Participants were recruited via e-mail, from undergraduate level laboratories, and from advertisements on notice boards throughout the University. The participants were screened to ensure that they had no knowledge of Swedish nor of any other West Scandinavian language (defined here as Swedish, Norwegian, Danish, Icelandic or Faroese, and any dialect of these languages). Participants were excluded if they had knowledge of any of these languages. The reason for excluding the participants for having knowledge of the non-Swedish West Scandinavian languages was because these languages are very similar to Swedish, and finding the necessary number of Swedish words that were different to the words in all these languages (as well as English and German) would have been difficult. The gender and age of the participants was not expected to be a significant factor, and was not controlled for.

The participants did not necessarily have English as a first language, but they all spoke and read English fluently.

As compensation for the time taken to participate in this study, the participants were given one $5 grocery voucher, with a 50% chance of winning a further $5 grocery voucher at the conclusion of the experiment. Participants were also given a choice from a range of confectionery, which included a popular brand of chocolate egg, as well as various brands of chocolate biscuits. These procedures complied with the requirements of the University of Canterbury Human Ethics Committee.
All of the participants who were recruited completed the experiment.

For further demographic statistics regarding the participants, see section 3.5.

2.3. Hardware used in this experiment
All simulated events were generated using a desktop computer, and presented to the participants via a desktop monitor. The relevant specifications of the computer were as follows:

Viewsonic 17” (43.2cm) CRT monitor. The screen resolution was at all times 1024 x 768 pixels.
A 2.0-GHz AMD Athlon 2400+ processor.
Microsoft Windows XP operating system with Service Pack 2.

The refresh rate of the display was 60 Hz. An optical two-buttoned mouse with scroll wheel was used by the participants to navigate the VE.

The system described here was roughly standard for contemporary medium-level home computers.

2.4. The Virtual Environment.
2.4.1. The environment itself.
The environments in conditions one and two were not virtual. The participants were presented with paper maps that simulated the virtual rooms that were to be used in conditions three and four, respectively. The maps were laminated A3 paper that presented the participants with a
birds-eye view of the VE.

The virtual environment consisted of eight textured rooms. In condition three, these were not colour-coded, nor were the items to be remembered arranged in any special pattern (see section 1.1.3.4.). In condition four, four of these rooms were colour coded blue and the other four pink. For the blue rooms, this meant that the walls of these rooms had a distinctive blue colour, although this blue was of varying hues (for the purposes of greater discrimination), and the same was true for the pink rooms. The rooms were navigable using a mouse or a keyboard, although it was expected that the participants would find the navigation by mouse much easier, and in practice all participants chose the mouse. Navigation was possible either by translational (forward or backward) or rotational (clockwise or anti-clockwise) movement. These two types of movement are typical for user control in a VE (Tlauka, 2007).

2.4.2. The words to be learned.

Displayed in 32 discrete locations throughout the virtual environment was text that took the form of a Swedish word that the participants were tasked with trying to remember by using the loci method. These words were selected by the experimenter from a Swedish lexicon with assistance from his wife (a native Swedish speaker), and were screened to ensure there is no obvious linguistic similarity between the word and its translation in either English or German.

The 32 words to be learned consisted of eight groups of four words each. The eight groups shared semantic similarities that marked each of them as a distinct group (see section 2.2.2. for some expected benefits of this). For example, one of the groups contained solely words that related to travel. All four words in each group were nouns. This choice of word class was motivated by the fact that the differing word classes have differing levels of ease of learning, and the nouns are the easiest of all (Ellis & Beaton, 1993). It was thought that choosing only nouns
would lower the influence of chance factors in finding effective keyword images. Furthermore, the words were all concrete nouns, as opposed to abstract ones, at the existing literature suggests that such words are more effectively learned by using the keyword method (Paivio & Desrochers, 1981).

Swedish was chosen as the target language due to the experimenter’s familiarity with it, and due to the low probability of participants having prior experience with a language spoken by so few people. (The current number of native speakers of Swedish is variously estimated to be around 8-9 million, depending on whether Skånska, the language spoken in the south of the country, is considered a dialect or a separate language.) Nonsense words were not used, as one of the goals of the experiment was to replicate real-world learning conditions, and it was felt that learning nonsense syllables might negatively affect the participant’s motivation.

The Swedish words that students were required to learn were screened to ensure that they were not especially similar to their English or German equivalents. In cases where a proposed Swedish word was judged by the experimenter to be similar in sound or spelling to its German equivalent, it was omitted and a new word selected. The rationale for screening the Swedish words to ensure dissimilarity with German was due to expected high levels of German language proficiency amongst the student population.

Table 2.4.2.1. below shows the grouping of the words in conditions one and three.
Table 2.4.2.1. The words that the participants were required to learn in conditions one and two.

<table>
<thead>
<tr>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>en tallrik = a plate</td>
<td>en groda = a frog</td>
</tr>
<tr>
<td>ett hängsle = a suspender</td>
<td>en linjal = a ruler</td>
</tr>
<tr>
<td>ett blixtlås = a zipper</td>
<td>ett träd = a tree</td>
</tr>
<tr>
<td>ett löv = a leaf</td>
<td>en orm = a snake</td>
</tr>
<tr>
<td>en laganda = a team spirit</td>
<td>ett odjur = a monster</td>
</tr>
<tr>
<td>ett matos = smell of cooking</td>
<td>en pärn = a folder</td>
</tr>
<tr>
<td>ett kryp = an insect</td>
<td>en sko = a shoe</td>
</tr>
<tr>
<td>en smörgås = a sandwich</td>
<td>en skytt = a signpost</td>
</tr>
<tr>
<td>ett moln = a cloud</td>
<td>en gata = a street</td>
</tr>
<tr>
<td>en grindvakt = a wicketkeeper</td>
<td>ett skafferi = a pantry</td>
</tr>
<tr>
<td>en dator = a computer</td>
<td>en hiss = an elevator</td>
</tr>
<tr>
<td>ett hjul = a wheel</td>
<td>en rännsten = a gutter</td>
</tr>
<tr>
<td>ett hav = an ocean</td>
<td>ett fordon = a vehicle</td>
</tr>
<tr>
<td>ett flygplan = a plane</td>
<td>ett mål = a goal</td>
</tr>
<tr>
<td>ett poäng = a point</td>
<td>ett tåg = a train</td>
</tr>
<tr>
<td>en mössa = a cap</td>
<td>en blyertspenna = a pencil</td>
</tr>
</tbody>
</table>

Table 2.4.2.2. below shows the grouping of the words in conditions two and four. Note that these groups were not labelled as such when they were presented to the participants.
Table 2.4.2.2. The words that the participants were required to learn in conditions two and four.

