UNIVERSITY OF TURKU

An Analysis of a Finnish 5th Grade German Class's Use of an Open-Source CALL tool for Vocabulary Learning:

A Case Study

Language Specialist Degree Programme, Digital Language Studies
Master's thesis
School of Languages and Translation Studies
Faculty of Humanities
Markus, Suominen

03/28/2023
Turku

Master's thesis
Language Specialist Degree Programme, Digital Language Studies
Markus Suominen
An Analysis of a Finnish 5th Grade German Class Use of an Open-Source CALL tool for Vocabulary Learning
63 pages, 1 appendix
03.28.2023.

The field of computer assisted language learning (CALL) studies the usage of computers in language learning. The field got its start, many decades ago, through the creation of a few projects, such as PLATO, in which the computer was used as an educational system/program. Many of these tools are available online through a web browser. However, using a computer tool via the internet is only a method among many. For example, software can be created for mobile phones and computers in many different forms and for different use purposes. One of these purposes is language learning. Thanks to the advancement of different types of computer technology people have access to a wide variety of different advanced language learning tools, for example Kahoot! and Duolingo. Additionally, programs can be created to help with smaller sub-tasks, such as vocabulary learning. In this case study, I analysed how a multimodal CALL program was received by a small, Finnish, 5th grade class studying German. Unlike, say, Duolingo, this CALL program was not used via the Internet. It was programmed using open-source, in other words free, software and was available for use on a single personal computer, in the classroom. The program was only available on a single computer, due to administrative restrictions related to the school's computers. The CALL program provided the students with additional modes of meaning to benefit their language learning process. By using the program students were able to view German language words in text form. On top of this, they were also able to hear the word spoken out loud by the program and they were supplied with a visual depiction of the word, in the form of a 2D picture. The possible benefit of this multimodal output would be the added context provided by the multiple modes. This way students would encounter a single word in three different ways, at once. Students interacted with the program, by using a touchscreen. During the study, the participating students took part in two vocabulary tests, for which they were able to study using the CALL program. Data on their performance in these two vocabulary tests was gathered. Data was also gathered on their pre- and post-study vocabulary test performance. Additionally, data was gathered from the students and the teacher of the classroom, via feedback forms. The results of this study show that the students were open to the use of the CALL program and that, according to the teacher, the students were motivated by the introduction of the program. The feedback showed that many of the students would be willing to use computer technology in language learning, rather than traditional textbooks. However, due to the small number of available participants, it was impossible to draw any wider statistical conclusions based on students' vocabulary test scores or their multimodality preferences. Nevertheless, despite the program's limitations, from a contemporary technological viewpoint, the language learning classroom was able to benefit from its use of an open-source language learning tool.

Key words: CALL, computer tool, human-computer-interaction, multimodality, vocabulary learning.

Table of contents

1 Introduction

During the past few decades, the processing power of computers has increased greatly, and as a result so has the number of innovative ways, in which computers can be utilized for information sharing. Computational processes that used to require computers the size of entire rooms can now be done using mobile phones, which fit in the palm of your hand. It should be remembered, however, that as with traditional books, the main purpose of computers has been to present information to a human user via letters and other such interpretable symbols. Even seemingly more complex constructs like video and audio are constructed based on binary ones and zeroes, on the file level. In this way, both books and computers still share the characteristic of being linguistic constructs. Computers are complex machines, the workings of which rely on the syntax of a variety of symbols, arranged in a particular order. Computers can even contain whole libraries of books and one can think of computers as evolved books. In fact, computers have been used as learning tools for decades: “computer education in the 1950s centered primarily on numerical methods and the ability to express solutions in the new language of computers” (Lee 2006, p.1). In the United States, the first computer science classes started in the 1960s (2006, p. 9). Language appears to be a central element when it comes to interacting with both computers and humans. What about using computers to help humans learn different things, such as natural languages?

Most people have likely, at some point, encountered computer programs, which claim to be able to help the user learn different subjects or skills. An example of such a program would be Moppi, originally developed by the Finnish company Mikrolinna Oy, which was a PC program designed for young school children, to help them practice their math skills. Other popular, modern examples of learning programs are the widely popular, commercial, online language learning tools Duolingo and Kahoot. There exists a wide variety of modern computer technology, which can be utilized to create learning software like this. Bateson and Daniels (2012, p. 127-138) list the following four main categories (in bold) and their subcategories:

Multi-server technologies:

This category includes; online groups and communities, online resource sharing, online conferencing, online collaboration and resource creation, online LMSs (Learning Management Systems) and teaching services, online virtual worlds and gaming.

Single-server technologies:

This category includes; learning management systems

Single personal computer (PC) technologies:

This category includes; file editing tools, quiz authoring software, screen capture tools, self-study CDs or DVDs

Mobile technologies:

Mobile content delivery and mobile apps.

The aforementioned Moppi program is an example of a single personal computer technology, while Duolingo serves an example of a multi-server technology, used via the Internet. Learning software exists on many digital platforms. The potential offered by such a multitude of technologies does sound promising. Replacing older methods, like textbook teaching, with new digital ones sounds like a no-brainer solution to perceived problems. Students keep forgetting to bring their books to school? Make a digital version downloadable through a website. Problem solved? Perhaps, perhaps not.

Often times, in modern popular discourse and media, we hear talk about the amazing potential of utilizing technology in the classroom. Studies have shown that even very young children are able to successfully use computers and click on objects on the screen (Donker 2007, p. 615). Using computers in modern elementary school teaching appears to be a viable option. There have been many studies made about the many aspects of computer usage of young school children. One such study (Vungthong et al. 2017) explored the use of pedagogic multimedia tablet apps by primary school children, in Thailand. The apps contained simplistic English language songs, accompanied by pictures and written words, which the children could view, as part of their lessons. One of the things the researchers found was that, although, the multimedia tablets were useful as learning tools, the teachers' guidance was still vital to the learning of new words, especially those with more abstract meanings (2017, p. 54).

Computer Assisted Language Learning or CALL is a field of research, which studies the effects of computer technology on language learning. The field has been rapidly expanding in recent years and many research articles have been published (Hubbard 2009, p. 1). In the

digital age, many schools have begun view computers and digital media as useful tools to help students in their studies. Adopting computers as learning tools on a large scale could prove to be beneficial. This adoption of technology has not been not without its critics, however. For example, in his book *Oversold & Underused – Computers in the classroom*, Larry Cuban, former president of the American Educational Research Association, presents the viewpoint that, in reality, students do not use computers frequently, in the classroom. His findings were based on his own observations of computer usage in different American schools. Cuban found that: "For the most part, teacher-centered instruction was the norm, even in computer-based classes." (Cuban 2001, p. 95). With this in mind, we might argue that the idea that the effective use computers in the classroom is as simple as placing them there can be called into question. It is not as simple as; computer plus class equals results. However, Cuban's book was written over 20 years ago and in our modern world, the demand for more digital education tools just seems to keep on increasing.

There is a list of factors that have led to more demand for the digitalization of educational materials. Examples of this are the desire to use less paper and the demand for more universal access. It seems that many are ready to accept the notion that digital beats paper. On the one hand, it is arguably easier to provide, for example, online e-books than it is paper books. Additionally, digital computational tools have the added benefit of multimodality. Features like audio, improved visuals, user input and output, often times, offer more variety to the language learner than just a book.

A recent real-world example showcasing the apparent benefits of digital tools in learning, would be the worldwide adoption of online teaching, during the 2020 - 2022 COVID era. During which, due to widespread legal restrictions, students in many countries were not able to attend face-to-face school classes. Instead, students attended online classes, using different applications like the popular video chat application ZOOM. Remote learning was a widely talked about phenomena, at the time. Exemplified by articles like "*Coronavirus Is Shutting Schools. Is America Ready for Virtual Learning?"* (Goldstein 2020) and "*5 Reasons Zoom Schooling Is Detrimental To Children*" (Ringelstein 2020). Many news articles were written about it, by the media. In many cases, teachers and students faced difficulties, when using these proprietary, applications. This is turn, sometimes, led to a decrease in lesson quality. On the other hand, however, teachers and students also discovered many new benefits to online teaching. One such example of this would be the availability recorded online classes, which could be viewed whenever. More students could now revisit classroom teaching recordings,

which had not been previously made. If a student missed a lesson for whatever reason, they were now able to view it despite their absence. However, all of these experiences varied from place to place, class to class, teacher to teacher, student to student etc. Needless to say, trying to come up with a definitive list of both the negative and positive effects of the use of computers as learning tools would be impossible. Despite this, we should at least be able to analyse some of the general results of the use of CALL software made for language learning. One observable product of a student's language learning process is their vocabulary test score. We should be able to observe the outcomes of language learners using a CALL tool in classroom vocabulary learning.

The benefits of analysing the effects of CALL software in an area like language learning should be clear. If we are to further adopt more digital means of learning in our educational system and gradually make the classroom more familiar with digital tools, the effectiveness and use of a variety of different digital methods should be analysed and studied.

In this case study, I explored how of a small, open-source multimodal CALL tool, with an easy-to-use graphical user interface, intended for vocabulary learning, was received by a classroom of Finnish fifth grade elementary school students studying German. The CALL tool was constructed by the author of this study, specifically for this study, using the freely available, open-source Python programming language. The program itself gave the students output in a multimodal fashion, taking advantage of the multimodal capabilities of a computer program. The output presented a queried word in three different modes: spoken, visual and written. The program was available for use, via a single personal computer equipped with a touchscreen. The students used the CALL tool to practice for two vocabulary tests. Data about the students' pre-study vocabulary test score averages was also gathered. On top of this, data on students' post-study vocabulary test averages, from the Fall of 2022, was gathered.

Additionally, feedback was gathered from the classroom students and the teacher. This feedback data was used in further analysis. Based on the feedback form, I analysed the participants' perception of the program as a language learning tool. Was the program seen as incompatible? What about the test scores of the students? Could looking at both the feedback and the vocabulary test scores reveal something to us, about the use of the CALL tool? Is a simplistic, open-source CALL tool available on a single person computer, a viable learning tool in a modern classroom? The setting of this study was also compared to the results of similar, earlier studies in the field of computer assisted language learning. In this case, focus

was on the use of a CALL tool to learning German vocabulary. German is an optional subject in Finnish schools. Many earlier CALL studies have utilized English language learners or learners of official languages in the country where the study has taken place. In short, the main questions that this study sought to answer was: *How is an open-source CALL tool, available on a single personal computer, received by a small language learning classroom and its students, who are more familiar with the use of textbooks when learning an optional language AND what effects, if any, might it have on their learning?*

Based on earlier CALL studies which have reported positive findings, we can hypothesize that a CALL tool can, in fact, be a valid tool for language learning. However, many of these earlier studies have focused on more "popular" languages and have relied on more readily accessible/available technologies, such as *Kahoot*, which can be used via a web browser. Other studies have utilized cell phone technology. An over-reliance on network-bound and proprietary tools may have some limitations. One possible example of a shortcoming would be the necessity of always having access to a working server. What if a service is not available, during a planned classroom lesson, for example? Or what if the Internet is down? In this case study, I decided to utilize freely available open-source software, for the compiling of an open-source CALL tool for vocabulary learning. It is important to assess whether a small open-source CALL tool, not used via the Internet, constructed on a non-existent budget, can still offer something to a 2022 language learning classroom.

One might, initially, think that conducting a study on the use of CALL in a modern classroom is as easy as turning on a light switch or pressing a button. However, the reality of conducting such a CALL study, is more complicated than one might initially think. Technology simply existing on its own does not guarantee that it will be useful or even used at all, even in the age of broadband Internet. Perhaps, Larry Cubans observations still apply today? During this study, plans often had to give way to the practicalities of real school work. Just as any class room of language learners, no implementation of digital technology is completely identical to another such situation. Therefore, on top of other analyses, later, I discuss the more practical side of conducting a study of this kind. Naturally, these insights are anecdotal in nature, however, I do believe that they provide valuable supplemental information for the reader about the study's overall context. Mainly, these observations relate to the practical and technical limitations, related to the implementation of the CALL software that was used during this study.

In this introductory chapter, I have introduced the overall premise and theme of this study and how it fits into the field of CALL. In the second chapter, I will go over some examples of earlier CALL related studies. I will also go over some CALL terminology and other relevant concepts, such as human-computer-interaction and multimodality. In the third chapter, I will explain, in-depth, the methodology used to conduct the study. I go over the materials used and describe the study's classroom setting. In the fourth chapter, the findings and results of the study will be presented. In the fifth chapter, I discuss these findings and also compare this study to earlier CALL studies. In the sixth and final chapter, I will give my concluding thoughts on the findings of the study and their possible implications.

2 Theory

2.1 CALL

In his book *Computer-Assisted Language Learning – Context and Conceptualization*, Michael Levy offers his perspective on the multi-dimensional field that is **Computer Assisted Language Learning** or **CALL**. According to Levy (1997, p. 1), CALL "may be defined as the search for and study of applications of the computer in language teaching and learning", a field that touches on computer science and linguistics. Just as there is no shortage of acronyms used in the field of computer education, there is also wide variety of acronyms related to the field and study of computers in language learning (1997, p. 77-80). Levy lists the following five: CALL (Computer-Assisted Language Learning), CAI (Computer-Assisted Instruction), ICALL (Intelligent Computer-Assisted Language Learning), CELL (Computer-Enhanced Language Learning), TELL (Technology-Enhanced Language Learning). According to Levy: "Each term suggests a particular focus which tries to encapsulate the use of the computer in language learning". Often, the difference between terms like CALL and TELL can boil down to a difference in what aspect of computer learning is focused on and the researcher(s) own point of view. Levy himself opts for the use of the widely used term CALL. Among his reasons for his preference for the acronym, is that it conveys the wide range of role computers can have in language learning (1997, p. 82). It should be noted, that with the development of new computational technology, many more CALL adjacent acronyms have emerged. An example of this would be DGBL (*digital game-based learning*), used by researchers like Hwang and Wu (Hwang & Wu 2010). However, this term is strongly tied to modern video games, so its applicability is not as flexible nor is its focus as wide scale as CALL's. In their study *Game On With Kahoot!: Effects on Vocabulary Learning and Motivation* (2021), Eric D Reynolds, Richard W. Fuchs and Peterson, which uses yet another acronym GBRS (*Game-based student response systems*), state that: "The landscape of technology in language classrooms is changing so quickly that it's hard for us to keep pace" (2021, p. 40). For the purposes this study, CALL was chosen as the term used, when referring to computer technology used in language learning.

As a field, CALL has been around for a while. Levy (1997, p. 15) points to The Plato Project, started in the year 1960, as a starting point for CALL. The initial PLATO (Programmed Logic for Automatic Teaching Operations) system spawned three sequels. Another early example CALL technology was created, at Brigham Young University, in 1971; the Time-Shared,

Interactive, Computer Controlled Information Television or TICCIT (1997, p. 18). Development in CALL and computers in general continued in the years that followed, largely thanks to the creation of the microcomputer, in the early 1980s (1997, p. 22). New inventions like LAN (Local Area Networks) allowed for larger, more widespread computer networks to be created than ever before, enabling the spread of many more learning systems. In the 1990s, many developments in CALL were largely tied to the rapid explosion of the worldwide Internet (1997, p. 31). With the help of the Internet and other developments like the creation of the smartphone, more people had access to new different CALL materials and technologies, via the web and other means. This advancement and development of new software and technologies has continued to this day.

There are a multitude of ways to construct a piece of CALL software, thanks to the "essential neutrality" of computers (1997, p. 84). Nowadays, there exist a number of freely available software tools, which can be used to create new CALL technologies. These "open-source" software tools enable people in different fields to create educational content. However, creating high-quality CALL content can often end up being an expensive endeavour. (Johnson, Brine 2012, p. 92-94).

Narrowing down an exact definition of the field seems a daunting task. Like most modern fields of research, CALL benefits greatly from shared interdisciplinarity with fields such as linguistics and human-computer interaction. Levy (1997, p. 49) supplies a long list of disciplines and fields which have relevance to CALL. He narrows these down, to five smaller group categories: psychology, artificial intelligence, computational linguistics, instructional technology and design, and human-computer interaction studies.

Due to the development of more sophisticated and powerful computer technology, in the 1990s, there emerged a desire to make computers more user-friendly (1997, p. 69):

> Now the focus is on design that is sensitive to the needs of users, and the cognitive principles that lead to more intuitive computer systems. The discipline of Human-Computer Interaction addresses such concerns

In the study of **human-computer-interaction** or **HCI**, the focus shifts from the computer to the user. In other words, HCI can be thought of as the study of humans in the role of a computer user. When analysing computer assisted language learning, human-computer-interaction is a large point of interest. Aside from the traditional keyboard and mouse, modern computers offer a wide range of methods for a user to interact with a computer, for example

the touchscreen. By 2016, such technologies were seen as permanent fixtures in the lives of many (Elkind, 2016).

In recent years, advanced techniques have been developed to make the interaction between humans and computers even more seamless and "easy". The field of multimodal human-computer interaction is another interdisciplinary field that utilizes knowledge from other fields, such as cognitive science and psychology. This interdisciplinarity allows for the study of things such as eye tracking and sentiment recognition. (Alejandro & Sebe, 2007, p. 116). However, for this study, I considered human-computer-interaction as simply a descriptive name for the classroom situation, in which students press buttons on a computer touchscreen. Here, we are observing simple input producing output. Information about this human-computer-interaction was gathered, via a written feedback survey containing relevant questions. In this study's pre-study survey, information about the participants' frequency of computer usage, outside of school, was gathered. It is possible that a student's prior computer skills may factor in on their ability, to better adapt to using a CALL program.