<table>
<thead>
<tr>
<th>TRAVEL</th>
<th>FOOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ett hjul = a wheel</td>
<td>ett skafferi = a pantry</td>
</tr>
<tr>
<td>ett tåg = a train</td>
<td>ett matos = smell of cooking</td>
</tr>
<tr>
<td>ett flygplan = a plane</td>
<td>en tallrik = a plate</td>
</tr>
<tr>
<td>ett fordon = a vehicle</td>
<td>en smörgås = a sandwich</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NATURAL WORLD</th>
<th>ANIMALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ett löv = a leaf</td>
<td>ett kryp = an insect</td>
</tr>
<tr>
<td>ett träd = a tree</td>
<td>ett odjur = a monster</td>
</tr>
<tr>
<td>ett hav = an ocean</td>
<td>en orm = a snake</td>
</tr>
<tr>
<td>ett moln = a cloud</td>
<td>en groda = a frog</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SPORT</th>
<th>INFRASTRUCTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ett mål = a goal</td>
<td>en hiss = an elevator</td>
</tr>
<tr>
<td>ett poäng = a point</td>
<td>en rännsten = a gutter</td>
</tr>
<tr>
<td>en grindvakt = a wicketkeeper</td>
<td>en gata = a street</td>
</tr>
<tr>
<td>en laganda = team spirit</td>
<td>en skylt = a signpost</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OFFICE EQUIPMENT</th>
<th>CLOTHING</th>
</tr>
</thead>
<tbody>
<tr>
<td>en linjal = a ruler</td>
<td>ett hängsle = a suspender</td>
</tr>
<tr>
<td>en blyertspenna = a pencil</td>
<td>ett blixtlås = a zipper</td>
</tr>
<tr>
<td>en dator = a computer</td>
<td>en sko = a shoe</td>
</tr>
<tr>
<td>en pärn = a folder</td>
<td>en mossa = a cap</td>
</tr>
</tbody>
</table>
en rännsten
a gutter

Figure 2.4.2.1. The words were presented to the participants in the above format, in all conditions. In the virtual conditions this image took the form of a floating white panel. In the non-virtual conditions, these words were simply printed on the paper as described in section 2.8.

In conditions three and four (described in 3.6.3. and 3.6.4. below), the white display panels showed the English translations immediately below the Swedish words. This was so that the participants knew the English translations of the Swedish words without needing to refer to material outside the virtual environment.

2.5. EXPERIMENTAL DESIGN

In each of the four conditions, participants were given what was essentially the same task (i.e. to encode, and upon prompting with the English translations, to recall the 32 Swedish words using some variant of the loci method. Their recall scores were then compared using two-factor ANOVAs to test whether or not there was any significant difference in recall scores. Factor 1 in the analysis is the organisation of the information into semantic groups. Factor 2 is the use of VR as a means for presenting the loci.

All participants received a short familiarisation trial to accustom them to the virtual environment. The non-VR participants were asked to remember two words in an imaginary environment. Most participants at this point in the procedure chose their own house or a house that was familiar to them, probably because this was the example given in the explanation sheet (see section 6.3).
The participants were asked to verbally explain what they were doing, in terms of the mental strategy that they employed, and to describe the mental images they were forming, so that the experimenter could be certain that they understood the nature of the task.

For the VR participants, there was a short trial in which they entered a small virtual world consisting of one open room with two words to be learned. As in the familiarisation test for the non-VR conditions, the participants were asked to verbally explain their thoughts and what strategies they employed. Each participant was then given a short verbal test to ensure that these two words had been learned correctly. If they failed to do so, they were asked to repeat the familiarisation procedure.

After the verbal test, the experimenter gave more examples that were designed to demonstrate how it was possible to learn the same keyword by forming a different image, to ensure that the participants understood the nuances of the use of the method.

2.5.1. Dependent variables.

There are three dependent variables in this study: the number of Swedish word bodies recalled; the number of genders of Swedish words recalled; and the number of English words recalled. These were recorded on answer sheets given to the participants after the learning phase (see 2.10. for more detail on the testing phase of this experiment). The answer sheets were scored by the experimenter (see 2.11. for more detail on how the answer sheets were scored).

2.5.2. Independent variables.

The independent variables in this experiment are the environments (for elaboration, see section 2.6.).
2.5.3. Performance measures.

The term 'performance measure' in this thesis refers separately to the recall scores of the genders of the Swedish word, the body of the Swedish word, and the equivalent English word.

2.6. Conditions

The experiment involved the use of four different learning environments. All participants were required to learn the same set of 32 Swedish words, and they were all given the same introduction in use of the loci method. In all four conditions, the words were arranged into eight rooms of four words each, as shown in Tables 1 and 2.

A description of the four conditions follows:
2.6.1. Condition one. No VR, no special organisation.

The first condition served as a control condition, because it used neither a higher degree of organisation nor VR.

In this condition, the words were presented to the participants on a black and white A3-sized map. This version is considered to be the control because this is the form that contained the least experimental manipulation. The purpose of this condition was partly to test the effect of organisation by comparing it to condition two, and partly to test the effect of VR by comparing it to condition three.

![Image 2.6.1. the map of condition one, as presented to the participants.](image-url)
2.6.2. Condition two: No VR, special organisation.

As was the case for condition one, this condition did not make use of VR. However unlike condition one, the words were arranged into semantic groups. The chart used the same layout and design as the virtual environment in condition four, including both the colour coding and spatial grouping of semantically similar words. The chart formed a pictorial representation of how condition four would look if viewed from above. The purpose of this condition was partly to test the effectiveness of organisation by comparing it to condition one, and partly to test the effect of VR by comparing it with condition four.

![Image 2.6.2. The map of condition two, as presented to the participants.](image-url)
2.6.3. **Condition three. VR, no special organisation.**

This condition made use of VR to generate a virtual environment in which it was intended to generate the loci of the loci method. This version required the participants to be responsible for the spatial organisation of the visual images that they were to form. Consequently, there was no colour coding and no locational coding. The words to be learned were displayed on floating white panels that were located in various regular places throughout the virtual world. The purpose of this condition was to test the effectiveness of organisation of information by allowing it to be compared to condition four.

Image 2.6.3: View of the VE presented to the participants in condition three.
2.6.4. Condition four. VR, special organisation.

For convenience 'South', and other compass directions, are herein arbitrarily defined according to
the assumption that the initial direction the participants are facing in the virtual environment is
North.

Condition four organised the virtual environment in a way that made use of the potential of VR.
The southern half of the virtual environment was coloured pink and the northern half blue. Four
of the eight rooms contained words that all had the same gender, and four contained an even mix
of common and neutral gender words, referred to as “en-nouns” and “ett-nouns” respectively. Of
the four rooms that contained words of all the same gender, the two rooms that contained ett-
nouns were both in the blue part of the map, and the two rooms that contained en-nouns were
both in the pink part. In this way, it was intended that the participants could use the colour blue
to code for ett-nouns and the colour pink to code for en-nouns. The participants were not told of
this distinction, and it was of interest to determine whether they might notice it themselves.
Figure 2.6.4: View of the VE presented to the participants in condition four.

The reason for the manipulation with coloured rooms and genders was to test whether or not the combination of both a colour and a locational cue resulted in better recall than simply the colour cue. In the rooms where the colour code matches the gender of the noun, the participants might have been able to recall the gender of the noun by either recalling the colour of the room or the colour of the object, whereas in the other rooms, only the colour of the object itself will be an effective cue.

The user was permitted to navigate in and out of the virtual rooms in conditions three and four as desired. The four nouns in condition four were placed according to semantic groupings, as in condition two.

In this way the potential of VR to capitalise on category learning, increased cue discriminability,
as well as the exploitation of certain neural structures that are used in spatial localisation, were tested.