Levy brings up the idea of a tutor-tool framework, in CALL (Levy 1997, p. 178-191). Meant to function as "a means of providing an organizational schema in a diverse and complex filed". Given the numerous different ways that CALL projects and programs can be constructed, it is important to try to form effective methods of categorization. The tutor-tool framework includes two terms to describe two possible roles of a computer: the computer tutor and the **computer tool**. The essential difference between the tutor and the tool is that the tutor evaluates, whereas the tool does not. The computer tutor can judge user input to be either right or wrong. In other words, the tutor evaluates *correctness*. The term computer tool refers to a perceived fundamental role of the computer, as a means of improving the working efficiency of the user. From the perspective of the computer tool, the **focus is on what the users do and how a computer can help them perform their task better**. When a learner uses a computer tool, they are, in a sense, **interrogating the computer**. In this case, **the user assumes a more pro-active role within the interaction**. With their own inputs, the user is responsible for the production of output. Levy refers to this as **conjectural learning**. The opposite of this would be revelatory learning. In this case, the information is supplied by the computer to the user as output (1997, p. 191-192). Levy states that both of these roles have their own drawbacks, evident in different scenarios. One clear weakness of the computer tool is that it offers no feedback for the user (1997, p. 205-209).

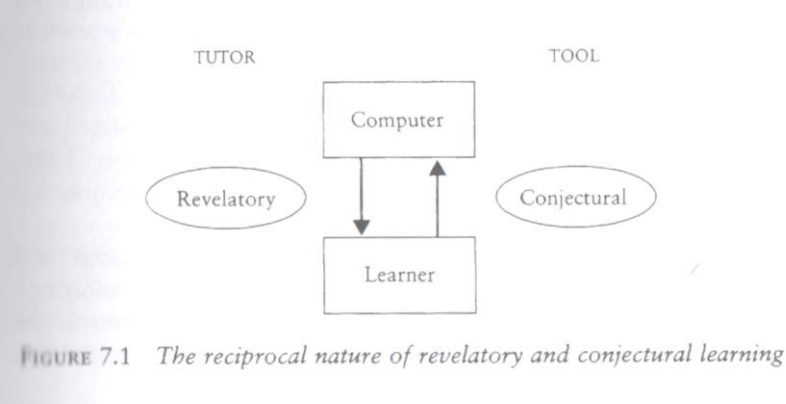


Figure 1. The tutor-tool framework. (1997, p. 191)

In this study, the CALL program the participants interact with, can be considered to function in the role of computer tool. However, to simplify matters, I have chosen to refer to the CALL program used in my study, as a *CALL tool* (*Computer assisted language learning tool*). We can think of the term CALL tool as referring to a computer tool, created for purpose of enhancing language learning. I believe this term best conveys the exact nature of the tool used in this study.

2.2 CALL studies on vocabulary learning

In his 1997 book, Levy described the scope of CALL, rather accurately, as broad (1997, p. 152). However, Levy's book was written over 20 years ago and things have obviously progressed. By 2013, the CALL field had evolved even further, gaining additional perspectives from different researchers. Thomas et al. (2013, p. 29-30) argue that modern CALL cannot afford to think too optimistically about the often spoken of benefits of CALL, nor can CALL simply be used to replace actual teachers in the classroom. In their eyes, modern CALL research needs to pay closer attention to the learners themselves and not bind itself to technological advancements or restraints. Contemporary CALL has advanced to a point where international CALL associations have been formed (2013, p. 25).

Many different studies on the effects of CALL, or software similar in nature to CALL, have been conducted in the 2000s. These studies have sought to take advantage of the multimedia capabilities of CALL software, in one way or another. These studies frequently involve the analysis human-computer-interaction, where a computer is used as a tool or a tutor, to help the learning process. Some studies have focused on the use of cell phone technology, while others have used different computational means, like laptops and tablets. Virtually any of the technology categories, listed by Bateson and Daniels (2012) can be utilized in CALL studies. Given the lack of a clear, all-encompassing, universal theory for CALL and its interdisciplinary nature, these studies and the way they have been conducted have varied tremendously.

A study by Segers and Verhoeven (2003) analysed the effects of computer software on the Dutch language learning of kindergarteners living in the Netherlands. At the time of the study, most of the earlier studies concerning computer assisted language learning had focused on "*living books*". The study analysed the effects of gamified vocabulary learning on 164 kindergarteners from the Netherlands. The software used in the study contained a storytelling function, used to teach the students new words, accompanied by drawn images and other graphical features (2003, 560). The software was on a CD-ROM and used on a computer, therefore, it can be categorized, according to Bateson & Daniels' (2010) technology classes, as a piece of single PC technology and more accurately a self-study CD or DVD. The software included parrot and a pirate character, which offered the student helpful tips, in case they were having problems. This would indicate that the software was built to function as a computer tutor, rather than a computer tool. The program also came equipped with vocabulary

training games, which the children used to train their vocabulary skills. The children's vocabulary gains were then tested, later on.

During the vocabulary tests, the children were shown photographs, corresponding to a word (2003, p. 560). During training, the computer assisted training program had used drawings, instead of photographs to query words. This was done to prevent the students from "– viewing similar materials during testing and training". The study found the CD-ROM software to have had a positive effect on the learning process (Segers, Verhoeven, p. 564). These positive effects were also seen during the second period of testing, where less controlled computer time was afforded. Additionally, the study found that: "There was no correlation between the number of games played and the children's learning progress - ".

CALL studies have also compared different user groups to one another. Rob Hirschel and Erik Fritz's 2012 study compared the differences between three language learner groups. All of the 141 participants were university students from Japan, learning English. Out of the three, two of these groups were instructed to study "using a CALL program or using a vocabulary notebook". The CALL program, designed for the purposes of the study, was hosted on the website Praxised.com. The third group received no explicit learning instructions. The researches also wanted to see which learning technique resulted in "the largest gains as measured by a pre and post-test". Overall, the study focused on the long-term retention of 36 specific words, included in the course materials (2012, p. 643). After the five-month period of pre- and post-test the results were looked at (2012, p. 649):

> The pre- to post-test gains achieved by the CALL group (41%) and the vocabulary notebook group (38%) were not statistically different, meaning that both treatments fared equally well in the short term.

The control group, on the other hand, fared worse and experienced lesser gains than the two other groups. The study found that students using the CALL program made adequate vocabulary gains.

CALL studies are not tied to a single category of computer technology. Lu's 2008 study *Effectiveness of vocabulary learning via mobile phone* (Lu, 2008), involved 30 Taiwanese high school students and it analyzed the effects of SMS messaging on their vocabulary. In Taiwan students' exposure to the English language was seen as lacking (2008, p. 515). This motivated the researchers to find out whether students vocabularies could be improved using texting, as mobile phones were readily available in Taiwan. This study investigated SMS

messaging, a form of mobile technology and more accurately, mobile content delivery (Bateson, Daniels 2012, p. 137). The students were split into two groups and the other group was provided with the SMS messaging containing targets words to be learned. The next week, the groups would switch places and the other group would be studying using the SMS messages (Lu, 2008, p. 517). The SMS lessons were sent twice a day, every day from Monday to Thursday. The results of the study appeared to be positive (p. 519):

> Overall, the mobile phone groups have greater vocabulary gains than their paper-group counterparts in both immediate and delayed post-tests.

The students' attitudes towards the learning strategy were also found to be mostly positive (2008, 521). The study found that "the immediacy and novelty of SMS lessons and its manageable amount of information can foster students' vocabulary learning." The researchers' earlier worries about the potential downsides of using small mobile phone screens, did not seem to have a noticeable effect on the learning outcomes.

Different formats or modes of media can also be utilized in CALL. In another study from Taiwan, Lin and Lu-Fang analyzed the use and effects of "a video-based CALL program". 44 more or less skilled and 39 less skilled second language learners took part in the study. These participants were then split into four sub-groups, based on their earlier English scores (2010, p.205). The CALL program used in this study, serves as an example of single PC technology as it included multimedia resources (Bateson, Daniels 2010, p. 135). The CALL program used was a piece of video software which included dual-subtitled video footage (Lin, Lu-Fang 2010, p. 206). The program had two main visual aspects; the text interface and the video interface. Many helpful tools were also available for the user of the program. The users could change the speed of the spoken audio, using the "Audio speed" option. A digital dictionary was also available for the students. Some of the other identifiable tools included in the program were the following: reference words, pronunciation practice, repeat, and role-play.

The researchers found, that "The statistic results suggest that this activity significantly enhanced less-proficient participants' incidental vocabulary acquisition and comprehension." (2010, p. 199). They also found that adjectives and nouns were more effectively acquired than adjectives, by the participants.

Advanced computational constructions, like video games, have also been the subject of numerous CALL studies. A study in 2010 found that, between 2001 and 2010, the amount of research done on digital game-based learning (DGBL) had increased significantly (Hwang,

Wu, 2010, 9). The number of studies published on the subject between 2006 and 2010 was four time the number of studies published between 2001 and 2005 (2010, p. 7). It was also found that most of the studies were focused on "the investigation of students' motivations, perceptions and attitudes toward digital games in the decade" (2010, p. 8). Most modern games make use of the 4 technological categories mentioned by Bateson and Daniels (2012).

One analytical study on the effects of DGBL was done in 2012, by Liou. 25 third-year English language college students participated in this study. The students engaged in language learning activities, using the still popular, online game Second Life (Liou 2012, p. 370). In the game, users were able to interact with the virtual environment and other users using customizable, digital avatars. The game supported both text and voice chat. The digital space also enabled the use of context offering environments (2012, p. 372):

> The students could ride animals in zoos or play the facilities in amusement parks. In addition, some places that are designed realistically added authenticity for the participants. For instance, when students went to Paris, they could see the Eiffel Tower and the Arc de Triomphe, which were similar to the real ones.

In the questionnaire, the students remarked on the unique learning environment offered and possibilities offered by the game Second Life (2010, p. 374). The participants' experiences using the game were mostly positive. The digital environment offered by the Second Life game was concluded, by the author of the study, to be a beneficial learning tool. Second Life is an example of a Bateson and Daniels' modern online group and community technology category, which utilizes the Internet.

A 2018 study focusing on the "effects of digital game-based learning (DGBL) on vocabulary" found that the style of game may have an effect, on the rate of vocabulary acquisition. The researchers use the term Digital Game-Based Learning or DGBL, to refer to a field of research, which studies the use of games in learning. The study found that players' vocabulary acquisition is highest in games which "require higher mind functioning such as critical thinking, problem solving and task engagement" (Chen at al. 2018, p. 73). Adventure games were thought to be better at this than non-adventure-based video games.

The CALL tool used in this study was much less complex than your average modern video game. It utilized older programming libraries and did not contain things such as 3D animation. As a tool, it also lacked a narrative element. Comparatively, however, creating a

CALL tool of this kind, with “older” technology is a much less time and resource consuming effort. This is one of the main benefits of using it.

Despite the rising popularity of complex new technologies, like videogames and other recent innovations in computer technology, modern CALL researchers have not neglected the study more simplistic, tool like CALL software. A 2016 study by Takeshi Sato analyzed the potential of using a multimodal dictionary as a learning tool. The dictionaries depicted prepositional words like; “above”, “below” etc. with picture to better illustrate their meanings. (2016, p. 5) The bilingual programs used in the study contained both 2D and 3D images. Two separate studies were conducted. In the first study, fifty-two, non-English major, Japanese university students were instructed to use a program on a Moodle site.

The program would display what Sato refers to as “visual glosses”, in other words, visual representations of individual words, alongside sentences (Sato, p. 7).

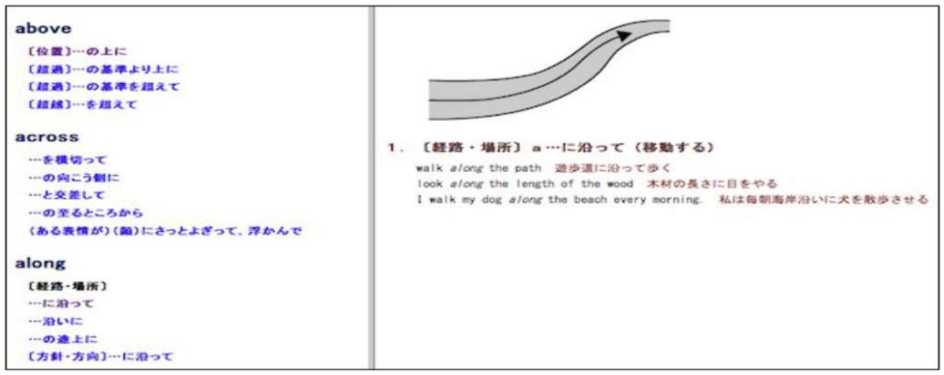


Figure 2. A screenshot of Sato's program

In his paper, Sato states (2016, p. 9) that the results of the study possibly indicate that the use of simplistic 2D images may be better suited for CALL programs, because 2D pictures “allow the learners to change the images in their minds to apply the images to each context -- “.

In 2018, Cakmak et al. Conducted a comparative study on people learning English as a foreign language. The learners were randomly split into two groups: The experimental group (EG) and the control group (CG) (2018, p. 1). Via electronic devices, the EG group was given access to “Mnemosyne Software”, which included a selection of flashcards designed for learning (2018, p. 3-4).

> The Mnemosyne program is an outcome of the project and runs as a flashcard program that aids in learning specific subject matters and general world knowledge through custom flashcards.

A pre-test and a post-test were conducted to measure the participants vocabulary gains. The results of the study and the post-test results indicated that the EG group had benefit from the CALL learning tool and had performed better on average than the CG. The researchers conclude their paper by hypothesizing that combining similar CALL software with mobile application technology (MALL) could potentially improve language teaching all around the world (2018, p. 6). Here we can see the appearance of yet another, new, CALL-adjacent acronym: MALL.

In recent years, a great number among the different CALL programs related to language learning have been developed. The functions and formats of these programs also vary and fit into most of the categories pointed out by Bateson, Daniels (2010). One recent example of a CALL program, designed for a somewhat "niche" purpose, would be the Japanese Rakugo program (Yamada, et al., 2019). The program was designed to overcome some of the limitations present in more traditional Japanese L2 teaching materials. Many teachers have stated that finding good teaching material for the L2 classroom about humor, for example, is difficult (2019, p. 149). The decision to choose Rakugo as the focus of the CALL program is explained as follows (2019, p. 150):

> Because of its cultural references and wordplays, Rakugo is potent material for improving pragmatic competence if framed in a learner-friendly format. Rakugo presents an ideal platform for exploring the possibility of integrating language and cultural learning that ultimately has the potential to improve pragmatic competence.

The Rakugo CALL program included multiple video recordings which offer the L2 learner helpful, context providing, additional material in their studies. The program's main goal was to offer the learner an opportunity to view "authentic humor material" (2019, 153). Despite not being able to experience a real-life environment, the L2 would still be able to "acquire linguistic and cultural comprehension to improve their pragmatic competence in Japanese" (2019, p. 153). Understanding Rakugo, requires more than pure linguistic knowledge.

As we can see, many CALL studies, such as the one by Sato (2016), have tapped into the fact that CALL can utilize multiple modes of communication in order to better facilitate the occurrence of the language learning process. Sato's program was also not reliant on state of the art gaming technology. Many of these studies have involved software, on different

platforms, specifically designed to address different language learner needs. Often times, these tools have been used by students to learn languages required by a school curriculum or potential employer. Many modern CALL studies have benefitted from the existence of advanced games with 3D graphics and such. However, when it comes to for-profit games, such as Second Life, language learning is not always the main focus of the developer. CALL studies like Liou's 2012 study, often rely on the existence of readily available programs. Taking into consideration the seemingly never ceasing popularity of video games among people, it is important to assess whether cheaper, more "niche" and "simplistic", less resource heavy CALL tools can still function as effective learning tool in a modern classroom. An obvious benefit of such custom, open-source programs would be their flexibility and low price. Presumably, the multimodal benefits of such CALL tools do not simply vanish, as a result of technological progress.

2.3 Multimodality to benefit CALL?

Dual coding theory or *DCT* is a well-established general theory, which claims that human cognition processes things in two different ways. The theory states, that there exists one cognitive system, which specializes in processing *language* and another system for dealing with *the nonverbal side of meaning making*. Essentially, DCT states that when a word is read, in addition to the brain processing the literal written form of the word, it also forms a mental, non-verbal representation of the same word (Sadoski & Paivio, 2013, p. 28-29). Sadoski and Paivio summarize that DCT is a theory, which can be applied to multiple different frameworks, such as constructivist and cognitive theory or as a theory of literacy (2013, p. 1-6). For the purposes of this study, we can assume DCT as proof of the idea that human beings are able to process language in more ways, than one.

Multimodality refers to the concept of using more than one "mode" to communicate information. Bezemer, Jewitt and O'Halloran outline three different core characteristics of multimodality (2016, p. 3):

> 1. Meaning is made with different semiotic resources, each offering distinct potentialities and limitations.
>
> 2. Meaning making involves the production of multimodal whole.
>
> 3. If we want to study meaning, we need to attend to all semiotic resources being used to make a complete whole.

Essentially, multimodality refers to the use of multiple different forms of communication. For example, instead of simply having to use text alone, meaning can be conveyed using other methods, such as imagery. This definition of multimodality seemingly accepts the premise of the dual coding theory, since it uses non-verbal means for communication.