2.7. Hypotheses

2.7.1. VR and loci method facilitation.

The purpose of this research was to see whether VR could be used as a substitute for participant-generated virtual worlds in the context of foreign language learning. Therefore, it was predicted that the recall scores of the two non-VR conditions would not be significantly different to the recall scores of the two VR conditions. It is possible for the reasons detailed in section 1.2. above that the learners in the VR conditions would out-perform the learners in the non-VR conditions, so this outcome was considered.

2.7.2. Organisation and loci method facilitation.

It is predicted that the higher levels of organisation used in conditions two and four would facilitate use of the loci method, as well as making it easier for the participants to recall the relevant information. Therefore, we expected to see a main effect of organisation of information, with the two specially organised conditions resulting in higher recall than the two randomly organised ones.

2.7.3. Interaction effects

It was predicted that the interaction of the main effects mentioned in sections 2.7.1 and 2.7.2 would result in conditions two and four providing the highest recall. In the event that VR assisted the learners in their task (as opposed to simply not being detrimental) then, it is predicted, given the prediction of the organised conditions providing higher recall, that condition four would provide the highest recall of all conditions, and condition one the lowest.
2.7.4. Other hypotheses, other interesting points

It was predicted that the recall of the English words would be higher than the recall of the body of the Swedish words. Despite evidence that performance in cued recall in paired-associate learning tests does not depend on probe direction (Caplan et al, 2006), it is expected that the previous knowledge of the English words would result in higher recall.

It is possible that the participants would find the navigation of the VE problematic, and that participants with greater previous simulator experience would be able to familiarise themselves with the controls faster than others, thus allowing them to spend more time learning the words and less time navigating the VE. Therefore, there is the possibility that previous simulator experience might result in higher recall scores, a problem mentioned by Frey et al (2007). For this reason, this experimenter recorded the levels of previous simulator experience.

2.8. Materials and Apparatus

In conditions one and two, the words to be learned were presented to the participants on a single A3 sheet of paper, upon which a map consisting of eight rooms with four words in each room were included. The words were presented with their Swedish translations above the English, in vertical columns.

In conditions three and four, the words to be learned were presented on a 17 inch (43.2cm) computer monitor, in the form of a three-dimensional virtual environment that could be navigated by the participant. The virtual environments of conditions three and four were representations of conditions one and two, respectively. As these were intended to simulate use of the loci method, the participants were required to navigate these at "ground level" by means of an invisible avatar that was itself controlled by the mouse. The virtual conditions were constructed by the student using the VRML (Virtual Reality Modelling Language) programming
language. The experimenter himself wrote the scripts that created the virtual worlds. These scripts were then interpreted by the *Mozilla Firefox* web browser, which utilised a plug-in called *Cortona*. This allowed the virtual worlds to be presented in full-screen format. Despite the use of a web browser to view them, the VRML scripts were not run from the Internet, as this was not necessary. They were instead stored on the hard drive of the computer used, which improved the speed and reliability of the procedure. A terminal problem affecting the hard drive is much less likely than a problem affecting internet connectivity.

The virtual environments consisted of a number of textured walls that formed the eight rooms, and a large textured plane that formed the ground. The words to be learned in conditions three and four were presented as a series of 32 white panels that appeared to float above the ground. These panels displayed the Swedish translation of each word to be learned with the English translation immediately below it.

The use of the mouse by the participants to navigate the virtual conditions was simple. If the participant clicked and held the left mouse button and pushed the mouse forwards, their viewpoint in three-dimensional space would appear to move forwards, and if they dragged it back, they would appear to move backwards. The participant could turn the viewpoint of the avatar to the left and right by clicking and holding the left mouse button and pulling the mouse to the left or the right. A second method involved clicking the right mouse button, which then brought up a list of ten different pre-determined locations. Selecting one of the locations with the cursor and releasing the right mouse button would cause the avatar to 'teleport' to that location. These locations were arranged into two basic groups. The first group consisted of two locations, one of which provided a view of the northernmost four rooms, the other providing a view of the southernmost four. The second group consisted of eight locations, one location for each of the eight rooms. These locations each provided a frontal view of its respective room. This method of
interaction, whereby the actions of the user result in a predictable and logical change in the appearance of the virtual environment, is one of the essential features of creating a sense of presence in a virtual environment (Vince 2004, Reid 2005).

The participants were provided with a sheet of blank paper and a pen to use if they wished to make notes. As noted earlier, a consumer-level Nvidia 6200 graphics card provided the rendering of the virtual environments.

2.9. Procedure

The participants in all conditions were seated at a desk, where they were first given an information sheet about the experiment to read through, and then a consent form and a demographics form to complete. Then they were given a fourth sheet describing the use of the loci method and giving several examples (see section 6.4). The participants were asked if they understood the loci method and it was then emphasised that during the course of their participation in the experiment, they would have to use the loci method, as it was described on the sheet, to learn the words they were given, even if they thought another method was more effective. The participants were then given the opportunity to ask further questions.

The procedure that followed was varied depending on whether the participants were in the non-virtual conditions (one or two) or the virtual conditions (three or four). In either case, the participants were given a short trial as an opportunity to use the loci method in a practice setting, and in the cases of conditions three and four, also to practise navigation of the control system inside a virtual world.
2.9.1. Trial levels

In conditions one and two, the short trial consisted of the participants being presented with a sheet of A4 paper on which two Swedish words were written with their English translations underneath each. The participants were then asked to verbally describe to the experimenter how they would remember the two Swedish words by using the loci method. The purpose of this task was to allow the student to ensure that the participant understood the nature of the task before them, and to give them a last chance to ask further questions or to ask for further examples. The experimenter listened to the descriptions and explained two basic techniques for creating a keyword image: 1) either try and find a word in English that sounds similar to the Swedish word, and use that as the basis for the image, or 2) try and break the Swedish word up into component parts, and then think of English words that sound similar to each of these parts.

In conditions three and four, the short trial involved the participants being presented with a virtual world that acted as a trial room. This world consisted of a flat, green, textured 'ground' and two short, red brick textured walls.
Figure 2.9.1. View of the VE presented to the participants in the trial level.

In this world there were two floating white panels, exactly the same as the ones used in the learning conditions of conditions three and four, i.e. rectangular, with a word in Swedish displayed on the top half and the English translation directly underneath. The participants were asked to navigate the world in order to learn the control system, and when the experimenter considered them proficient at this task, they were asked to navigate towards the first of the white panels. As in the non-virtual trial, they were then asked to verbally describe to the experimenter how they would remember the two Swedish words by using the loci method.

The trial level produced two additional benefits in the virtual conditions. Not only did it give the experimenter an opportunity to ensure that the participant understood the nature of the task they
were given, it allowed them to familiarise themselves with navigating in a VE, which should have helped to even out some of the pre-existing differences in simulator experience that may have existed among the participants (see section 2.7.4.). The trial levels also gave the experimenter a chance to ensure that the participants did not suffer from cybersickness (see Frey et al, 2007).

2.9.2. Experimental levels

After the participants had completed the trial levels to a degree where the experimenter was confident that they understood the task, they began the actual learning phase.

In conditions one and two, the participants were presented with the A3 sheet of paper described in section 2.8. They were then familiarised with the layout of the rooms on the map, in particular they were told that the map represents eight rooms of four words each, with the Swedish translation above and the English translation below, and that the words had one of two different genders (either 'en' or 'ett'). The participants in condition two were not told that the words in their condition were organised into semantic groups. They were then asked if they had any further questions. If not, the experimenter advised them that they could ask questions at any point during the learning phase, that they had 20 minutes to learn as many words as they could, and that their 20 minutes had started. This time was measured on a stopwatch held by the experimenter. The experimenter remained in the room in order to be able to answer any questions that might arise.

In conditions three and four, the experimenter ran the VRML file that displayed the virtual world on the desktop monitor. The participants were shown how to use the right mouse button to 'jump' to certain locations, these locations being either directly in front of one of the eight rooms, or further back so that four rooms were visible at a time. This feature was not available during the trial levels, as there was nowhere to jump to. This feature was put in place as an example of the
potential benefits of using VR. The participants were then shown a bird's-eye view of the virtual world, which looked very similar to the view that the participants in the non-VR conditions had been given on their A3 paper maps. As in the non-VR conditions, the participants in condition four were not told that the words in this task were organised into semantic groups. As in the non-VR conditions, they were asked if they had any further questions, were told of the 20-minute time limit and were asked to begin.

2.10. Testing.

In conditions one and two, the maps with the words to be learned on them were taken away from the presence of the participant. In conditions three and four, the test computer was switched off. Each participant was given a test sheet that listed, in the same order as on the learning sheet, each of the English translations of the Swedish words they had been given. There was a blank space in place of the Swedish words, in which the participants were asked to write the Swedish translation of the English word, complete with its gender. They were told that they had ten minutes to recall as many words as they could, but that if they finish before 10 minutes or could not recall any more words then they could tell the experimenter, who would then give them the next test sheet. When the participant felt they were finished, or when the ten minutes had expired, the first test sheet was taken away. They were then given a second test sheet that listed, in the same order as on the learning sheet, each of the Swedish translations of the English words they were given. Again, they were told that they had ten minutes to recall as many words as they could. When they were finished with this sheet or if their ten minutes had expired, they were thanked for their participation and were given their participant incentives.

2.11. Scoring

Some participants did not provide an answer for the main body of the some words, but did give a gender by itself. Often the participant gave a gender for every one of the 32 words, even if they
gave a main body for far fewer. As the goal of this experiment was to compare different learning conditions and not different strategies for guessing the gender of vocabulary items, no score was awarded for a gender recalled in isolation from any attempt at recalling the main body of the word.

In the English recall section of the experiment, a point was awarded for a correct answer and nothing was awarded for an incorrect answer.

The spelling of the words was only considered important for the main body of the Swedish word. If the English word was misspelled but if it was nevertheless clearly the correct word, a full point was awarded. This occurred on five separate occasions.

The scoring system for the body of the Swedish words was based on one devised by Pressley & Levin (1981). Each answer given for the Swedish body of the words was categorised to be either correct; correct except for one letter added, deleted, or replaced by another letter; correct except for two letters added, deleted or replaced by other letters; or incorrect.
CHAPTER 3.

RESULTS

3.1. Overview of results and miscellaneous notes

A literature review by van Hell & Mahn (1997) suggested that learners with previous experience in learning foreign language vocabulary might benefit less from keyword method instruction than is the case for inexperienced foreign language learners.

Ellis & Beaton (1993) raised the possibility that learners that already knew several languages should be able to learn vocabulary in a new language faster. "New skills or knowledge invariably initially build upon whatever relevant abilities or knowledge are already present; then, as they are used, they legitimate and make more relevant those prior skills and knowledge, and so in turn cause their further development."

Due to the high numbers of the body of the Swedish words that were recalled with one error, in the present study it was decided to group the perfectly recalled words together with the words recalled with one error into a single, unweighted variable. A similar approach was adopted by Pressley & Levin (1981). It should be noted that in this section the phrase "recall of the body of the Swedish word" should be read as "recall of the body of the Swedish word either correctly or with one error".

Note also that no data from the training phase was collected; therefore, no analysis will be made regarding this aspect of the experiment.
3.2. Analysis of Variance.

3.2.1. Body of Swedish words

3.2.1.1. Descriptive statistics

The following statistics refer to the number of Swedish words that the participants could either recall correctly or with one error.

**Table 3.2.1.1. Mean recall scores and standard deviations of participants when recalling the main body of the Swedish word.**

<table>
<thead>
<tr>
<th>Condition number</th>
<th>Mean recall score (maximum 32)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition one</td>
<td>14.20</td>
<td>6.39</td>
</tr>
<tr>
<td>Condition two</td>
<td>18.73</td>
<td>6.30</td>
</tr>
<tr>
<td>Condition three</td>
<td>16.87</td>
<td>7.23</td>
</tr>
<tr>
<td>Condition four</td>
<td>19.13</td>
<td>6.37</td>
</tr>
<tr>
<td>All conditions</td>
<td>17.23</td>
<td>6.71</td>
</tr>
</tbody>
</table>
Figure 3.2.1.1.1. Bar graph of recall scores for the main body of the Swedish words, all conditions.

The distribution of recall scores for the main body of the Swedish words clearly followed a normal distribution curve.
Figure 3.2.1.2. Bar graph of recall scores for the body of the Swedish words, for each condition.

Inspection of the distribution curves for the 15 participants in each of the four conditions suggests that even with the small samples sizes in question, the curves approximate, to varying degrees, a bell curve. This is especially clear in condition three (bottom left of Figure 3.2.1.2.). Also evident from this graph is that none of the participants in the two organised conditions (both on the right) scored lower than four, which was the case for neither of the unorganised conditions.
3.2.1.2. ANOVA of recall of Swedish words by participants.

<table>
<thead>
<tr>
<th>Manipulation</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtuality</td>
<td>1</td>
<td>35.27</td>
<td>0.8138</td>
<td>0.370863</td>
</tr>
<tr>
<td>Organisation</td>
<td>1</td>
<td>173.40</td>
<td>4.0013</td>
<td>0.050323</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>19.27</td>
<td>0.4446</td>
<td>0.507653</td>
</tr>
</tbody>
</table>

3.2.1.3. Virtuality.

The ANOVA showed that whether or not the words were presented in VR format was not significant \[ F (1, 56) = 0.8138, p = 0.370863 \].

![Main effect of virtuality on recall of Swedish words](image)

**Figure 3.2.1.3.** Graph showing the effect of the manipulation of virtuality on recall of the main body of the Swedish language words.
3.2.1.4. Organisation.

The ANOVA showed that whether or not the words were organised when they were presented was marginally significant [F (1, 56) = 4.0013, p = 0.050323].

![Main effect of organisation on recall of Swedish words](image)

**Figure 3.2.1.4.** Graph showing the effect of the manipulation of organisation on recall of the main body of the Swedish language words.

3.2.1.5. Interaction.

The ANOVA showed that the interaction of the factors of virtuality and organisation was not significant [F (1, 56) = 0.4446, p = 0.507653].
Figure 3.2.1.5. Graph showing the effect of the interaction of the manipulations of virtuality and organisation on recall of the main body of the Swedish language words.

3.2.1.6. Further analysis

Although there was no significant difference in recall scores between the virtual and non-virtual conditions taken together, a repeated-measures t-test showed that the difference in recall scores between conditions one and two was marginally significant ($p = 0.060$, df = 28). Of course, this means that the difference in recall scores between conditions three and four were not significant ($p = 0.37$, df = 28).