Modern computer programs utilize a graphical user interface and can produce sound. This enables them to take advantage of more than one mode of output. In other words, computers can make use of multimodality, ergo, so can a CALL tool. Many of the studies, already mentioned earlier, utilized multimodality, in different ways. The use of multimodality, especially with technology combined with learning has received a lot of interest, due to its potential benefits and has been the subject of many studies, in recent years. Norbert Pachler (2001, p. 24) states that:

> CALL can be seen to represent a natural extension of 'traditional' resources, such as print-based encyclopedias, dictionaries and grammar books, harnessing some of the potential of new technologies, in particular, convenience and speed of access to reliable information as well as multimodality

Magnusson and Godhe (2019) argue that the use of multimodal compositions has established itself as an everyday method thanks to the increase in the use of digital technology, which allows us to create such materials, for example. Warschauer, M., & Healey, D. (1998, p. 83) list a number of benefits to the use of computers in language learning, for example, the so-called "fun factor" and the variety of resources available. Researchers can use a variety of survey questions, to assess the effectiveness of different CALL tools (Dunkel, p. 211). These questions can relate to topics like student opinion on CALL (p. 218).

A 2014 study by Bisson, Marie-Josée et al. found that, in incidental learning situations, learners can acquire lasting knowledge about a lexical item's form and meaning, from as few as two encounters with relevant multimodal stimuli. This would suggest that, with CALL, vocabulary acquisition can be made easier for foreign language learners. They also noted that this sort of fast acquisition of knowledge is maybe more likely to occur with less abstract word classes, such as proper nouns. In situations where the meaning of a word is not as apparent, more repetition may be necessary for learning to occur (Bisson, Marie-Josée et al. 2014, p. 870-873).

A study by Chun & Plass (1996) found that students' vocabulary scores improved, when they studied materials, where words were annotated with both text and a picture, as opposed to just plain text or text accompanied by video. This would seem to indicate, that the specific type of multimodal mode selected can have an effect, on the learning process and rate of learning. A 2003 study by Yuli and Wang, in which three different versions of a language learning program were compared, that the most effective version of the program was one in which annotation was a combination of text with a still picture (2003, 139-140). The other two versions were one with only text annotation available and one with text annotation accompanied by an image and spoken audio. Another study also found that "information coded both verbally (textually) and visually (pictorially) is more effective for learning than information coded singularly" (Yoshii & Flaitz, 2002, p. 40).

Cárcamo, Melisa Millaray Acuña et al. (2016, p. 150) also found that the use of multimodality in a classroom context improved students' ability to retain new words: "The use of multimodality in the classroom context significantly benefits the students in their process of

learning new vocabulary". The study also found that majority of students preferred multimodality. It would seem, that multimodality could be a useful tool, for enhancing the vocabulary learning rates of language learners. There also appears to be evidence for some learners' desire to use multimodality in learning.

On the other hand, Boers, Frank et al. (2017) conducted a study on adult L2 readers, which found that using pictures as marginal glosses did not improved the learners' retention of those glossed words. The researchers found more evidence for the possibility that it is, in fact, the repeated viewing of glosses that creates a better imprint into memory. The study also found that when glosses contained both pictures and text, learners spent more time inspecting those glosses (2017, p. 719-721). Chapelle, (2003, p. 141) states that by including repetition into CALL, learning efficiency can be improved.

Research has also found, that the addition of more modes of glossing does not always benefit learning. Glosses containing more than two modes may not prove to be beneficial in a practical teaching scenario: "Given the effectiveness of dual glosses and the extra workload involved in adding a third mode, teachers can rely on using dual glosses rather than triple glosses for teaching vocabulary" (Ramezanali, Nasrin Takumi Uchihara and Farahnaz Faez, 2021, p. 127-128). When studying the effectiveness of vocabulary learning using the popular Kahoot! program, researchers found that "while the instruction that incorporated Kahoot! did assist students in learning vocabulary, it was neither more nor less effective than the "regular" instruction in the control comparison group". (Reynolds, Fuchs, Johnson, 2021, p. 48)

It seems that some research has found the use of multimodality to be beneficial, while some studies have been more cautious in their conclusions, or have pointed out other factors which may have led to improved vocabulary retention, such as repetition. Some studies, such as Segers & Verhoeven's 2002 study, have noted that learning gains simply tend to improve, as a result of more time having been spent overall on exercises (Segers & Verhoeven, 2002, p. 219). Presumably this also applies to computer skills in general. For this study, data was gathered on the participants' own preferences, when it comes to multimodality related to the CALL tool with which they interacted.

2.4 Vocabulary test scores

No specific official Finnish guidelines for scoring L2 vocabulary tests in elementary schools exists. However, there are laws regarding on official grading system (Perusopetusasetus, 1998). According to Finnish law, grades can range from four all the way up to ten. With this guideline in mind, many teachers opt to have their vocabulary tests have a maximum score of ten points.

In Finnish schools, a separate larger exam has the final say on a student's language grade, with their performance in vocabulary tests having either a slight positive or negative effect on their final overall grade. In Finnish schools, vocabulary test scores do have an impact on a student's grade in a subject like German, for example. We can take these scores, as data about the students' vocabulary test performances.

3 Materials & Methods

In this CALL study, I set out to see how a simple, open-source CALL tool used for vocabulary learning is received by a small language learning classroom, studying an optional language and whether it had any effect on their learning. Data for this quantitative and qualitative case study was gathered by using different methods, in the year 2022. This was done in order to construct a more "multi-dimensional", analyzable data whole, which could be analyzed from multiple points of view, both quantitative and qualitative. The gathering of different types of data from the participants of this study, allowed me to analyze the students' reaction to the CALL tool and other things, not just numbers.

3.1 Experimental Setup

For this study concerning CALL tool usage, a small class of Finnish, fifth grade German students adopted a multimodal language learning tool, into their lesson plan. With the help of the CALL tool, students were able to receive multimodal output (*image, text and sound*), to help them better understand a given word's meaning, hypothetically. In some of the earlier studies, mentioned before, multimodality has been shown to enhance the language learning experience. However, contradicting evidence has also been found by some research.

Figure 3. A screenshot of the CALL tool's GUI.

A single personal computer was placed into its own designated workstation, on an available table. The students would, one by one, go through the vocabulary list presented by the CALL tool, by pressing words on the screen, using the touchscreen. They were also instructed, by

the teacher, to repeat the words out loud or in their mind, to help with memorization. In the case of this school, using the CALL tool on the school's own computers was not possible, due to administrative network restrictions.

During the study, the students took part in two specially prepared vocabulary tests. In these tests they would be tested on their knowledge of words, which they had encountered while using the CALL tool. Before each of the two tests, the students attended three classes, during which they were able to familiarize themselves with the vocabulary, using the CALL tool. These results of these two tests were then compiled into an anonymized data set for the study. Data was also gathered about each of the students' performance in German vocabulary tests, before and after the study.

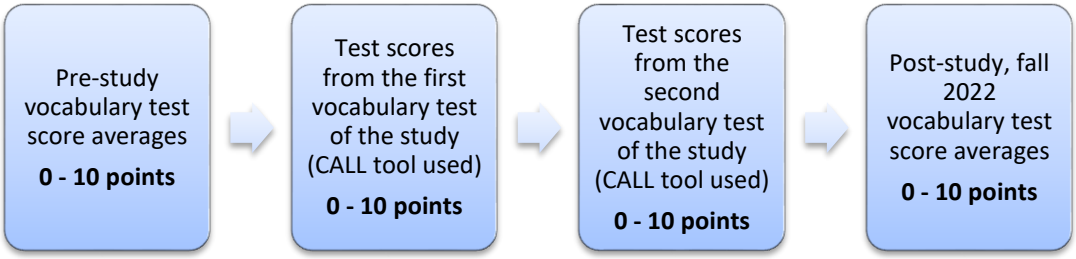


Figure 4. Overview of the four instances from which data on vocabulary test scores was gathered.

On top of the tests, feedback data from questionnaires was also collected from the students and teacher. This data was then used, in further analysis. The feedback form asked students questions about their experience of interacting with the CALL tool and about their opinion of CALL overall. The answers to these questions were then used to form potentially analyzable scoring groups.

3.2 Participants

The participants of this study were students from a Finnish, small town elementary school, participating in German class. In total 11 students participated in the full study. The class had more than 11 students, unfortunately, some additional students were unable to take part in all parts of the study. In some earlier studies, for example Lu 2008, the number of participants in similar studies has been relatively low (30). On the other hand, other studies, like Lin and Lu-Fang's 2010 study, involved a large number of students. The numbers of participants taking part in this study was mostly limited by the relatively small size of the school and lack of German students.

On the other hand, this did make the study more manageable to analyze, in terms of scope and time. It is worth noting that the number of German language students in Finnish schools is already low, since German is a not a mandatory subject. Many CALL studies, such as some of the ones described earlier, have focused on English language learners or learners of a language required by a nation's curriculum. As fifth graders, the students were already familiar with the concept of studying for German vocabulary tests. As indicated by their pre-study vocabulary test averages, the students in this class were highly skilled at German vocabulary tests. Additionally, a smaller school gave an opportunity to observe how CALL tool is received in a smaller classroom, in a rural area school.

As, mentioned earlier, these students were highly skilled at taking German language vocabulary tests. In this school, students had the opportunity to start studying German, in the fourth grade. Therefore, the students already had some experience with studying German. On top of this, German differs from English, another language taught in Finnish schools, in that it is voluntary. Therefore, we can assume that the German language learners of this class were at least somewhat motivated to learn German. The vocabulary which was learned using the CALL tool was selected, due to it being compatible with the rest of the curriculum in terms of being basic vocabulary and appropriately challenging. In the fifth grade, the German students were mostly learning basic grammar and basic "everyday" vocabulary. These words belonged to the noun word class. All of the data collected from the students was anonymized for privacy reasons.

3.3 Use of the open-source CALL tool

Here, I will describe the basic functions of the CALL tool, which was used in this study. To clarify the functionality of the program, I will supplement the explanation with pictures of the program in different stages of use and of the different individual parts of the program.

A screenshot of the CALL tool initial state, before any words has been loaded into it, is shown in figure 5. This is how the program looks when it is initially started. The graphical user interface has been compiled using modules like Tkinter and PIL. Among other things, the Tkinter module enables the creation and use of visible entry boxes, text labels and buttons, shown on the screen. By default, the tool was configured into full screen mode.

Figure 5.

As shown by figure 6, The word *der Fuß* has been loaded into the user interface, seen in the upper left corner of the screen. Using the white entry box, located in the top left corner of the screen, a user can load the words they want from a list, to be made into pressable buttons. They do this by typing out the word, into the entry box and pressing enter. This way, the user (or in this case the teacher) can have better control over which words they want shown during a lesson.

Figure 6.

All 21 words to be trained, in preparation for the second vocabulary exam, have been inputted (figure 7). Now the user is able to "*interrogate*" the tool, for the purpose of conjectural learning. This is done by pressing the buttons on the left side of the screen. Next, we will go over how the program responds to actions taken by the user.

Figure 7.

Figure 8 shows what the screen looked like, if the user were to have pressed the button on the screen containing the word "*der Fuß*".

Figure 8.

An image depicting a foot appeared on screen. In addition to this, two large labels appear above it, these contain the singular and plural German translations of the word. Below the picture, we see the Finnish translation of the word. Additionally, on the right side of the screen, we can observe a white label. This label contains an example of the showcased word used in a real phrase context and the Finnish translation.

When the button was pressed a text-to-speech voice speaks the German word out loud for the user to hear. Admittedly the voice module used here (pyttsx3) sounded rather robotic. However, the reason for using was that this voice module was one of the only ones available for open use, which did not require an Internet connection to function. Most of the more realistic sounding speech synthesizers, such as those maintained by Google, need to be connected to an external server, due to their inherent size and other logistical reasons. The continued functionality of the CALL tool was prioritized here. In other words, it was deemed necessary for the program to work even in an offline scenario or setting.

Next, I will go over the "widgets" that have appeared on screen and their purpose, in more detail. These two "labels", shown in figure 9, contained references to the written forms, both singular and plural, of the outputted German word, in this case der *Fuß* (*a foot*). For students of German, learning both the singular and plural forms of words is an important aspect of vocabulary learning, since the formation of the plural form is not as simple as adding an -s suffix to the end of the word. In this case, the creation of custom-made software is advantageous.

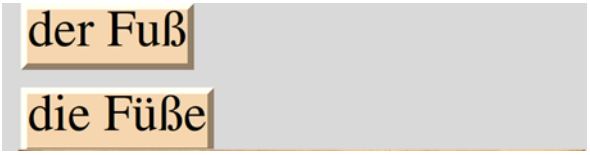


Figure 9.

Here, the user sees the chosen German language word, depicted by an equivalent 2D image (figure 10). The image is situated prominently, in the middle of the screen. Hypothetically, the inclusion of this mode / image should help the user better retain more information about a particular word's character, since they are seeing more than just plain text output on the screen.

Figure 10.

Below the image, another label is displayed by the CALL tool (figure 11). This label contained a reference to the Finnish translation of the German word. The translation, in this case, for der *Fuß* is *jalka*. Since the native Finnish students would already know how to form the plural form of *jalka*, it was not deemed necessary to include the plural form in the display.

Figure 11.

The white text label (figure 12), situated on the middle right, contains an example of the phrasal use of the German word and its Finnish translation. Here, the displayed phrase is *der grose Fuß*, translated as iso *jalka* (*the big foot*). This contextual phrase was added, to better help the user understand the word and how to use it.

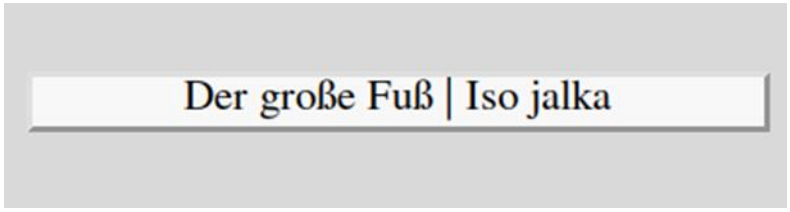


Figure 12.

Alternatively, here (figure 13) is how the screen would have appeared, had the student pressed, for instance, the button containing the word *die Hand*:

Figure 13.

There are also some additional widgets connected to different functions, visible on the screen. These were placed there out of practical necessity. I will quickly go over them here. First, in the bottom right corner of the screen, another entry box can be found. The purpose of this widget was to function as a password protection mechanism, to prevent the students from modifying or erasing what was shown on the screen. By typing in the password, the teacher was able to prevent the students from pressing anything other than the buttons containing the words. Secondly, in the top right corner there is a button containing the letter X. This button could be used to reset the program, in case of a bug or some kind of runtime error.

3.4 Teacher Feedback

For an alternate perspective on the classroom's use of the CALL tool, the teacher was also queried. As the main coordinator of the classroom environment the teacher had a unique perspective on the human-computer-interaction which had occurred, in the classroom. After the last vocabulary exam had taken place, the teacher filled out a more verbose feedback form, regarding the use of CALL tool. The form asked her to reflect on the positives, negatives, limitations (practical things, time, technical difficulties etc.) and surprises they encountered while helping and observing the students use the CALL tool. The teacher feedback form included the following questions:

Question one: What were the differences compared to traditional methods?

Question two: What were the positive aspects of CALL tool use?

Question three: What were the negatives aspects of CALL tool use?

Question four: What limitations related to the CALL tool did you observe?

Question five: Did something about the CALL tool surprise you?

The answers to these feedback questions provided us with information about the use of CALL from the perspective of a pedagogic professional. These answers were then used in later analysis.

3.5 Student data and feedback

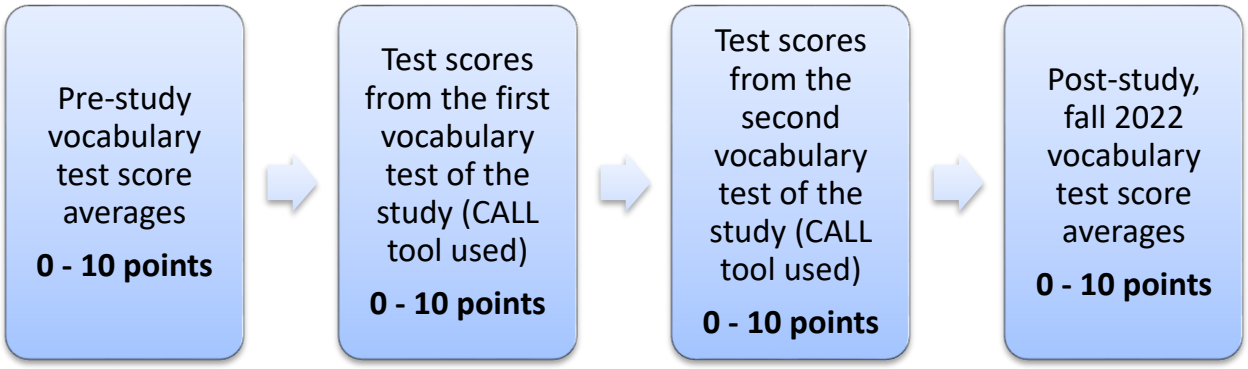


Figure 14. An overview of the different instances from which data was collected.