3.2.2. Gender of Swedish words

3.2.2.1. Descriptive statistics

The following statistics refer to the number of the gender of the Swedish words that the participants could correctly recall:
Table 3.2.2.1. Mean recall scores and standard deviations of participants when recalling the gender of the Swedish word.

<table>
<thead>
<tr>
<th>Condition number</th>
<th>Mean recall score (maximum 32)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition one</td>
<td>14.33</td>
<td>7.23</td>
</tr>
<tr>
<td>Condition two</td>
<td>18.40</td>
<td>6.31</td>
</tr>
<tr>
<td>Condition three</td>
<td>15.67</td>
<td>10.45</td>
</tr>
<tr>
<td>Condition four</td>
<td>16.40</td>
<td>7.94</td>
</tr>
<tr>
<td>All conditions</td>
<td>16.20</td>
<td>8.07</td>
</tr>
</tbody>
</table>

The distribution of recall also followed a clear bell curve, although a little flatter than the curve of recall scores from the main body of the Swedish words. It should be noted that at the far left end of the bell curve there are is an anomaly regarding the number of participants who recalled zero of the genders correctly. This is possibly because certain participants found the concept of gender so unusual and difficult that they made a strategic decision to ignore it from the beginning. None of the participants recalled either one or two of the genders.

It should also be noted that the standard deviation of the results for the gender of the Swedish words (8.07) is higher than that for the body of the Swedish words (6.71) or for recall of the English words (6.36).
Figure 3.2.2.1.1. Bar graph of recall scores for the gender of the Swedish words, all conditions.

The distribution of recall scores for the main body of the Swedish words followed a clear normal distribution curve.
Figure 3.2.2.1.2. Bar graph of recall scores for the gender of the Swedish words, for each condition.

A tendency towards normal distribution curves is less evident in this analysis. This may be because some of the participants were confused by the concept of linguistic gender and ignored it, whereas some others may have seen the patterns in the genders of the words that were given and learned them well because of this. Also possible is that some participants were unable to comprehend the concept of linguistic gender, so those that did often got both genders correct, whereas those that didn’t, got neither.
3.2.2.2. ANOVA of recall of gender of Swedish words.

Table 3.2.2.2. ANOVA of recall of gender of Swedish words.

<table>
<thead>
<tr>
<th>Manipulation</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtuality</td>
<td>1</td>
<td>1.67</td>
<td>0.0252</td>
<td>0.874511</td>
</tr>
<tr>
<td>Organisation</td>
<td>1</td>
<td>86.40</td>
<td>1.3049</td>
<td>0.258181</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>41.67</td>
<td>0.6293</td>
<td>0.430963</td>
</tr>
</tbody>
</table>

3.2.2.3. Virtuality.

The ANOVA showed that whether or not the words were presented in VR format was not significant \[F(1, 56) = 0.0252, p = 0.874511\].
Main effect of virtuality on recall of gender of Swedish words
Current effect: F(1, 56) = .025, p = .87451
Effective hypothesis decomposition
Vertical bars denote 0.95 confidence intervals

**Figure 3.2.2.3.** Graph showing the effect of the manipulation of virtuality on recall of the gender of the Swedish language words.

**3.2.2.4. Organisation**

The ANOVA showed that whether or not the words were organised when they were presented was not significant [F (1, 56) = 1.3049, p = 0.258181].
Main effect of organisation on recall of gender of Swedish words

Current effect: $F(1, 56)=1.3049$, $p=0.25818$

Effective hypothesis decomposition

Vertical bars denote 0.95 confidence intervals

---

**Figure 3.2.2.4.** Graph showing the effect of the manipulation of organisation on recall of the gender of the Swedish language words.

3.2.2.5. Interaction.

The ANOVA showed that the interaction of the factors of virtuality and organisation was not significant [$F (1, 56) = 0.6293$, $p = 0.430963$].
Figure 3.2.2.5. Graph showing the effect of the interaction of the manipulations of virtuality and organisation on recall of the gender of the Swedish language words.

3.2.2.6. Further Analysis.

There was again no difference in recall scores between the virtual and the non-virtual conditions, but two separate t-tests showed that the difference in recall scores between conditions one and two \((p = 0.11, \text{df} = 28)\) was again much greater than the difference in recall scores between conditions three and four \((p = 0.83, \text{df} = 28)\).

The difference in recall scores between conditions one and three \((p = 0.69, \text{df} = 28)\) was not greater than the difference in recall scores between conditions two and four \((p = 0.45, \text{df} = 28)\).

3.2.3. English words

3.2.3.1. Descriptive statistics

The following statistics refer to the number of English words that the participants could correctly
Table 3.2.3.1. Mean recall scores and standard deviations of participants when recalling the English word.

<table>
<thead>
<tr>
<th>Condition number</th>
<th>Mean recall score (maximum 32)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition one</td>
<td>23.20</td>
<td>7.07</td>
</tr>
<tr>
<td>Condition two</td>
<td>25.73</td>
<td>4.61</td>
</tr>
<tr>
<td>Condition three</td>
<td>25.27</td>
<td>7.34</td>
</tr>
<tr>
<td>Condition four</td>
<td>26.60</td>
<td>6.21</td>
</tr>
<tr>
<td>All conditions</td>
<td>25.20</td>
<td>6.36</td>
</tr>
</tbody>
</table>
Figure 3.2.3.1.1. Bar graph of recall scores for the English words, all conditions.
Figure 3.2.3.1.2. Bar graph of recall scores for the gender of the English words, for each condition.

The distribution of recall of English words exhibits a clear ceiling effect on the right hand side of the main histogram, and this is especially evident in the histogram for condition four. This is probably because recall of the English words from the Swedish translations is easier than vice-versa; this was expected to be the case. Howard (1995) wrote that recognition is much easier than recall. Although the situation in this experiment is not as simple as recognition versus recall, the participants asked to recall the English words only had to recall a word that they were familiar with based on recognition of the Swedish word that was associated with it, rather than recall a pattern of unfamiliar Swedish letters based on the English word associated with them.

Almost a quarter of the participants (14/60) managed to recall 30 or more of the 32 English
words. Many of the participants remarked to the experimenter that they found recall of the English words far easier. It should be, however, emphasised that in all cases the participants were asked to recall the English words from the Swedish translations immediately after they were asked to recall the Swedish words from the English translations, so there is the strong possibility that there was an effect attributable the extra rehearsal time, creating a bias towards higher recall of the English words. In any case, English recall scores were not compared against recall scores for either the body of the Swedish words or the gender of the Swedish words.

3.2.3.2. ANOVA of recall of English words.

Table 3.2.3.2. ANOVA of recall of English words.

<table>
<thead>
<tr>
<th>Manipulation</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtuality</td>
<td>1</td>
<td>32.27</td>
<td>0.7884</td>
<td>0.378379</td>
</tr>
<tr>
<td>Organisation</td>
<td>1</td>
<td>56.07</td>
<td>1.3699</td>
<td>0.246778</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>5.40</td>
<td>0.1319</td>
<td>0.717792</td>
</tr>
</tbody>
</table>

3.2.3.3. Virtuality.

The ANOVA showed that whether or not the words were presented in VR format was not significant [F (1, 56) = 0.7884, p = 0.378379].
Figure 3.2.3.3. Graph showing the effect of the manipulation of virtuality on recall of the English words.

3.2.3.4. Organisation

The ANOVA showed that whether or not the words were organised when they were presented was not significant [F (1, 56) = 1.3049, p = 0.258181].
**Figure 3.2.3.4.** Graph showing the effect of the manipulation of organisation on recall of the main body of the Swedish language words.

3.2.3.5. *Interaction.*

The ANOVA showed that the interaction of the factors of virtuality and organisation was not significant \[ F (1, 56) = 0.1319, p = 0.717792 \].
**Figure 3.2.3.5.** Graph showing the effect of the interaction of the manipulations of virtuality and organisation on recall of the English words.

### 3.2.3.6. Further Analysis.

Once more, although there was no significant effect of virtuality on the recall of English words, the difference between conditions one and two \((p = 0.26, \text{ df } = 28)\) was much greater than the difference between conditions three and four \((p = 0.6, \text{ df } = 28)\).

The difference in recall scores between conditions one and three \((p = 0.44, \text{ df } = 28)\) was greater than the difference in recall scores between conditions two and four \((p = 0.67, \text{ df } = 28)\).

### 3.2.4. Further observations regarding mean recall scores

In all three of these groups (Swedish recall, gender and English recall), the participants in condition one returned the lowest average recall score. In the recall of the body of the Swedish
word and in recall of the English word, the participants in condition four returned the highest score. In all three groups, the mean recall score for condition four was higher than either score for condition one or two.

There was a correlation (product-moment) of 0.26 between gender and total number of languages able to be read by the individual, at any skill level. This correlation was significant at the p < 0.05 level.

There was a significant difference between the mean recall score of the English words and the mean recall score of the body of the Swedish words. This was analysed with a single means t-test. The English recall scores had a mean of 25.2 and a standard deviation of 6.36; the recall of the body of the Swedish words had a mean of 17.23 and a standard deviation of 6.71. In both cases n was 60. This difference was significant at the p < 0.01 level.

3.3. Mean recall scores and their correlations with other variables.

3.3.1. Previous language ability

There was a significant correlation between previous language ability (measured by total number of foreign languages able to read at any skill level) on recall of the body of the Swedish word either correctly or with one error (p < 0.05, df = 4). The correlation here was 0.37. The correlation between previous language ability and recall of the gender of the Swedish word was 0.28, which was also significant at the p < 0.05 level. The correlation between previous language ability and correct recall of the English word was not significant.

This suggests that perhaps previous language ability enabled the participants that possessed it to learn the Swedish language vocabulary items faster. It is also possible that an intervening or confounding variable, in this case general interest in foreign languages and/or a corresponding
desire to learn them, explains the correlation. It is reasonable to expect that a person who is interested in foreign language learning will have learned a foreign language and will be more motivated to learn the foreign words used in this experiment. This increased motivation and/or interest in foreign language learning may in itself improve recall (Higbee, 1977).

The observation that there was a significant correlation for recall of the gender of the Swedish word and previous language experience is not surprising, as participants who are familiar only with English would not have encountered the use of linguistic gender. There are a few exceptions in English, such as the use of the pronoun she for describing countries and ships.

3.3.2. Previous simulator experience

In an effort to control for prior simulator experience on the recall results, the participants were asked how much previous simulator experience they had (See Appendix 6.2).

![Prior simulator experience amongst participants](image)

**Figure 3.3.2: Bar graph showing differing levels of previous simulator experience amongst participant population, pooled across all participants.**
Note that these measures may not accurately reflect levels of simulator experience in the population as a whole. Because the mean age of the participants was 24.03, with a standard deviation of 7.09, the levels of simulator experience reported must be viewed alongside the fact that video gaming and interest in computing is something that is more popular and common amongst the young.

3.3.3. Gender of participants

The 20 males who participated in this study had an average recall score for the body of the Swedish words either correctly or with one error of 17, whereas the 40 females had an average recall score of 17.35. This difference in recall scores was not significant at the $p < 0.05$ level ($p = 0.850782$, df = 58).

The 20 males had an average recall score for the gender of the Swedish words of 16.7, whereas the 40 females had an average recall score of 15.95. This difference in recall scores did not reach significance at the $p < 0.05$ level ($p = 0.737328$, df = 58).

The 20 males had an average recall score for the English word of 25.35, whereas the 40 females averaged 25.125. This difference in recall scores was not significant ($p = 0.898495$, df = 58).

The females in this study were able to read, at any skill level, an average of 1.575 other languages, whereas the males averaged one other language exactly. This difference was significant at the $p < 0.05$ level, ($p = 0.043091$, df = 58). Although there was a significant correlation between previous language ability and recall of the body of the Swedish word, females with not score significantly higher than males on this measure.
3.3.4. Age of participants

There was no significant correlation between the age of the participants and recall on any of the three performance measures. Recall of age and the gender of the Swedish words had a correlation of 0.04; recall of age and the body of the Swedish words had a correlation of -0.02; and recall of age and the English words had a correlation of 0.02. All of these scores were not significant at the $p < 0.05$ level.

The correlation between age and previous language ability, at -0.04, was also not significant.

3.4. Correlations between the performance measures themselves.

To the extent that this experiment is a test of the individual language learning capacity of the participants, it would not be surprising to find significant correlations between the performance measures.

3.4.1. Correlation of body of Swedish word and gender of word

It can be safely assumed that the incidence of some participants learning the body of the Swedish words but not the gender, and vice-versa, was low. The Pearson correlation between correct recall of the gender of the Swedish words and recall of the body either correctly or with one error was 0.68, which was significant at the $p < 0.05$ level.

3.4.2. Correlation of English and Swedish recall

The correlation between correct recall of the English translation and recall of the body of the Swedish word was 0.79 (significant at $p < 0.05$). Despite the fact that the means of the recall scores themselves were significantly different, the correlation between these scores was still very strong. It can therefore be assumed that the participants, as a whole, generally did not focus on learning to recall the words in one direction to the exclusion of the other, i.e. they did not learn to
recall English words when given the Swedish prompts significantly better or worse than recalling Swedish words when given the English prompts. This is consistent with the hypothesis in 2.7.4.

This finding is not surprising, as it simply suggests that language ability in learning foreign language vocabulary provides benefits to both active and passive recall. The magnitude of this correlation suggests that there were very few cases where a participant would score highly in one form of recall and poorly in another.

3.5. Demographic Statistics

Table 3.5: Demographic statistics, all participants.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Age</td>
<td>24.03</td>
<td>7.09</td>
</tr>
<tr>
<td>Number of males</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Number of females</td>
<td>40</td>
<td>-</td>
</tr>
<tr>
<td>Average age of males</td>
<td>25.05</td>
<td>6.48</td>
</tr>
<tr>
<td>Average age of females</td>
<td>23.53</td>
<td>7.40</td>
</tr>
<tr>
<td>Average Beginner Languages</td>
<td>0.77</td>
<td>0.72</td>
</tr>
<tr>
<td>Average Intermediate Languages</td>
<td>0.35</td>
<td>0.55</td>
</tr>
<tr>
<td>Average Advanced Languages</td>
<td>0.27</td>
<td>0.52</td>
</tr>
<tr>
<td>Average Languages, any level</td>
<td>1.38</td>
<td>1.04</td>
</tr>
</tbody>
</table>
Females outnumbered males by 2:1 in this study, the major reason being that the bulk of the participants were recruited by means of advertisements to undergraduate level psychology students, a group which comprises vastly more females than males. A majority of the participants were familiar with at least one language other than English. The educational level of the participants was not measured further than asking for their previous level of language ability.