3.5.1 Pre-study data

Prior to starting the use of the CALL tool, numerical data on the participating students' previous vocabulary test score averages were gathered. In this instance, these scores had the potential to be numbers ranging from zero to ten. This range of scores was selected because it somewhat mirrors widely accepted conventions in vocabulary test scoring, in Finnish schools. The range of points in the vocabulary tests conducted during this study followed the same range. Prior to the study, the students had studied for vocabulary tests using textbooks.

3.5.2 Post-study data

The two CALL vocabulary tests took place before the summer break of 2022. Data was also gathered on the students' vocabulary test averages from the following fall semester of 2022. This was done, in order to ascertain whether or not any changes had happened to the students' average vocabulary test scores, after the use of the CALL tool and the study itself. The scoring scales for the vocabulary tests remained the same.

3.5.3 Data from the two vocabulary tests

11 German language students took part in two vocabulary tests, during the study. The themes and words, that had been previously taught, to these fifth-grade German students, related to basic vocabulary. Earlier studies have indicated, that basic nouns are more effectively learned, using multimodality, than more abstract words. The themes of the words that the students studied using the CALL tool were as follows: names of body parts, names of pieces of clothing and names of animals. In other words, this was basic noun vocabulary. The nature or

difficulty of the vocabulary that the students learned by using the multimodal CALL tool did not differ drastically, from what they had been previously studying, with the aid of a textbook.

During class, the students each got to use the CALL tool in the classroom. In the beginning, they were shown, step by step, how to use the touchscreen program. Enough time was allotted, so that all of the students had ample time to familiarize themselves with the use of the program. The students were told to go through each of the words, presented on the screen as buttons. They were also instructed to focus on the output of each button press and to repeat each presented word in their mind or out loud. This was done, to motivate the students to be more proactive, when using the program and to not just have them simply pressing random buttons, without any thought. Of course, totally ensuring that each student followed these recommendations to the letter was impossible.

In preparation for the first test, the students learned different animal words and clothing nouns, 21 in total. In the vocabulary test, students were presented with ten of these words, for which they would then have to present a translation. In preparation for the second test, the students learned body part vocabulary, another 21 nouns were presented by the CALL tool. As with the first test, in the second vocabulary test the students were tasked with selecting the correct translation for another 10 words. Depending on their accuracy, students could receive a total score ranging anywhere from zero points to ten points. Scores from both of these vocabulary tests were then be manually written down on a datasheet. All of the data on the students' vocabulary test scoring, before, during and after the study, was manually compiled into a single spreadsheet table, containing additional information about the individual participants.

3.5.4 Post-study questionnaire

During their second vocabulary test, the students were also instructed to fill out a final feedback questionnaire. This questionnaire contained questions to gauge the students' experience with the CALL tool. The questions and answer choices were as follows (translated from the original Finnish version):

Question one: The computer program was easy to use?

a) Yes *b) No*

Question two: would like to use these kinds of multimedia programs more often while studying:

a) Yes b) No

Question three: I retained the words in my mind more easily with the program:

a) Yes b) No c) Same as before

Question four: While using the program, I paid most attention to the:

a) image b) text c) sound

Question five: While using the program, I paid the least attention to the:

a) image b) text c) sound

Question six: In the future, I would rather study, using a:

a) textbook b) computer

(At the bottom of the form, the students were also free to provide an additional, optional comment on their experiences of using the CALL tool.)

These answers were used in the analysis part of the study in different ways. First, by simply analysing the answers on the feedback form alone, we discover contextual information about the participants experiences and thoughts. By looking at the answers to the first question, we, for example, can see whether the students perceived the CALL tool to be easy to use, in general. Naturally, this does not tell us whether this CALL tool was an easy program to use. The answers simply tell us about the students' own perceptions. By diving the students based on their answers to this first question on the feedback form, we can form at least two groups A and B. The same can then be done based on the answers to question number four, number five and etc. This way we can see if, for example, students who paid more attention to sound scored numerically higher than those students that focused on the images presented by the program. In some cases, students circled no option or multiple options. For the purposes of this analysis these two types of answers are both treated as meaning "can't answer definitively" and marked as an empty line.

3.6 Python

During this study, participants used a simple GUI (Graphical User Interface) Python program, written for the purpose of vocabulary training during this study. In other words, the students used a simplistic computer program, with a graphical interface, written with the Python programming language. Python is one of the most widely used programming languages available. A language like Python can be used, by anyone, to build and design different kinds of free open-source software. There are also modules available for the Python programming language, which offer additional tools for Python programmers. Documentation and additional information on the structure of the Python programming language can be found on the official Python website; https://www.python.org/ (The Python Foundation, 2021). The website also includes information on compatible modules which come with a standard Python installation, for example Tkinter. Which was used to build the CALL tool used during this study. Documentation on third party modules, like pyttsx3 can be found on the Python Package Index at https://www.pypi.org/ (Python Package Index, 2021). The current, supported and updated version of Python is called Python3.

4 Results

4.1 Student feedback results

Table 1. The data gathered from the students, before, after and during the study.

id	gender	weekly use	pre-study avg.	1st test	2nd test	Fall 2022 avg.	Q1	Q2	Q3	Q4	Q5	Q6	Comment
1	g	rare	9	8	9	9	a	a	c	c	a	b	n
2	b	rare	10	10	8	10	a	a	b	c	a	b	n
3	g	>10h	9	8	7	10	a	a	a	-	c	b	y
4	b	rare	10	8	9	10	a	a	c	a	b	b	y
5	b	>10h	10	5	9	10	a	a	c	a	b	a	y
6	g	rare	9	4	5	10	a	a	c	b	c	b	y
7	g	rare	10	6	8	10	a	a	-	-	c	a	y
8	b	rare	9	3	6	9	a	a	b	c	b	b	n
9	g	rare	10	9	9	9	a	a	c	c	b	-	n
10	b	>5h	10	8	9	10	a	a	a	a	b	b	n
11	b	>10h	10	6	7	10	b	b	b	b	a	a	n

In total, 11 students filled out the post-study feedback forum, which contained six questions and room for an additional comment. Next, we will look at the answers given by the students. After this, we will also go over the feedback about the use of the CALL tool, provided by the classroom teacher.

Table 2. The students' answers to the feedback form questions. Q1 refers to Question one etc.

id	Q1	Q2	Q3	Q4	Q5	Q6	additional comment?
1	a	a	c	c	a	b	no
2	a	a	b	c	a	b	no
3	a	a	a	-	c	b	yes
4	a	a	c	a	b	b	yes
5	a	a	c	a	b	a	yes
6	a	a	c	b	c	b	yes
7	a	a	-	-	c	a	yes
8	a	a	b	c	b	b	no
9	a	a	c	c	b	-	no
10	a	a	a	a	b	b	no
11	b	b	b	c	a	a	no

4.1.1 Question one

The computer program was easy to use?

a) Yes b) No

First, I wanted to know whether the participants found the use of the CALL tool to be easy. Many computer programs are designed for tasks, which require the ability to perform complex actions. However, a CALL program use by fifth grade German students should not be made to be overly complex.

According to the results of the post-test questionnaire, ten out of the 11 students that took part in the exam found the program to be easy to use. We can surmise, that the majority of the class found the CALL tool to not be difficult to use in language learning. Only one of the 11 students found the program to not be easy to use. The same student also reported that they did not enjoy the use of the software.

4.1.2 Question two

I would like to use these kinds of multimedia programs more often while studying:

a) Yes b) No

The questionnaire then asked the students, whether they would like to use multimedia programs like the CALL tool they had been using during the study, more often, while studying. Mirroring the results of the previous question, ten out of 11 students said yes to this question. The answers to these first two questions of the questionnaire would seem to indicate that majority of the students were not against the continued use of the CALL tool, at least, in their own classroom.

4.1.3 Question three

I retained the words in my mind more easily with the program:

a) Yes b) No c) Same as before

The third questions asked the students how they felt about the ease of retaining new words with the CALL tool when compared to traditional classroom learning. Out of the three answers, c was the most popular answer with five students out of 11 picking it. Next b was the

second most picked answer with three students choosing it. Lastly, two students chose a as their answer. Additionally, one student simply wrote “I can’t decide” as their answer. Evidently, this third question began to divide the students’ opinions more than the previous two.

4.1.4 Question four

While using the program, I paid most attention to the:

a) image b) text c) sound

Next, the students were asked which of the three aspects of the multimodal dictionary program they paid the most attention to. The main advantage of using CALL tool in language learning is the possibility of utilizing multimodality. Option c was the picked by five students. After that, three of the students picked option a. Option b was only picked by a single student. Additionally, two of the students had picked multiple choices with one student choosing a and b and one student choosing all options. Evidently, the students’ opinions on the multimodal aspects of the CALL tool varied.

4.1.5 Question five

While using the program, I paid the least attention to the:

a) image b) text c) sound

The students were also asked which of those same three aspects, mentioned above, they paid the least attention to. This time around, more than half of the students, six in total, picked option b. Option a was chosen by three students and option c was chosen by two. Overall, majority of the students found the text of the CALL tool to its least attention-grabbing part. Unlike in the case of question four, all the students were able to pick a single answer to the fifth question of the feedback forum. In other words, all students were able to pick out the singular aspect of the CALL tool they cared for the least.

4.1.6 Question six

In the future, I would rather study, using a:

a) textbook b) computer

The last question on the feedback forum asked the students to pick which one of the two options, a textbook or a computer, they would rather learn things with, in the future. Out of the 11 students, seven picked option b, while three students picked option a. Additionally, one student ended up circling both answers.

4.1.7 Additional feedback

At the end of the feedback form, the students were also able to provide an optional, additional comment. In total, five students provided these. In a pleasant surprise, all five additional comments were positive about the use of the CALL tool in a classroom setting, with one student requesting more use of CALL during classes. This seems to further demonstrate the students' desire, for more CALL in the classroom.

4.2 Post-study teacher feedback

During the classes, the classroom activities were guided by a teacher. For the benefit of the lessons, it was imperative, that the newly introduced CALL tool did not interfere with the flow of the lesson and lessen the effectiveness of the teaching. It is of interest to this study, to find out what possible effects the CALL tool had on the practical teaching and the lesson overall. Feedback was gathered from the teacher, in the form of a feedback form. Unlike the feedback form the students filled out, this form contained no binary yes/no questions. Instead, the teacher was instructed to give more verbose feedback based on their experience of the class taking part in the study.

When asked how the use of the CALL tool in the classroom differed from the use of more traditional methods (Question one), the teacher wrote that:

> Any method that includes some kind of activity and/or thinking is an improvement to just reading the words.

Evidently, they discovered that the use of CALL tool added to the quality of the language learning process. Next (Questions two and three), they were asked to point out some of the perceived positives and negatives of using the CALL tool during the lessons. The first noted positive aspect was an apparent increase in the *motivation* of the students. Secondly, it was noted that the program worked the best, as a teaching tool, when the image shown on screen was clearly depicting the word in question. The third noted positive was the fact that the program turned out to be "*surprisingly effective*". The negatives of using the CALL tool were seen as being overwhelmingly technical in nature and they were also seen as the limitations of the CALL tool (asked about in Question four):

> It takes time to get everyone started when working with computers, so if you only have 10 minutes there's no point starting. Also, if technical problems occur, it easily destroys the whole lesson.

These practical technical problems also became apparent, while conducting this study. On one occasion, for example, a student had, through no fault of their own, simply touched a corner of the screen. This led to the unexpected opening of an assistant program and caused the window of the CALL tool to disappear. As a result, the students were unable to access the program for the remainder of the lesson. Fortunately, this occurred near the end of the lesson. On one occasion, the class was taught by a substitute teacher, who noted that due to the hectic

nature of the classroom, it was sometimes hard or impossible to monitor the use of the CALL tool. During this lesson, however, technical difficulties were entirely absent.

After listing the positives and negatives, the teacher was then asked whether something about the use of the CALL tool surprised them (Question five). They noted that despite the program's simple design the students were still eager to use it during class. The teacher pointed out that this was even more surprising given the fact that, during their free time, some of the students enjoyed playing much more advanced and complicated online games. Lastly, the teacher gave one last comment about the project itself:

> This was extremely interesting. Usually, I have no idea how long pupils need to learn new words, because they do some of the learning at home. Now each pupil spent the same amount of time studying the words, and the differences in learning show. (Note: Obviously, the pupils in the group are fast learners)

From the feedback comments provided by the teacher, we can see that, from the teacher's perspective, the use of CALL tool during this study had perceived positive aspects and some slightly negative aspects. The teaching tool was seen as motivating and as being able to provide new perspectives to classroom planning and teaching.

4.3 Test score data

As mentioned earlier, not every single student's test scores were included in the results. This is due to the fact, that a few of students ended up being unable to attend/study for both vocabulary tests. The unavoidable absence of some students reflects the practical realities of language teaching. It is, therefore, somewhat ironic that the CALL learning materials used in this study were less accessible to the absent students in this case, than their own textbooks. Whatever the case, it should be seen as a success that the majority of the class was able to easily and fully participate in such a learning routine diverging study, despite these scheduling conflicts. In total, 11 Finnish, fifth-grade German students participated in the entire study, with six of them being boys and five girls. Now, I will present the students' vocabulary test score results.

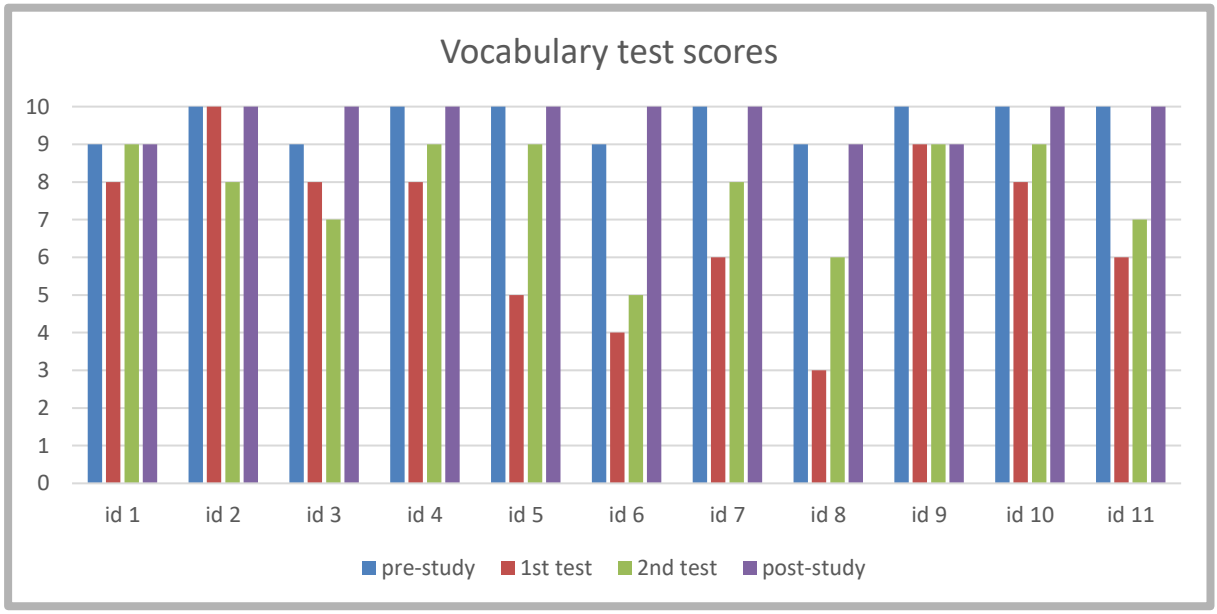


Figure 15. All of the students' four vocabulary test score instances, depicted in 11 histograms.

4.3.1 Pre-study test score averages 2021

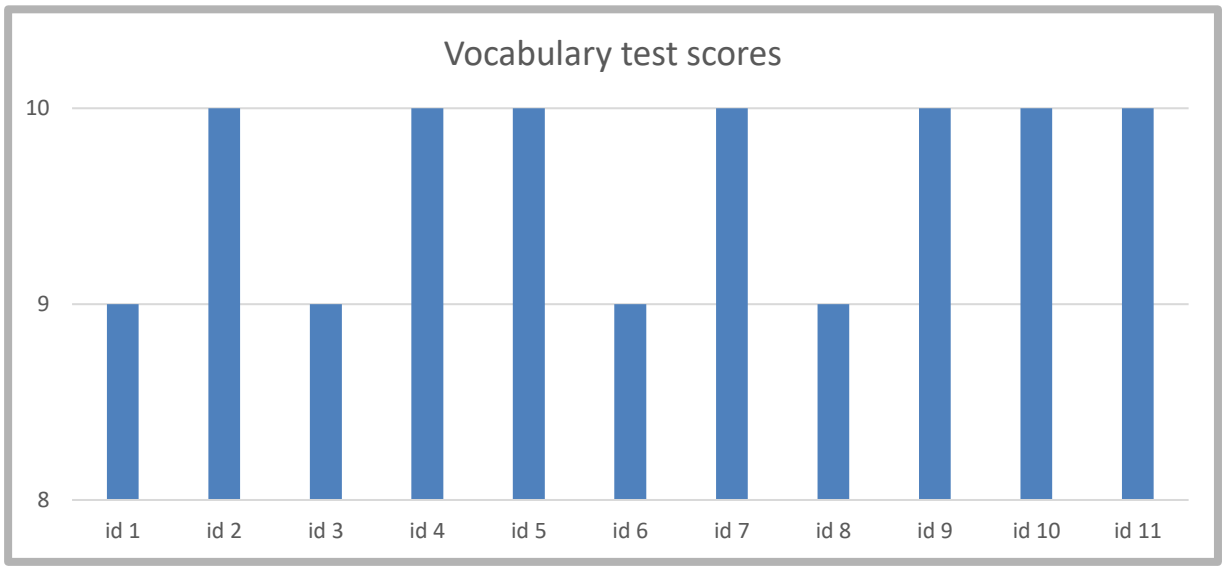


Figure 16. The students' pre-study vocabulary test score averages. A standard deviation of 0.504524979.