The variability of previous language ability was great. In the case of intermediate and advanced language ability, the standard deviation was larger than the mean.

3.6. Miscellaneous points.

None of the participants reported any kind of cybersickness, and all of them completed their participation to the end. This 100% completion rate is interesting in the context of the results found by Frey et al (2007), who found that 11% of the participants in their study suffered from cybersickness, most of whom had to abort the experiment. It is possible that some elements of the VE used in this experiment made cybersickness less likely, one possibility could be the gentle pastel colours employed in the textures of the VE. However, it is much more likely that the use of a non-immersive desktop monitor instead of a HMD was responsible for the low levels of cybersickness.
CHAPTER 4.

DISCUSSION

The results of this study raise a number of possible avenues for future research. In particular, a number of separate follow-up studies suggest themselves.

4.1. Organisation

The significant improvement in recall of the body of the Swedish words by the participants in the organised conditions compared to the non-organised ones suggested that there is potential for synergy between the use of the loci method and extra-mnemonic strategies. A self-directed learner of foreign vocabulary might initially be slowed by using these techniques, as they would have to organise the words into groups. However, given that organisation seems to allow for better recall, this initial impediment might be outweighed by the advantages in overall efficiency, especially over the course of learning several thousand words.

It is perhaps worth noting that neither the participants in conditions two nor four were told that the words in their conditions were organised into semantic groups. It is possible that with this knowledge, the participants may have been able to form more effective series of keyword images and that their mean recall scores would have been even higher than the non-organised conditions. However, it is also possible that many of the participants in these conditions realised without being told that the words were organised in this manner. It is unclear whether or not the participants' awareness of the organisation of the vocabulary items did or would affect their learning ability.
It would be interesting to investigate the effects of further levels of organisation. For example, a study with eight groups of words could be arranged into two "meta-groups" of four groups each.

4.2. Virtual Reality

Despite the VR conditions providing average higher recall scores than the non-VR ones, this difference was not shown to be significant. When reviewing this finding one must consider the low power of this study. A significant difference may have been found if the power of the study was greater, which could have been achieved by having two VR groups instead of four in total, that is, by either dropping the organised factor from the study, or by increasing the number of participants in the study.

It would also be useful to investigate whether improvements to the nature of the VR used in this experiment might assist the learner. The participants did not receive extensive familiarisation before beginning their participation in the study. It is possible that a greater familiarity with paper-based learning amongst the participants as a whole may have artificially increased the scores of conditions one and two. One could usefully run a higher-power study that compared the equivalents of conditions three and four only. Moreno & Mayer (2002) mentioned that environments with higher levels of immersion provide a learning advantage because the learner needs to allocate fewer attentional resources to gaining familiarity with the interface, and can therefore concentrate better on learning the material. A future study that uses a higher-immersion VR condition might be of value here, given the (non-significant) trend towards higher recall scores in the virtual conditions.

Furthermore, the images used in the VR conditions were not particularly detailed and could have been made more so with a little effort. The words were presented to the participants by being displayed on a floating white panel; a series of three-dimensional objects that elicited more
imagination might have worked better. This would be easily achievable within the framework of a VR system.

One factor that may have limited the recall scores of the participants in the virtual conditions is raised by the results of a study by Godden and Baddeley (1975). They found that the recall scores of some groups of participants were strongly influenced by whether or not they recalled the information in the same environment that they were in when they had learned it. Howard (1995) proposed that this might be explained by the environment itself acting as a set of recall cues for the information desired. If this is correct, then perhaps the participants in the VR conditions were at a relative disadvantage when they had to recall their words on paper as opposed to on the screen or by using a computer. Relevant to this point is the fact that the answer sheets for all conditions in this experiment were very similar in appearance to the learning sheets in the non-virtual conditions. The learners in the non-virtual conditions may have therefore gained some positive transfer by being able to recall the position of certain words on the answer sheet. The learners in the virtual conditions would not have been able to do this.

4.3. Limitations of the study/ further research.

Most of the following points relate to shortcomings in the scale of this experiment to the task of actually learning a foreign language.

4.3.1. The type of words being learned.

This study looked only at nouns, and these were mostly concrete nouns. Further research is required to investigate the effectiveness of the system in this thesis on learning abstract nouns, verbs and adjectives. Simple animations could be used in VR to help the participants learn the words required. A learner of a foreign language must acquire a vocabulary of thousands of words, and so a system that can only help to learn nouns is incomplete.
4.3.2. Higher-quality VR.

There is considerable research investigating interference and discriminability. It was thought that the use of VR might be able to enhance the participant's use of the loci method, as the computer-generated images may provide a large number of discriminable cues. However, the differences in recall scores between the VR and non-VR conditions proved not to be significant. With a higher level of technology (in particular greater graphics display capabilities), a virtual environment could be created with a much richer environment that itself had loci that were easier to discriminate from each other. The computer-based nature of the experiment would allow sounds to be used and associated with the loci, which might further increase their discriminability.

4.3.3. Various types of mnemonic learning.

Several researchers have found evidence that suggests that the long-term retention of information learned using the keyword method was no greater than that learned using rote learning (e.g. Olton, 1969; Thomas & Wang, 1996). However, this information related to the use of keyword learning only, not to the greater structuring and organisation of the information using a method such as the loci method. It would be useful to study the long-term retention of information learned using the loci method compared to short-term retention of information using the same method, as well as long-term retention of information learned using other mnemonics.

In the context of foreign language learning, the value of prior organisation of the words could be usefully investigated further, especially when the benefits afforded by computer-aided learning are considered. Perhaps an experiment could be conducted where the participants are able to firstly organise the words into whichever format they consider the easiest before the actual learning task begins.
4.3.4. Duration of experiment.

As mentioned in the Introduction and in section 1.1., the purpose of this experiment was to investigate the effectiveness of VR on foreign language learning. There is further need to investigate the effects of VR on learning a far greater body of words. It is possible that the effects of VR would be maximised when it comes to longer-term studies, because of factors such as motivation and getting used to the system and interface. The experiment reported by Harless et al. (2005) lasted for several hours spread over several days. Given the usual amount of time that a learner must expend in order to learn a foreign language, this thesis is for practical reasons rather smaller in scale than would be ideal for providing definitive answers.
CHAPTER 5.

CONCLUSIONS

This study was not simply an investigation into human interface technology. It was also a study into the ability of learners to apply sophisticated memory techniques to the task of learning foreign language vocabulary in the context of human interface technology.

5.1 Hypothesis 2.7.1

The hypothesis given in section 2.7.1. predicted that the recall scores of the VR conditions would not be significantly different to the recall scores of the non-VR conditions, although it was possible that the recall scores for the VR conditions may be higher. The results showed that for all three dependent variables the recall scores were not significantly different. However, in almost every case (five out of six), the recall scores of the VR condition was higher than its corresponding non-VR condition, although not significantly higher. The one exception was for the recall of the gender of the Swedish word, where the recall scores for condition two (organised, non-virtual) were higher than for condition four (organised, virtual). With greater power, for example, with a larger number of participants, the results are likely to have reached statistical significance.

5.2 Hypothesis 2.7.2

The hypothesis given in section 2.7.2. predicted that there would be a main effect of the organisation of the words given, namely the in the conditions where the words were organised into semantic groups, the recall scores would be higher. Although this was true in every one of the six cases where an organised condition was compared with a non-organised counterpart, the
results showed that this difference was only marginally significant in the case of the recall of the
body of the Swedish words, and elsewhere the difference was non-significant.

5.3 Hypothesis 2.7.3
The hypothesis given in section 2.7.3 considered the interaction of the main effects of virtuality
and organisation, and predicted that condition one would result in the lowest recall scores for
each of the three dependent variables and condition four the highest. Condition one did result in
the lowest recall scores for each of the three dependent variables. Condition four resulted in the
highest recall scores for two of the dependent variables, and the second highest in the case of
recall of the gender of the Swedish word. The ANOVAs showed that there was no significant
effect of the interaction between virtuality and organisation in any of the three dependent
variables.

5.4. Minor Hypotheses, in section 2.7.4
The minor hypotheses in section 2.7.4 considered recall of the English words versus recall of the
Swedish ones, and the effect of previous simulator experience on recall scores. The recall of the
English words was predicted to be higher, and the results showed that they were.

Overall, there is a suggestion in this research that the creation of a suitable virtual environment
for enhancing the learning of Swedish vocabulary by the method of loci has the potential to yield
improved learning outcomes. The effect of providing a virtual environment however, does not
appear, from the evidence available from this study, to be stronger than other non-virtual
enhancements to the method such as the prior grouping of the words into related semantic
categories.

It was considered that the data regarding prior simulator experience was not useful because of
the small sample sizes of participants recording higher levels of same.
6. APPENDICES

6.1. Participant consent form

Participant consent

1. I have read the participant information sheet provided and understand the description of the
   above-named experiment.

2. On this basis, I agree to participate as a subject in the experiment, and I consent to publication of
   the results, with the understanding that anonymity will be preserved.

3. I understand that I may withdraw my participation at any time, including the withdrawal of any
   information I have provided.

Name:

Signed:

Date:
6.2. Demographics form

General demographics

Please circle the appropriate answer or fill in the spaces provided:

1. Sex: M / F
2. Age: ________
3. Handedness: Left / Right
4. Simulator experience:

"Simulator" includes full-motion simulators, head-slaved virtual reality equipment, and PC based 3D first-person perspective computer games, such as Tomb Raider, Unreal Tournament, Halo, Doom etc. as well as flight simulators.

None.
Less than 10 hours.
10 - 100 hours
100 - 1000 hours
More than 1000 hours
5. Foreign language experience:

Please write the languages you can read in the spaces below, as well as your level of proficiency. Rate your proficiency as either 'Beginner', 'Intermediate' or 'Advanced'.

<table>
<thead>
<tr>
<th>Language</th>
<th>Skill Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.3. Participant Information Sheet

Participant information

The experiment you are about to participate in is investigating the usefulness of virtual reality for people trying to learn a number of words in a foreign language.

You will first be trained in the use and application of the Loci Method, a mnemonic technique utilising imaginary three-dimensional space for the purpose of recalling information. You will be asked to demonstrate understanding of the concept of the Loci Method before beginning the main component of this experiment.

In the main component, you will be asked to learn a total of 32 Swedish-language words, and your recall ability will be tested at the end of your participation. The total session length is expected to be around 40 minutes, and you will have 20 minutes to use the Loci Method in order to learn the required words. At the end of the session you will be asked for your general impressions and feedback on the experiment.

All information and data collected in this study are kept private and confidential to the experimenter and the experimenter’s supervisor. Individual participants will not be identifiable in results or publications. Your participation in this study is optional and you may withdraw your participation at any time, including the withdrawal of any information you have provided. However, by signing the consent form attached, it is understood that you have consented to participate in this experiment and to publication of the results, with the understanding that anonymity will be preserved.
If you have questions related to the study itself or the results obtained from your participation, please feel free to contact Vince McLeod, Department of Psychology, University of Canterbury, in Room 605, by phone at 364-2987 ext 7173, or by email at vjm30@student.canterbury.ac.nz.

This research has been reviewed and approved by the University of Canterbury Human Ethics Committee.

Please take this sheet with you when you leave.
6.4. Introductory guide to the Loci Method

The Loci Method - a brief user's guide.

During your participation in this experiment you will be asked to use a mnemonic technique called the Loci Method. It is possible that you have encountered some form of this technique before, as it is reasonably well-known.

The idea of the Loci Method is to form an association between two visual images, one generally being a three-dimensional visual image and the other being any other kind of memorable image. The three-dimensional visual image should represent a location that you are reasonably familiar with.

First imagine the house where you live. See the front door as you walk in. Then imagine yourself walking through every possible location in your house. Then choose a number of specific places in your house. This number should equal the number of pieces of information you want to remember. For example, you may choose the following:

1. Front Door
2. Kitchen
3. Hallway
4. Toilet
5. Bedroom

Now comes the interesting part. The trick with the Loci Method is to form a series of
imaginative visual images so that *when you imagine yourself walking through a certain location, the appearance of that location acts as a cue for the image.* For example, if you wanted to remember a shopping list that consisted of: tomatoes, bread, milk, grapes and chocolate, you might imagine the following:

You walk up to your front door as someone throws a giant tomato at you from behind. You see it splatter into the door and dribble towards the front step in watery red streaks. As you go through the door into the kitchen, you notice that someone has replaced your refrigerator with an enormous loaf of bread, which rests against one wall. You go and prod it and feel that it is much softer than the fridge was. You leave and go into the hallway, and a splashing noise on the floor alerts you to the presence of a stream of milk that runs out of a giant milk bottle embedded in the far wall. A glance to your left reveals a toilet bowl full of grapes, which overflow onto the floor. Bemused by the strange appearance of your house, you go into your bedroom for a lie down, only to find that your bed has been replaced with an enormous block of chocolate.

You should find that recalling these items is easy if you imagine yourself retracing your route through your house.

Learning vocabulary items in a foreign language is slightly different but not any harder. For example, if you wanted to remember the Swedish word 'Bord' (meaning 'Table') you might imagine a man sitting at a table with his head in his hands and a bored expression on his face. When you see this image it ought to cue the word 'bored' which ought to in turn cue the Swedish 'Bord'. You are asked to use this technique during the course of your participation in the experiment. Please ask the experimenter if you would like more information or examples.
7. REFERENCES


Littlewood.


Herrmann, D., Rea, A., & Andrzejewski, S. (1988). The need for a new approach to memory training. In M.M. Gruneberg, P.E. Morris & R.N. Sykes (Eds.), *Practical Aspects of*


Concrete and Abstract Foreign Words by Experienced and Inexperienced Learners.

*Language Learning, 47:3, 507-546.*


Directions. Information Age Publishing, Greenwich, CO., USA.