The fifth-grade German students, who took part in this study, were reportedly "*fast learners*", according to the teacher. The students' pre-study vocabulary test averages would also suggest that this was true. All the students in the class had a pre-study test score average of either nine or ten, prior to this study. The class average pre-study test score was 9.636~ with a standard deviation of 0.5045~. This further suggests that the students of this classroom were already highly skilled at taking German vocabulary tests. Additionally, since German is an optional school subject, in Finnish schools, we can assume that many of the students were already somewhat motivated to learn German.

4.3.2 Results of the first CALL vocabulary test

Looking at the results of the first vocabulary test, one can see a lot of variation between the student's test scores:

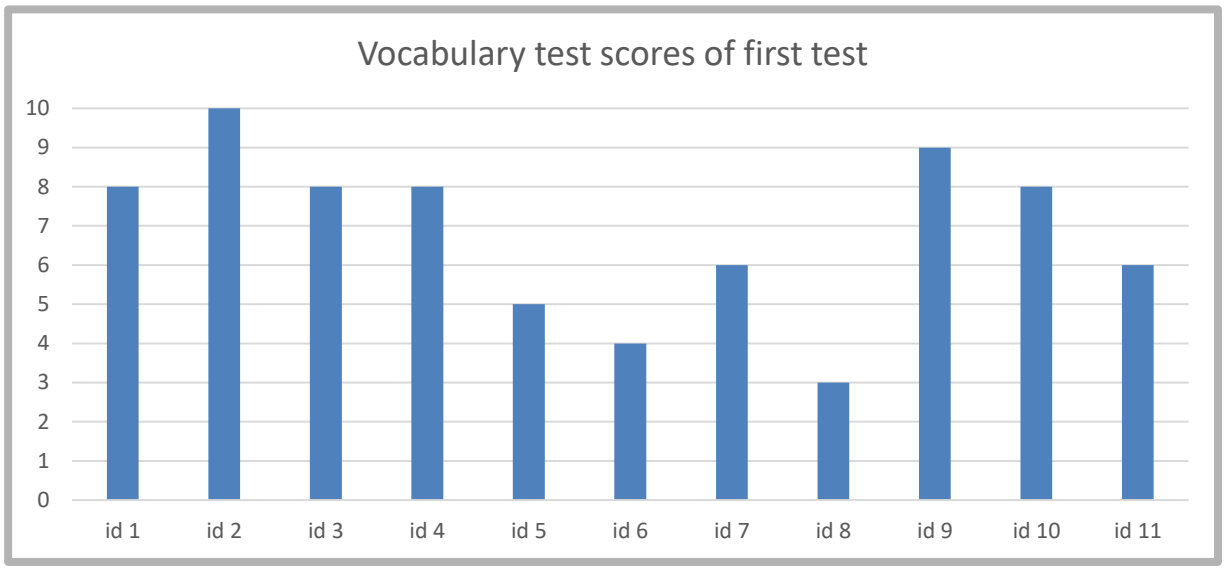


Figure 17. The students' vocabulary test scores from the first vocabulary test. A standard deviation of 2.182575626.

A single student received a "perfect" score of ten. A score of eight was received by four students in the group, making it the most frequent total score of the first vocabulary test. The second most frequent total score was six, given to two of the students. No student ended up with a test score lower than three. The average score for whole class, for this first test, ended up being 6.818~, with a standard deviation of 2.1825~.

4.3.3 Results of the second CALL vocabulary test

Variation among test scores is also apparent in the results of the second CALL vocabulary test:

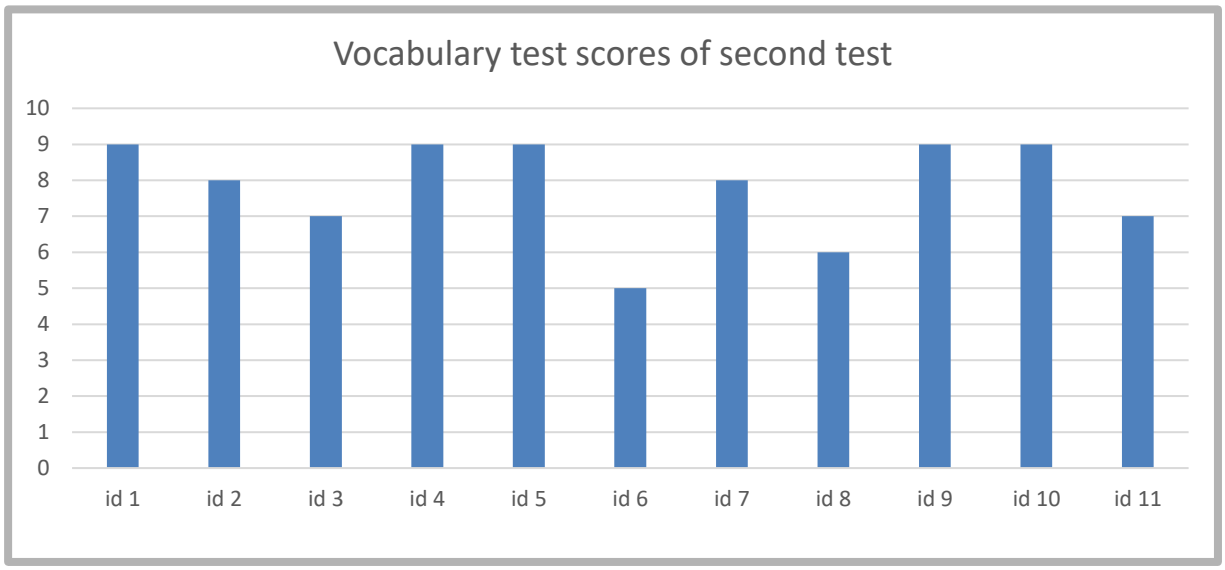


Figure 18. The students' vocabulary test scores from the second vocabulary test. A standard deviation of 1.401298099.

Not a single student received a "perfect" score, however, five students received a test score of nine. In other words, nearly half of the students taking the second vocabulary test managed a near perfect score. There were no scores lower than five. The average score for whole class, this time around, was 7.818~ with a standard deviation of 1.40129~.

4.3.4 Post-study Fall 2022 Average

After the study, the students returned to using more traditional vocabulary learning methods. By Fall 2022, the students had advanced to the sixth grade. Surprisingly, the students' average vocabulary test score had further increased, from the already relatively high 9.636~ to a an even higher score of 9.727~ with a standard deviation of 0.4670~. Most of the students, now, had an average vocabulary test score of ten.

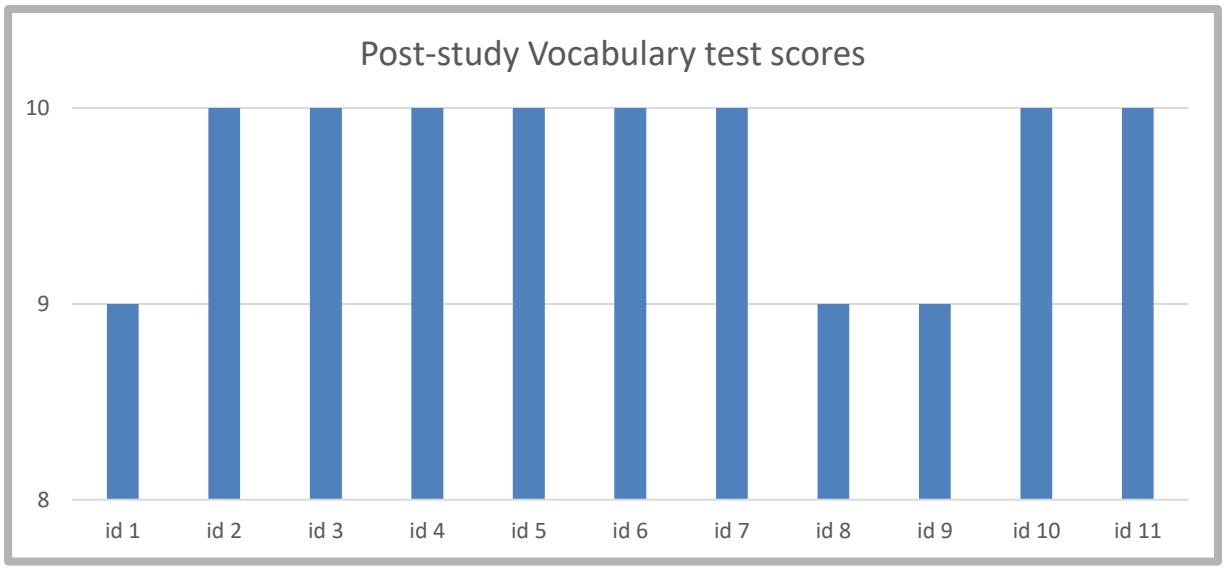


Figure 19. The students' post-study Fall 202 vocabulary test score average. A standard deviation of 0.467099366.

4.3.5 Difference between 1st test and pre-study

The difference in test scores between the pre-study average and the first vocabulary test are very apparent. The students' earlier averages only ranged from nine to ten out a possible ten. However, when looking at the scores of the first vocabulary test conducted during the study, we can see that only two students out of 11 scored either a nine or a ten. Scores of eight, six, five, four or three were nowhere to be seen in the earlier pre-study average scores. However, many lower scores were present in the first test.

A single student, out of 11, scored a ten out of ten. Another student managed to score a nine in this test.

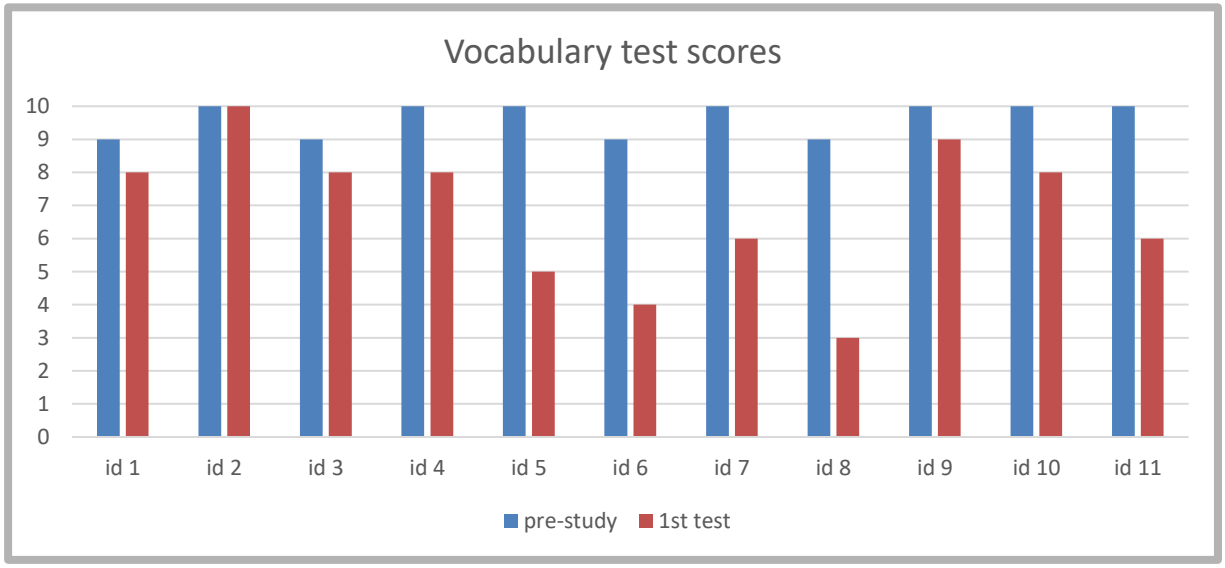


Figure 20. The difference between the pre-study scores and scores of the first test. A standard deviation of 2.54058

The scores received by the students in the first test were lower, overall. For example, three students scored half or less than half of the total points. The total difference between the students previous 2021 average and the average of this first vocabulary test was 2.818~ with a standard deviation of 2,54~. In other words, there was a decrease in the average of almost three points.

4.3.6 Difference between 1st and 2nd test

The difference between the results of the first vocabulary test and the second is also apparent. In total, five students managed to score nine points out of a possible ten, quintupling the number of nines. While, no student received a ten, this time. The lowest score received by a student changed from a three to a five, an increase of two points.

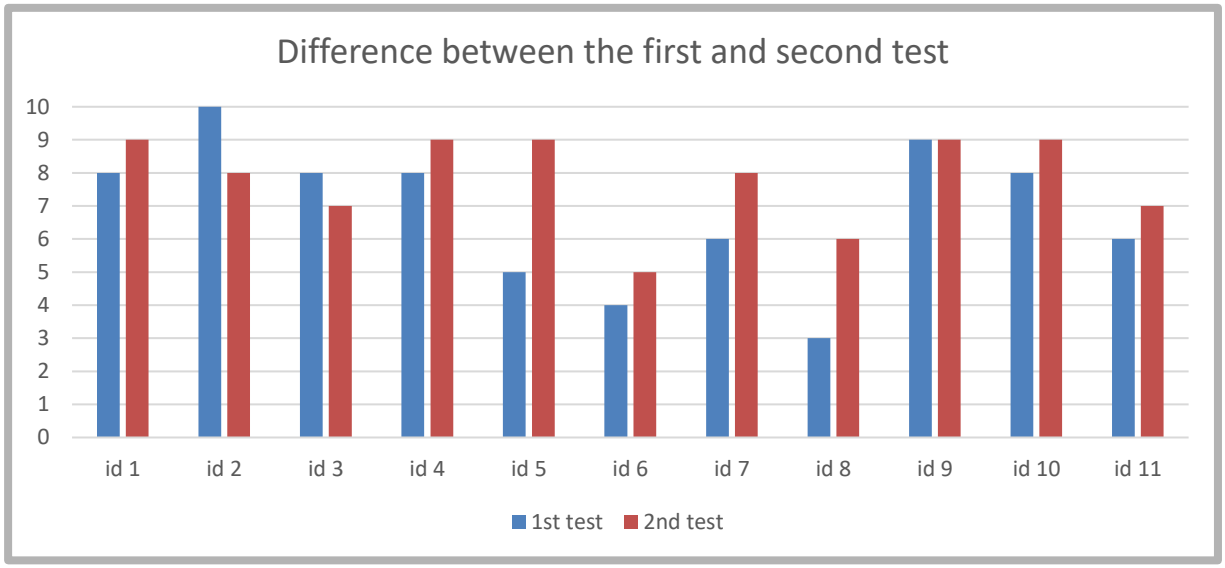


Figure 21. The difference between the scores of the first and second test. With a standard deviation of 1.67332

Overall, the students scored numerically higher in the second vocabulary test. nine out of 11 students saw an increase in their test score, when compared to the first vocabulary test. One student received the same score as before and another student received a lower score. The group average score increased by 1.0 with a standard deviation of 1.673~.

4.3.7 Difference between pre-study and 2nd test

The score gap between the pre-study average and the second vocabulary test score is smaller than the one between the pre-study average and the scores of the first test. An average difference of 1.727~ with a standard deviation of 1.1037~.

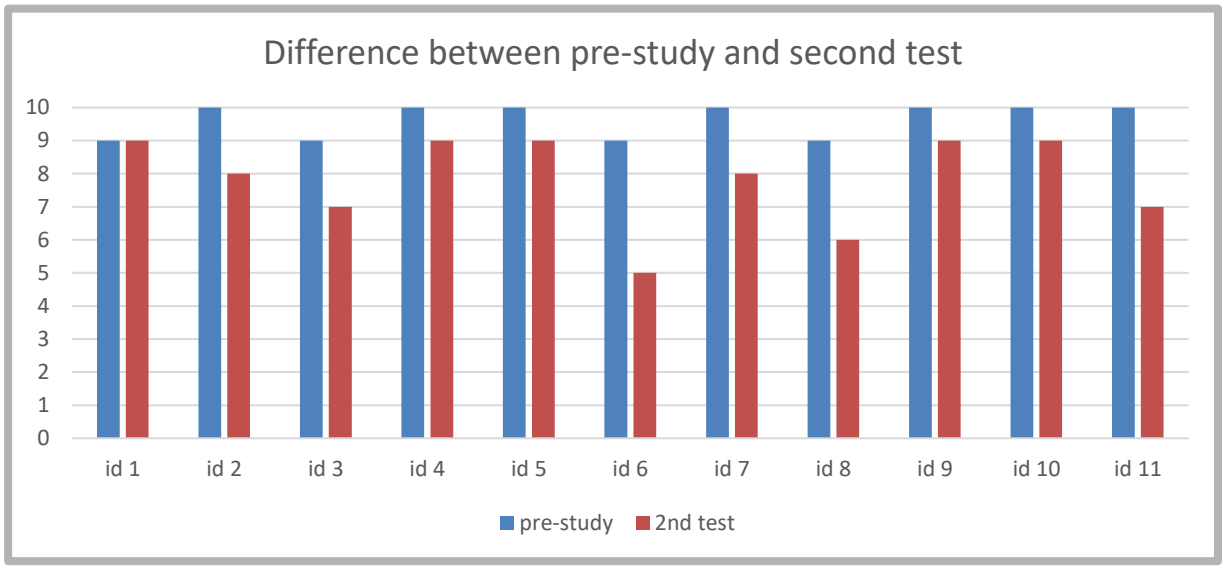


Figure 22. The difference between the pre-study scores and scores of the second test. With a standard deviation of 1.103713

4.3.8 Post-study score averages, Fall of 2022

The participants' Fall 2022 average score 9.727~ is quite close to their earlier 2021 average of 9.636~, the difference only being 0.0909 points. The students' vocabulary test scores ended up "recovering" to their former state. Let's also look at the score sum totals. The sum of all the 2021 average points wound up being 106 points and the sum of the Fall 2022 average points was 107 points. The sum of points for the first vocabulary test was 75 (approximately 70% of 106 points) points and 86 (approximately 81% of 106 points) for the second test.

4.3.9 Student averages based on all four scores and the two vocabulary tests

If give each student an average score based on all of four of their test scores, the classroom average based on the students' averages on all four scores ends up being 8.5 with a standard deviation of 0.901388. The class average, for the average of the two vocabulary tests, conducted during this study was 7.31818~ with a standard deviation of 1.632~. The average difference between these two averages was 1.1818~ with a standard deviation of 0.775~.

Table 3. id refers to a single student.

id	Average of all four scores (pre-study, 1st test, 2nd test, post-study)	Average of the two vocabulary tests	Point difference between the two averages
1	8,75	8,5	0,25
2	9,5	9	0,5
3	8,5	7,5	1
4	9,25	8,5	0,75
5	8,5	7	1,5
6	7	4,5	2,5
7	8,5	7	1,5
8	6,75	4,5	2,25
9	9,25	9	0,25
10	9,25	8,5	0,75
11	8,25	6,5	1,75

4.3.10 Histograms of individual scoring patterns

The final score histograms, depicting the progression of student averages, ended up being quite varying in shape. Out of the 11 histograms, only two were completely identical, pointwise. In these two histograms, the score progression was: ten, eight, nine, ten. The remaining nine histograms all differed in their scores, when compared to the others. The effects of the CALL software on the students score progression were extremely varied, if nothing else.

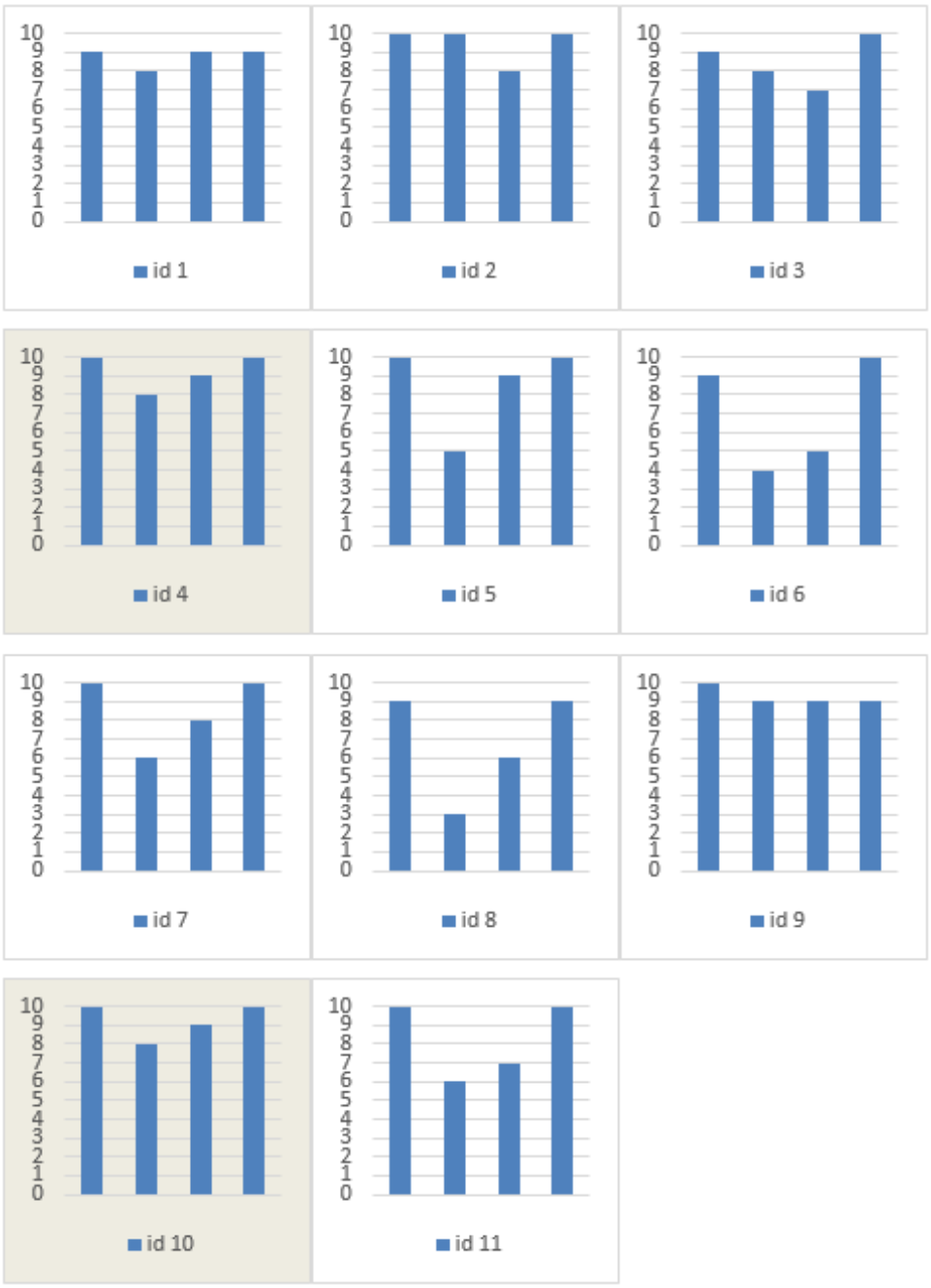


Figure 23. All of the individual students score histograms. Ids 4 and 10 share all of the same values. The rest of the histograms vary from one another.

4.4 Results in groups

Here, we look at the students' test scores, from the perspective of the feedback form. Due to the lack of statistical analysis, these perceived differences cannot be judged to be statistically valid. In other words, the following are hypothesizing observations.

4.4.1 Based on weekly use

Four out of the 11 students in the classroom, reportedly used computers for more than five hours per week, during their free time. It would stand to reason, that, perhaps, students with prior familiarity with computers could fare better with the CALL tool. The average scores of these frequents computer users and the remaining, more "infrequent" computer users were almost identical, in the first vocabulary test. The average for the frequent group was 6.75 points and for the infrequent group it was 6.8 points.

The score average of the frequents computer users for the second vocabulary test was 8.0 and for the group of infrequent computer users it was 7.857~. Again, these averages seem very close to one another. These groups' average scores ended up being similar, despite differences in prior computer habits. Next, let's look at and compare the students' scores, from the perspective of their answers to the final feedback form.

4.4.2 Based on answers to questions one and two

Almost every single student (ten out of 11) answered a for both questions one and two. There does not appear to be a point in analyzing test score differences based on these two questions. Since, this would be akin to comparing the scores of one, randomly selected, student's scores to the scores of the other ten.

4.4.3 Based on answers to question three

Participants who answered c, felt that their ability to retain new words was unaffected by the CALL tool. This group had an almost identical pre-study average of 9.6 as the "affected group", who answered either a or b, which averaged 9.666~ points. One student was unable to give an answer to this question.

The scores of the first vocabulary test also contained this similarity. The affected group had an average of 6.8333~ points, while the reportedly unaffected group scored an average of 6.8 points. The gap between the score averages had decreased by an additional 0.033~ points.

When looking at the results of the second vocabulary test, however, this seemingly small gap between the averages increases. Both groups scored higher, on average, in the second test. The affected groups' average score was 7.5 and the unaffected group managing an average of 8.2 points. This leaves us with an overall difference of 0.7 points. Overall, the group of unaffected students, who felt that their learning had not been affected by the CALL tool, scored higher in the two vocabulary tests.

4.4.4 Based on answers to question four

Now, we move on to question number four. Were there differences between students' scores, based on which sensory aspect of the program they paid most attention to? Almost half of the students, five in total, reported that they paid the most attention to the sound (c) produced by the CALL tool. The remaining six students picked something other than c, as their answer. Two of these students were not able to provide a definitive answer, to this question. Let's compare the vocabulary test scored of group c aka. the sound group to the rest of the class or "the other group".

For the first vocabulary test, the sound group received an average score of 7.2 points. Meanwhile, the other group scored an average of 6.5 points. However, after the second test the score gap between these groups decreased substantially. This time around, the sound group received an average score of 7.8 points, while the other group received a barely higher score of 7.833~.

Interestingly, the three students from the other group, who reported to have paid the most attention to the image aspect of the CALL tool (a), all scored nine points of out ten, in the second vocabulary test. All three students also had a pre-study average of ten and maintained this average of ten, in the Fall of 2022.

4.4.5 Based on answers to question five

According to the feedback, five of the 11 students reported that they paid the least amount of attention to the text (b) presented by the CALL tool. Initially, this group received an average score of 6.6 points, from the first vocabulary test. However, they ended up receiving an average of 8.4 points from the second test, a noticeable increase of almost two points. In fact, four out of the five students, in the group scored nine points out of ten, in the second vocabulary test.

Another group of three students that reportedly paid the least attention to the images (a) scored an average of 8.0 points, in both vocabulary tests. So, no change in the average, between tests, aside from variation in individual scores. The remaining group of three students, who reportedly paid the least attention to sound (c), received the lowest average scores of all the groups, in both tests. In the first test, their average score was 6.0 and in the second test it was 6.666~.

4.4.6 Based on answers to question six

In the last question of the feedback forum, students were asked which method they would prefer to use for learning, in the future (textbook or computer). One student reported that they would be fine with either method. Three students said that they would prefer to keep learning with a traditional textbook. This group of three had a pre-study average of ten. However, their average score in the first vocabulary test average was 5.666~, much lower than previously. All three students scored higher in the second test. Their average score, in the second test was 8.0. The rest of the class, seven students in total, indicated that, in the future, they would prefer the use of computers over traditional textbooks. This group's average score in the first vocabulary test was 7.0 and 7.547~ in the second test. Interestingly, the one student who was fine with either method of learning was also the only student that received the same score (nine points) from both vocabulary tests. Additionally, this student was also the only student who had a Fall 2022 average lower than their pre-study 2021 average. The other ten students either scored higher in the second test or they received a lower score.

4.4.7 Care to comment?

Lastly, the feedback form contained space for an additional comment. All the students who provided an extra comment had, in the Fall of 2022, an average vocabulary test score of ten. Out of the six students who chose not to elaborate further, in the form of a comment, three had an average test score of nine, in the Fall of 2022. Regardless, not much can be extracted from this information.

5 Discussion

5.1 Participants' reception of the CALL tool

Was the CALL tool successfully adapted into the German language learning classroom? The majority of the students, participating in this study, found the CALL tool to be easy to use. All students in the class possessed the ability to competently use a multimedia touchscreen device. Evidently, according to the pre-study survey data gathered, some of the students were already quite familiar with the use of computers during their free time.

Before the study, the students of this class were accustomed to using physical textbooks to learn German words. What about after the study? Did students become interested in the use of CALL? At least according to the feedback, the majority of the students reported that they would like to use multimedia programs, similar to the CALL tool used in this study, in the future. This would seem to indicate, that the students enjoyed using multimodal materials and tools to learn vocabulary items. Evidence for learners' desire to engage with multimodality has been found in some earlier studies (Cárcamo, Melisa Millaray Acuña et al. 2016). Furthermore, this would also seem to strengthen the claim, that there exists a need for more CALL, in schools, among young people.

Regardless of the students' computer use habits, they were all able to learn to use the program for vocabulary training, rather quickly, with little effort. In this sense, the students took part in successful human-computer-interactions with the end goal of conjectural learning. Computer tools, like the touchscreen, have been a part of modern society for many years now. Therefore, it is not at all surprising, that even elementary school students are capable of competently using them in a classroom scenario. The students' successful use of the CALL tool during the study, would also seem to suggest that even young people are very much able to adapt to CALL methodology.

The teacher also found the use of a multimodal CALL tool to be beneficial and informative. They perceived the CALL tool as a good additional learning activity for the group of students. According to the teacher, the students' learning motivation also increased, as a results of interacting with the CALL tool. When we consider the fact that, nowadays, many children have access to advanced and complex software and video games, studied in fields such as DBGL, the fact that students, in 2022, had a positive experience interacting with the comparatively basic CALL computer program, is somewhat surprising. The CALL tool did

not have to utilize expensive state of the art multimodal human computer interaction technology, like virtual reality, to be seen as engaging and motivational. The tool also did not include any videogame-like narration. This positive attitude towards CALL has also be witnessed in some earlier studies, such as the 2010 study by Liou.

The positive feedback provided by the students also seems to hint that multimedia CALL tools do, in fact, possess a "*fun factor*", as suggested by Warschauer and Healey (1998, 83). Whether or not this is simply the result of the CALL tool seen as "*fun*" compared established routine or students just preferring computer use remains to be seen or some other factor. According to the feedback over half the students reported that they would prefer the use of computers over textbooks, in language learning. It is entirely possible that the students' opinion would change, if the CALL tool became a staple of their everyday learning routine. Nevertheless, in this case study, this opensource CALL tool successfully functioned in its role as a language learning tool, without interrupting lesson flow.

5.2 Vocabulary test scores with CALL

Can we say that the students' scores were affected by the use of the CALL tool? Were they able to use the CALL tool for the purpose German vocabulary learning? Glancing at the vocabulary test scores of the students, we can see that by using the CALL tool, the students were able to retain German vocabulary. Before the study, all of the students had scored very high marks in previous German vocabulary tests. Presumably, this is partly due to the fact that German is an optional subject and that the students were already motivated to learn German vocabulary. Additionally, it is true that elementary school vocabulary tests, at least in Finland, are not designed to be exceedingly difficult.

In preparation for the first test, the students had three lessons, during which they had access to the CALL tool. From the first vocabulary test conducted during this study, some students received numerically low scores. While some other students were able to score relatively well. This seems to strengthen the claim that, at least in some cases, a language learner only needs a few exposures to multimodal material, for learning to occur (Bisson, Marie-Josée et al. 2014, 870-873).

A likely reason for some students low scores was the students' unfamiliarity with the newly introduced CALL tool. Since the students had previously been using textbooks provided by the school, some time was spent on adjusting to the new learning method. Additionally,

compared to their previous vocabulary tests, the students had less time to familiarize themselves with the words. This is because, unlike with the textbook, students were only able to interact with the CALL tool in class, for a limited amount of time. In retrospect, each participant had very limited time to interact with the CALL tool. This was mostly due to the availability of only a single personal computer. The participants would have undoubtedly had more time to engage with the CALL tool, had classroom had access to more computers or some multi-server technology with which to use the CALL tool. What about the results of the second vocabulary test?

Eight out 11 students scored numerically higher in the second vocabulary test. There may be several reasons for this. The lowest score received by a student increased by two points, changing from a three to a five. One student increased their previous score by three points and another by four. Overall, the class average score was higher by one whole point. A difference between the first and second test was visible. One possible reason for the increase may have been the individual students own motivation to increase their test score. Since the CALL tool used in this study provided no feedback to the students, the tests graded by the teacher, following the first period of CALL use, may have functioned as good motivation for the students. Additionally, when training for the second vocabulary test, the students had already had adequate time to familiarize themselves with the new classroom learning tool. The use of the touchscreen did require the use of a different set of cognitive and physical skills, than the reading of a textbook. More repetition has been found to enhance people's language learning abilities, by studies such as Boers, Frank et al. (2017) and Chapelle (2003).

The fact that five students received a score as high as nine out of ten, from the second vocabulary exam is very surprising, given the limited access these students had to the CALL tool. Arguably, this can be seen as a return to the pre-study "nine to ten" point range, for these students. In other words, these students demonstrated their fast ability to adapt to the new CALL tool methodology used. Additionally, all of these high scoring students reportedly paid the most attention to either the images or the sounds produced by the CALL tool. On top of this, four out of these five students stated that they had paid the least attention to the CALL tool's text features. It could indicate that the students who paid the most attention to the multimodal aspects enabled by the use of the CALL tool performed well, in the second vocabulary test. Paying less attention to the text aspects of the CALL tool was beneficial to some. However, due to the small number of participants and lack of statistical testing, it is impossible to say whether this had any real effect on the students' scores.

After both of the vocabulary tests had been completed, the students returned to their normal classroom language learning routine. Did the students post-study scores remain on the same level, that they were before the study. By observing the students Fall 2022 vocabulary test score averages, we can plainly see that the group of students continued to score numerically high, in vocabulary tests. As was the case before this study, the students all had a vocabulary test average score of either nine or ten.

In summary, many of the students' scored numerically lower than their pre-study average, initially, during the first vocabulary test. However, their scores increased in the second test, with some students receiving quite high points. The participants' post-study 2022 averages mirrored their pre-study averages. The temporary switch to the CALL tool had no negative long-term effects on the students' ability to perform well in German vocabulary tests. From this, we can suggest that testing and implementing creative and new CALL tools for vocabulary learning can be a valid method of motivating students, in the everyday learning situation. Additionally, it was shown, that even with a limited time window, some learning occurred in the class with the help of the CALL tool.

5.3 Sense of retention with CALL

When we look at the data from the feedback, we can see that students own sense of vocabulary retention with the use of the CALL tool was somewhat accurate. On average, the class scores lower in the two vocabulary tests conducted during the study than they did before or after the study. nine out of the 11 students did not feel that vocabulary retention was easier with the use of the CALL tool. Only two students stated that they felt that retaining words was easier this way.

It is interesting to then contrast this with the fact that the vast majority of the students stated that they would like to continue to use such CALL tools in the future. The students would like to use the tool, even though many of them felt that it did not improve their ability to retain words, compared to their normal classroom methodology. Perhaps, this is further evidence of a CALL tools effectiveness, as an energizing motivational tool in the classroom. By implementing such a tool into the lesson plan, educators could bring more variety into an, already, established teaching routine.

5.4 Modality differences?

What about the students' perception of the multimodal aspects of the CALL tool? Judging by the feedback, we can see that the students differed greatly in what aspect of the CALL tool they paid the most attention to. Some students seemed to struggle with choosing exactly which of the three they paid the most attention to. Out of the three options, only one student reported to have paid the most attention to the text of the program. The rest of the class picked a different answer. It could be that, the pictures and sound modes provided by the CALL tool were seen as more novel, by the students. Lou's 2008 study pointed out novelty, as one of the upsides of using a cellphone as a learning. The novelty factor may have contributed to the increased attention paid to images and sounds, in favor of text.

The students had no knowledge as to how long the CALL tool would be available to them, so perhaps they saw fit to focus on the aspects most exclusive to it. According to the answer to question five, almost half of the students paid the least attention to the text presented by the program. This would seem to hint to fact that the text aspect of the program was its least interesting aspect.

Unfortunately, due to the small number of participants taking in this study, making any strong conclusions on the CALL tools effect on vocabulary test scores, based on the students' multimodal preferences, is an impossibility. Due to the small numbers of German students available for this study and the chosen methodology, the number of participants found for this study was small. Only so much data could be extracted from the vocabulary test scores of 11 German language learners, who only had access to a single computer equipped with a CALL tool used for conjectural learning. More conclusive data could have likely been found, had this CALL study been conducted in a different setting or utilized a computer tutor.

5.5 Comparison to earlier studies

Yes, the number of participants taking part in this study was, unfortunately, small. Therefore, making any sort of in-depth, statistical conclusions based on the available data would be ill-advised. Let us contrast this study's setting with earlier CALL studies. First, in this study elementary school students were tasked to prepare for a German language vocabulary test. In Finnish schools, German is an optional subject, meaning class sizes are smaller than usual, by default. Additionally, not all schools offer German lessons. Compare this factor to many of the earlier studies mentioned previously. In most of the studies discussed previously,

participants were dealt with a language which was widely taught in the country where the research took place. For example, Lu's (2008) study focused on Taiwanese English learners. Lu-Fang's (2010) study also dealt with English learners. Other studies, like that of Yamada et al. (2019) have also selected to work with languages with high learner numbers, in the country where the studies have taken place. As a global lingua franca, English is a compulsory subject in most countries, Finland and Taiwan included.

Finding participants for this case study would have probably been easier if the language of choice would have been more widely taught. Unfortunately, this is not the case for a smaller foreign language like German. Problems seem to arise, when attempting to conduct a CALL study, when the number of available language learners is small. In this study, involving 11 Finnish elementary school students studying German, the lack of participants led to an insufficient amount of data. The decision to have German as the language utilized by the CALL tool used in this study led to a decrease in the amount of data available. It should be noted, however, that conducting this study with the English language students would have led to some necessary modification of the CALL tool. For one, English students would not need to be presented with the plural form of the word, as forming the plural form of a noun in English is comparatively a much less complicated process than it is in German. Constructing and analyzing a CALL tool meant for a smaller language should remain a possibility despite lower learner numbers.

We can state that only using a single personal computer, here, to gather information about the effectiveness of a CALL tool on the vocabulary test scores of a small classroom did not produce enough quality data, for us to draw any conclusions on multimodal preferences in CALL. A more successful version of this study could have taken advantage of some type of multi-server technology, like Kahoot, for example. Opting to use of a more widely available technology, like a cellphone, could have also been beneficial in terms of increased data. Lu's 2008 study was able to gather more data, partly thanks to the widespread availability of cellphones. However, building a multimodal CALL tool to be used for vocabulary learning on a cellphone operating system could prove to be difficult, due to factors such as hardware limitations and memory.

Another way of tackling this problem regarding data, could have been to opt for the use of a computer tutor instead of a computer tool. As explained by Levy (1997) a computer tool, is simply a tool used by the user. Unlike the computer tutor, it does not evaluate user data. If the

CALL program, used in this study, would have been scripted, to function in the role of a computer tutor, it is possible that better data could have been extracted, regardless of the small sample size. This hypothetical CALL tutor program could have been programmed to store more user data and other metrics, such as repetition. It also could have been used to perform more vocabulary testing in a more effective matter, compared to the traditional pen and paper tests conducted here. This way, information could have also been gathered on things such as repetition and time spent on a single vocabulary item by a user. Based on the results of this study and hindsight, I propose that the simplistic CALL tool, only available on an offline single pc computer, is not an effective choice, for a study of this kind. In this case study, it did produce semi-analyzable data, however, due to the ultimately low amount of it gathered, many of the findings are doomed to be inconclusive.

However, this study does indicate that many young students are open to the use of CALL tools in language learning. On top of this, students are more than able to adapt to the use of a CALL tool. It is encouraging to see that even a program made using opensource software, with little to no budget, can function as motivational learning tools. Even if the use of the program did not produce “better” learning results, on par with previous methods, reception to it was mostly positive. In fact, the fact that students were able to retain vocabulary after only a limited amount of time with a CALL tool, is encouraging, to say the least. In addition, the data shows that trying out a new CALL tool does not have any temporary effects on students’ ability to switch back to previous ways learning. At the very least, a CALL tool like the one used here could serve as an extra learning tool, which could enhance the already established workflow of the classroom. It stands to reason, that such a multimodal CALL tool could be adapted by other subjects as well, such as biology.

The results of this study would also seem to reflect positively on more traditional learning methods. In this case, I’m mainly referring to language learning textbooks. In some sense, the students’ earlier vocabulary test scores would seem to imply that the use of a traditional textbook, as a learning tool, had been effective. Outside of the study, the students had constant access to their language learning books. In contrast, the CALL tool was only available to them in school. Comparatively, the textbook proved to be superior to the CALL tool, in terms of overall availability. This was due in no small part, to the technical limitations of this study’s setting. The apparent effectiveness of the textbooks does raise a question, however. If elementary students are capable of effective learning using non-electronic means, is the use of electronic methods, such as CALL tools, even necessary? While it is true, that customizing

and modifying a CALL tool is much easier and provides its own advantages, a well-designed and prepared textbook seems to perform its job as a language learning tool just as well. However, during this case study, it was found that the CALL tool did, in some ways, benefit the vocabulary learning process.

6 Conclusions

In this case study, I set out to observe how a multimodal, open-source CALL tool was received by a small, Finnish class of German students. I also aimed to see what the resulting vocabulary test scores for the students would be like, after they had used the CALL tool to learn new vocabulary. I looked at different test scores produced by students of the class. These included pre-study vocabulary test averages, scores from two specially prepared vocabulary tests and data from their later 2022 vocabulary test scores. Additionally, I also gathered and took into consideration feedback provided by both the students and their teacher.

So, can we say that students were able to use the CALL tool to learn vocabulary? The results of this case study showed that, at least with this group of students, there existed a capability and motivation to use such a CALL tool for vocabulary learning. Most of the students had a positive experience using the program. The majority of the students did not receive high test scores from the first test. A few students received a relatively high score. This is surprising, given the fact that the students did not have a long time to interact with the CALL tool. Yes, some students' scores were lower than others. The students scored higher in the second vocabulary test. This may have been the result of more experience with the CALL tool or repetition, for example. The lower test scores did not lead to any permanent decreases, as students' already quite high vocabulary test score averages continued, during the following fall semester.

It appears to be the case, that the participants of this case study were perfectly capable of performing well in German vocabulary tests, without the use of a CALL tool. For this class, using traditional textbooks seems to have served them exceedingly well, thus far. This is not to suggest, however, that no learning took place with the help of the CALL tool. The reception to the CALL tool was also overwhelmingly positive, with most students stating, that they would like to use such multimodal tools in the future. Some even stated that they would prefer the use of the CALL tool over their textbook. Perhaps, students were more motivated by the novelty of using a CALL tool in the classroom. Other students may have needed more time to become better acclimated with the newly introduced learning tool. The classroom teacher also stated that they benefitted from the use CALL tool in the classroom. The teacher agreed, that the students were motivated by the CALL tool. Technical problems, which could have interrupted the lesson flow were also minimal.

The most glaring problem related to the conduction of this study was the very small amount of data produced by the methodology of choice. I set out to also analyze the effects of students' multimodal preferences on their vocabulary test scores. Unfortunately, due to the size of the class and the use of only a single touchscreen computer, the amount of data gathered during this study ended up being small. Gathering more data from this small class of German students would have required even more time to be spent with the class. Organizing time for lessons, where students can prepare for a vocabulary test can take weeks. On top of this, there is no guarantee that all of the students would remain as students of the class for such a prolonged period of time. What was found in this case study is that a small period of human-computer-interaction with a CALL tool can serve to motivate a classroom of language students. However, opting for the use of a CALL tutor could have ended up providing us with more, better analyzable language learner data. For example, by implementing the vocabulary test itself into a CALL tutor program, we could increase the number of tests available.

Many other CALL studies have relied on more flexible and widely available technologies to conduct their studies. In the future, a study like this can, obviously, be improved upon. The number of participants taking part in this study was, in retrospect, quite low. Increasing the number of participants could provide a better, more varying collection of data. The most obvious solution for this would be to conduct a similar study in a larger school, perhaps a school in a more metropolitan area. On the other hand, this could lead to neglecting the conduction of such studies in smaller schools. The study of the use of CALL, in different learning environments can provide us with useful insights.

On the more practical, technical side of things, a future study would benefit from having access to more "customizable" and open computer equipment. Originally, the plan for this study was that the CALL tool would be installed on multiple computers in the classroom. However, this turned out to be an impossibility, due to some bureaucratic and technical restrictions and, on the other hand, my own shortcomings when it comes to aspects of Python programming knowledge. Essentially, these issues can be boiled down to two major ones. First, the computers that the school had did not allow for the installation of any custom programs, because of security concerns. Obviously, the security of multiple, expensive computers is a valid and understandable concern for schools. This lack of available computers could of course have been circumvented by making the CALL tool accessible through the Internet. However, making this a workable solution with Python was impractical in this case. It is true, that there exist online services that allow the execution of Python scripts through a

web browser. Services like Google's *Google Colab* make it possible to share and execute Python scripts in a browser environment. Unfortunately, though, as far as I am aware none of these services offer support for the use of GUI modules, like *Tkinter*. In other words, these Python programs can be executed, in such browser environments, in text form only. As the CALL tool used in this study had a heavy emphasis on visuals and other such features, using a web environment for this would be impossible. The possibility of utilizing multimodality in this case would be impossible.

Naturally, this problem could be solved by simply maintaining a dedicated server, running a virtual machine, with which the Python program could be used with the combination of some *Flask* or *Javascript*, for example. However, the accessibility and stability of such a server would be more than questionable and most likely would cause multiple issues for a study like this. For example, how would you guarantee that all students had similar access to the tool? Could browser incompatibilities lead to some users not being able to use the program at all, through no fault of their own? And even if all technical aspects of such an online tool were sound, what would guarantee that all the students had access to an appropriate place to study using a CALL tool or had enough free time, outside of school? What could even guarantee that all the students engaged with the software to begin with? All of this is to say, as was the case with this study, that future studies like this should exhaust available options, when it comes to the practical conduction of a CALL tool related study.

In retrospect, perhaps the most obvious solution to the lack of data, would have been the use of multi-server technologies, however, as discussed here, this would have most likely forced a change in the methodology. The CALL tool used in this study was built using opensource software. However, perhaps the most efficient way of conducting CALL studies like this would be to rely on the use of readily available, pre-existing language learning resources, like Duolingo or Kahoot, for example. Yes, when using these proprietary online services, the possibility of customization is rather minimal. However, what these programs lack in customization, they more than make up for in access and availability to thousands of users. More language learner data can help us better understand the effects and effectiveness of these computer assisted language learning tools. Unfortunately, in the present day, not even these online language learning tools are without their fair share of problems. For example, if one were to conduct a CALL study based on the user data of millions of people, what would guarantee that all of that data was all human in origin, to begin with and not compiled by bots? Yet another thing for the field of CALL to consider, as the result of technological

progress. In my study, the presence of human language learners was verifiable, thanks to the methodology used. As newer and newer CALL applicable software emerges, we should not forget the important, real-world roles played in the learning process, by the language, the school, the students, and the teachers.

References

Bateson, G., & Daniels, P. (2012). Diversity in technologies. In G. Stockwell (Ed.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 127-146). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139060981.008

Bisson, M.-J., van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2014). The Role of Repeated Exposure to Multimodal Input in Incidental Acquisition of Foreign Language Vocabulary. Language Learning, 64(4), 855–877. https://doi.org/10.1111/lang.12085

Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. Computer Assisted Language Learning, 30(7), 709–725. https://doi.org/10.1080/09588221.2017.1356335

Cárcamo, M. M. A., Cartes, R. A. C., Velásquez, N. E. E., & Larenas, C. H. D. (2016). THE IMPACT OF MULTIMODAL INSTRUCTION ON THE ACQUISITION OF VOCABULARY. Trabalhos Em Lingüística Aplicada, 55(1), 129–154. https://doi.org/10.1590/010318134842170942

Çakmak, F., Namaziandost, E., & Kumar, T. (2021). CALL-Enhanced L2 Vocabulary Learning: Using Spaced Exposure through CALL to Enhance L2 Vocabulary Retention. Education Research International, 2021, 1–8. https://doi.org/10.1155/2021/5848525

Chapelle, C. A. (2003). English Language Learning and Technology: Lectures on Applied Linguistics in the Age of Information and Communication Technology. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/lllt.7

Chen, M., Tseng, W., & Hsiao, T. (2018). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. British Journal of Educational Technology, 49(1), 69–77. https://doi.org/10.1111/bjet.12526

CHUN, D. M., & PLASS, J. L. (1996). Effects of Multimedia Annotations on Vocabulary Acquisition. The Modern Language Journal (Boulder, Colo.), 80(2), 183–198. https://doi.org/10.1111/j.1540-4781.1996.tb01159.x

Cuban, Larry. (2009). Oversold and Underused. Harvard University Press.

Donker, A., & Reitsma, P. (2007). Young children's ability to use a computer mouse. Computers and Education, 48(4), 602–617. https://doi.org/10.1016/j.compedu.2005.05.001

Dunkel, Patricia. (1991). The Effectiveness Research on Computer-Assisted Instruction and Computer-Assisted Language Learning. In Computer Assisted Language Learning.

Elkind, D. (2016). Touchscreens and Young Children: Benefits and Risks. YC Young Children, 71(1), 90–93.

Goldstein, Dana. (2020). Coronavirus Is Shutting Schools. Is America Ready for Virtual Learning? https://www.nytimes.com/2020/03/13/us/virtual-learning-challenges.html

Hirschel, R., & Fritz, E. (2013). Learning vocabulary: CALL program versus vocabulary notebook. System (Linköping), 41(3), 639–653. https://doi.org/10.1016/j.system.2013.07.016

Hubbard, P. (2009). Computer assisted language learning. New York: Routledge.

Hwang, G.-J., & Wu, P.-H. (2012). Advancements and trends in digital game-based learning research: a review of publications in selected journals from 2001 to 2010. British Journal of Educational Technology, 43(1), E6–E10. https://doi.org/10.1111/j.1467-8535.2011.01242.x

Jaimes, A., & Sebe, N. (2007). Multimodal human–computer interaction: A survey. Computer Vision and Image Understanding, 108(1), 116–134. https://doi.org/10.1016/j.cviu.2006.10.019

Johnson, E., & Brine, J. (2012). Diversity in content. In G. Stockwell (Ed.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 90-108). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139060981.006

Lee, J. (2006). History of Computing in Education. International Federation for Information Processing Digital Library; History of Computing in Education. 145. https://doi.org/10.1007/1-4020-8136-7_1

Levy, M. (1997). Computer-assisted language learning : context and conceptualization. Oxford: Oxford UP.

Lin, L.-F. (2010). A video-based CALL program for proficient and less-proficient L2 learners' comprehension ability, incidental vocabulary acquisition. Educational Media International, 47(3), 199–216. https://doi.org/10.1080/09523987.2010.518812

Liou, H.-C. (2012). The roles of Second Life in a college computer-assisted language learning (CALL) course in Taiwan, ROC. Computer Assisted Language Learning, 25(4), 365–382. https://doi.org/10.1080/09588221.2011.597766

Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. Journal of Computer Assisted Learning, 24(6), 515–525. https://doi.org/10.1111/j.1365-2729.2008.00289.x

Magnusson, P., & Godhe, A.-L. (2019). Multimodality in Language Education – Implications for Teaching. Designs for Learning, 11(1), 127–137. https://doi.org/10.16993/dfl.127

Pachler, N. (2001). Electronic reference tools for foreign languages: offline vocabulary programs. Language Learning Journal, 24(1), 24–29. https://doi.org/10.1080/09571730185200181

Perusopetusasetus 852/10 §. (1998) https://finlex.fi/fi/laki/alkup/1998/19980852

Python Foundation. (2021). https://www.python.org/

Python Package Index. (2021) https://www.pypi.org/

Ramezanali, N., Uchihara, T., & Faez, F. (2021). Efficacy of Multimodal Glossing on Second Language Vocabulary Learning: A Meta-analysis. TESOL Quarterly, 55(1), 105–133. https://doi.org/10.1002/tesq.579

Reynolds, E. D., Fuchs, R. W., & Johnson, P. (2021). Game On With Kahoot!: Effects on Vocabulary Learning and Motivation. International Journal of Computer-Assisted Language Learning and Teaching, 11(4), 40–53. https://doi.org/10.4018/IJCALLT.2021100103

Ringelstein Zak. (2020) ”5 Reasons Zoom Schooling Is Detrimental To Children.” *Forbes*. https://www.forbes.com/sites/zakringelstein/2020/09/15/5-reasons-zoom-schooling-is-detrimental-to-children/?sh=27b8c8051223

Sato, T. (2016). Could a Multimodal Dictionary Serve as a Learning Tool? An Examination of the Impact of Technologically Enhanced Visual Glosses on L2 Text Comprehension. EuroCALL Review, 24(2), 3–12. https://doi.org/10.4995/eurocall.2016.5236

Segers, E., & Verhoeven, L. (2003). Effects of vocabulary training by computer in kindergarten: Vocabulary training in kindergarten. Journal of Computer Assisted Learning, 19(4), 557–566. https://doi.org/10.1046/j.0266-4909.2003.00058.x

Segers, P. C. ., & Verhoeven, L. T. . (2002). Multimedia support of early literacy learning. Computers and Education, 39(3), 207–222. https://doi.org/10.1016/S0360-1315(02)00034-9

Thomas, M., Reinders, H., & Warschauer, M. (2013). Contemporary computer-assisted language learning. London ;: Bloomsbury Academic.

VUNGTHONG, S., DJONOV, E., & TORR, J. (2017). Images as a Resource for Supporting Vocabulary Learning: A Multimodal Analysis of Thai EFL Tablet Apps for Primary School Children. TESOL Quarterly, 51(1), 32–58. https://doi.org/10.1002/tesq.274

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. Language Teaching, 31(2), 57-71. https://doi.org/10.1017/S0261444800012970

Yamada, T., Sakai, T. & Bushnell, C. (2019). 6. Rakugo CALL Program for Japanese Language Learning: Its Development and Possibilities for Implementation. In E. Zimmerman & A. McMeekin (Ed.), *Technology-Supported Learning In and Out of the Japanese Language Classroom: Advances in Pedagogy, Teaching and Research* (pp. 149-170). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781788923514-009

Yeh, Y., & Wang, C.-W. (2013). Effects of Multimedia Vocabulary Annotations and Learning Styles on Vocabulary Learning. CALICO Journal, 21(1), 131–144. https://doi.org/10.1558/cj.v21i1.131-144

Yoshii, M., & Flaitz, J. (2003). Second language incidental vocabulary retention: The effect of text and picture annotation types. CALICO Journal, 20(1), 33–58. https://doi.org/10.1558/cj.v20i1.33-58

Appendices

Appendix 1 Suomenkielinen lyhennelmä

Johdanto

Viime vuosikymmenten aikana, tietokoneiden suorituskyky on kehittynyt huimasti. Työkaluna, tietokoneen yhtenä pääkäyttötarkoituksena on edelleen tiedon jakaminen käyttäjälle. Tämän piirteen tietokone jakaa perinteisten fyysisen kirjan kanssa. Tietokonetaitojen opettaminen aloitettiin Amerikassa jo 1960-luvulla (Lee, 2006, 9). Entäpä miten onnistuu tietokoneiden käyttäminen opetuksessa? Monet meistä ovat varmasti elämänsä aikana käyttäneen, vähintään, yhtä ns. opetussovellusta. Internetin kautta on saatavilla useita kieltenoppimiseen käytettäviä ohjelmia. Esimerkiksi *Duolingo* ja *Kahoot* ovat molemmat suosittuja kieltenoppimisohjelmia. Tietokoneiden historian aikana on jo ehditty kehittää useita opetustarkoituksiin suunniteltuja ohjelmia. Bateson ja Daniels (2012, p. 127-138) erottelevat tietotekniikan eri teknologiat neljään luokkaan: monipalvelinteknologiat, yhden palvelimen teknologiat, yksittäisen kotitietokoneen teknologiat ja mobiiliteknologiat. CALL eli *Computer Assisted Language Learning* (suom. Tietokoneavusteinen kielen oppiminen) on monialainen tieteenala, joka tutkii tietokoneiden käyttöä kielenopetusvälineenä (Hubbard, 2009, p.1). Erilaisia CALL-ohjelmia on mahdollista luoda edellä mainittujen kategorioiden laitteille. Digitaalisten työkalujen käyttö vaikuttaa hyvältä ratkaisulta eri opetusmateriaaleihin liittyviin ongelmiin. Esimerkiksi COVID-19:n pandemian aikana monen koululaisen oppitunnit tapahtuivat täysin ZOOM-sovelluksen välityksellä. Erilaisilla opetustyökaluilla on havaittu olevan erilaisia hyviä ja huonoja puolia. On tärkeää tutkia erilaisten digitaalisten opetustyökalujen käyttöä opetustilanteessa, jos niitä halutaan hyödyntää jatkossa. Tämän tutkimuksen pääkysymyksenä oli: ***Kuinka lukumäärältään pienen, suomenkielisen luokan oppilaat, jotka ovat tottuneet opettelemaan kirjan avulla, suhtautuvat yksinkertaisen, multimodaalisen CALL työkalun käyttöön Saksan sanaston opettelussa ja onko työkalun käytöllä jotain vaikutusta heidän oppimiseensa?***

Kyseinen sovellus oli oppilaiden saatavilla, luokkahuoneeseen sijoitetulla, tietokoneella. Sovellus ei ollut käytettävissä Internetin kautta ja se oli koottu vapaan lähdekoodin ohjelmistojen avulla, toisin sanoen ”ilmaiseksi”. Oppilaat osallistuivat tutkimuksen aikana kahteen sanakokeeseen, joita varten he saivat harjoitella Saksan sanastoa CALL-työkalun avulla. Tutkimuksen aikana, pistetiedot näistä sanakokeista kerättiin talteen. Tämän lisäksi

keräsin tieto oppilaiden menestyksestä ”normaaleissa” sanakokeissa ennen tutkimusperiodia ja sen jälkeen. Tutkimuksen lopuksi oppilaat ja opettaja täyttivät palautelomakkeen, jossa heiltä kysyttiin CALL-työkalun liittyviä kysymyksiä.

Tausta ja teoria

Tietokoneavusteinen kielten oppiminen

CALL eli Computer Assisted Language Learning” (**suom**. *Tietokoneavusteinen kielten oppiminen*) on tutkimusala, joka tutkii tietokoneiden käyttöä kielenoppimisessa. Viime vuosina, ajatus tietokoneiden käytöstä kouluopetuksessa on kasvattanut suosiotaan. Michael Levyn mukaan CALL sai, käytännössä, alkunsa vuonna 1960 aloitetun PLATO-projektin ansiosta (Levy 1997, 15). Tutkimusalana CALL on jatkanut kehitystään nykypäivään asti, erilaisten teknologisten harppausten ja innovatiivisten tutkimusten ansiosta. 2000-luvulla, CALL tutkimusten aiheina ovat olleet muun muassa: sanakirjaohjelmat (Sato 2016), SMS-viestien käyttö kielenopetuksessa (Lu, 2008) ja pelien rooli kielenoppimisessa (Liou 2012). CALL-tutkimuksia on siis tehty hyvin erilaisista näkökulmista ja eri metodein. On esitetty myös näkemyksiä siitä, että nykyaikaisen CALL-tutkimuksen ei tulisi olla liialti teknologisen kehityksen ohjaamaa (Thomas et al. 2013, 29-30).

Ihmisen ja tietokoneen välinen vuorovaikutus

HCI eli Human-computer-interaction (*suom. ihmisen ja tietokoneen välinen vuorovaikutus*) käsite viittaa tutkimusalaan, joka tutkii ihmisen ja tietokoneen välistä vuorovaikutusta. 80-luvun teknologisen kehityksen myötä, tutkijat alkoivat kiinnostumaan enemmän ihmisen roolista tietokoneen käyttäjänä (1997, 69). Tutkimuksessani keskeisenä kiinnostuksen kohteena oli oppilaiden ja niin sanotun CALL-työkalun välinen vuorovaikutus. Hyödynsin CALL-tutkimuksessani Michael Levyn ehdottamaa ”*tuutori-työkalu*”-näkökulmaa (tutor-tool framework) (Levy, 1997, 178-191). Levyn näkemyksen mukaan, tietokoneelle voidaan nähdä kaksi erillistä toimintaroolia. Tutoorin roolissa tietokone ohjaa käyttäjää ja antaa tälle palautetta CALL-ohjelman käytön aikana. Työkalun roolissa tietokonetta tulee taas ”kuulustella”, jotta siltä saadaan tietoa. Tutkimuksessani käytin osallistujien käyttämästä multimodaalisesta CALL-sovelluksesta käsitettä CALL-työkalu (CALL tool). CALL-työkalua käyttäessään, käyttäjä on vuorovaikutustilanteen aktiivisempi osapuoli (1997, 191-192). Toisin sanoen, käyttäjä ohjaa HCI-tilannetta hakemalla relevanttia tietoa itsenäisesti. Sovelluksen käyttöliittymä ei siis ohjaa käyttäjää hänen tehtävässään. Tietokoneen graafisen

käyttöliittymän ansiosta, ihmisillä on nykyään mahdollisuus hyödyntää erilaisia multimedia materiaaleilla varustettuja tietokoneohjelmia.

Multimodaalisuus ja CALL

DCT (dual-coding theory) on teoria, jonka mukaan ihmisen mieli kykenee käsittelemään sanoja (ainakin) kahdella eri tavalla: verbaalisesti ja nonverbaalisesti. Toisin sanoen, esimerkiksi, sanan *kenkä* ihminen ymmärtää sen kirjoitetussa muodossa ja non-verbaalisena mentaalisena representaationa (Sadoski & Paivio, 2013, 28-29). Multimodaalisuus olettaa tämän olevan totta. Multimodaalisuudella viitataan ideaan siitä, että kielellistä merkitystä voidaan kommunikoida monin eri tavoin, ei pelkästään sanoin (Bezemer, Jewitt & Halloran 2016, 3). Esimerkiksi, edellä mainitun sanan *kenkä* merkitys voidaan ilmaista myös kuvan, videon tai jonkun toisen moodin avulla. Sanasta kenkä voitaisiin, vaikkapa, luoda multimodaalinen representaatio liittämällä kirjoitetun sanan *kenkä* viereen kuva kengästä.

Useat CALL- tutkimukset ovat juurikin tutkineet multimodaalisten metodien hyödyntämistä kielenoppimisessa. Tutkimuksissa on tutkittu, muun muassa, osallistujien kokemuksia sovellusten käytöstä ja erilaisilla ohjelmilla tapahtuneen sanasto-oppimisen tehokkuutta. Monet näistä tutkimuksista ovat todenneet multimodaalisuuden tehostavan sanaston oppimista, esimerkiksi: Chun & Plass 1996. Toisaalta, on olemassa muita tutkimuksia, joiden valossa näyttää enemmän siltä, että esimerkiksi oppimiseen käytetyllä ajalla ja toistolla olisi oleellisempi rooli tehokkaassa sanaston omaksumisessa, esimerkiksi Boers, Frank et al. 2017. CALL-tutkimuksien pohjalta on siis olemassa hieman ristiriitaista tietoa multimodaalisuuden hyödyllisyydestä. Tämän tutkimuksen aikana tutkin yhden multimodaalisen sanakirjaohjelman käyttöä CALL-työkaluna, suomalaisessa luokkahuoneessa. Ohjelmaa käytettiin kosketusnäytön avulla ja sen kanssa oppilaat harjoittelivat kahteen, heidän sanastotietämystä testaavaa, sanakokeeseen. Sanakokeiden pisteytys käytti tavanomaista pisteskaalaa nollasta kymmeneen.

Aineisto ja menetelmät

Yhteensä tutkimukseen osallistui 11 suomalaista, viidennen luokan Saksan kielen opiskelijaa. Loin tutkimusta varten python-ohjelmointikielellä yksinkertaisen, graafisen käyttöliittymän omaavan, sanakirjaohjelman. Kyseinen sanakirjaohjelma hyödynsi toiminnassaan multimodaalisuutta. Opeteltavan sanan kirjoitetun muodon lisäksi, oppilaat kuulivat sanan puhuttuna ja näkivät sanan myös sitä vastaavan kuvan muodossa näytöltä. Oppilaan painettua

yhtä näytöllä olleista, Saksan kielisen sanan sisältänyttä nappia, ohjelman graafinen käyttöliittymä toi näytölle kyseisen sanan multimodaalisena representaationa. Jokaiselle oppilaalle neuvottiin kuinka sovellusta tuli käyttää. Ennen molempia sanakokeita, oppilailla oli kolme mahdollisuutta käyttää konetta, sille varatulla työpisteellä, sanaston opiskeluun. CALL-työkalun sisältämät sanat olivat Saksan perussanastoa. Sanat olivat substantiiveja ja niiden aiheet (eläinsanoja, kehonosat, vaatesanoja) eivät eronneet muusta opetusmateriaalista. Molempia sanakokeita varten harjoiteltiin 21 sanaa. Sanakokeissa oppilailta kysyttiin heidän CALL-työkalua käyttäessään kohtaamia Saksan sanoja. Oppilaat olivat jo ennestään osallistuneet sanakokeisiin, ennen koetta. Valinnaisen aineen opiskelijoina oppilaita voitiin pitää jo valmiiksi motivoituneina kielen opiskelijoina.

Tutkimuksen aikana keräsin tietoa tutkimukseen osallistuneiden oppilaiden sanakoetuloksista. Kaikilta 11 oppilaalta kerättiin tiedot heidän sanakoekeskiarvoista ennen tutkimusperiodia ja tämän jälkeen. Tämän lisäksi keräsin heidän sanakoetuloksensa kahdesta, erillisestä sanakokeesta, jotka järjestettiin tutkimuksen aikana. Molemmat näistä sanakokeista järjestettiin heille jo ennestään tuttuun tapaan. Tiedot näistä oppilaiden saamista pistemääristä ja muista asioista koottiin manuaalisesti taulukkoon. Tutkimuksen loppuvaiheessa oppilaat täyttivät lyhyen palautelomakkeen. Palautelomakkeessa heiltä kyseltiin erilaisia kysymyksiä CALL-työkalun käytöstä. Vastausten perusteella saatiin tietoa oppilaiden suhtautumisesta uuteen opiskelutapaan. Yhteensä lomakkeessa oli kuusi monivalintakysymystä. Kaikki oppilailta saadut tiedot muutettiin nimettömään muotoon. Myös opettaja täytti tutkimuksen lopuksi hieman yksityiskohtaisemman palautelomakkeen. Luokkaa ohjaavana henkilönä, opettajalla oli tarjota oma perspektiivinsä CALL-työkalun käytöstä.

Tulokset

Palautelomake

Vastauslomakkeen kahden ensimmäisen kysymyksen pohjalta voimme todeta, että suurin osa oppilaista suhtautui CALL-työkalun käyttöön positiivisesti. Kaikki oppilaat kykenivät käyttämään ohjelmaa, tutkimuksen aikana, ilman sen suurempia ongelmia. Myös opettajan palaute oli enimmäkseen positiivista. CALL-työkalun miinuspuolina opettaja näki tekniset ongelmat. Opettajan mukaan CALL-työkalu motivoi oppilaita luokkahuoneessa ja toi vaihtelua opetukseen. Vastausten perusteella, suurin osa oppilaista olisi halunnut jatkossakin käyttää vastaavanlaisia multimediaohjelmia koulussa. Kuitenkin ainoastaan kaksi luokan oppilaista koki, että opeteltavat sanat jäivät paremmin mieleen CALL-työkalun kuin

oppikirjan avulla. Tässä lyhyesti palautelomakkeen kysymykset ja oppilaiden ***[vastausmäärät]*** lomakkeessa esitettyihin kysymyksiin.

1. *Ohjelmaa oli helppo käyttää?*

   *(a) kyllä **[10]** b) ei) **[1]**)*

2. *Haluaisin käyttää vastaavanlaisia multimediaohjelmia jatkossa opiskellessani?*

   *(a) kyllä **[10]** b) ei) **[1]**)*

3. *Sanat jäivät paremmin mieleen ohjelman avulla*

   *(a) kyllä **[2]** b) ei **[3]** c) samoin kuin ennen **[5]**)*

   (1 oppilas ei osannut vastata)

4. *Ohjelmaa käyttäessäni kiinnitin eniten huomiota*

   *(a) kuvaan **[3]** b) tekstiin **[1]** c) ääneen **[5]**)*

   (kaksi oppilasta ei osannut vastata)

5. *Ohjelmaa käyttäessäni kiinnitin vähiten huomiota*

   *(a) kuvaan **[3]** b) tekstiin **[5]** c) ääneen **[3]**)*

6. *Tulevaisuudessa opettelisin asioita mieluiten:*

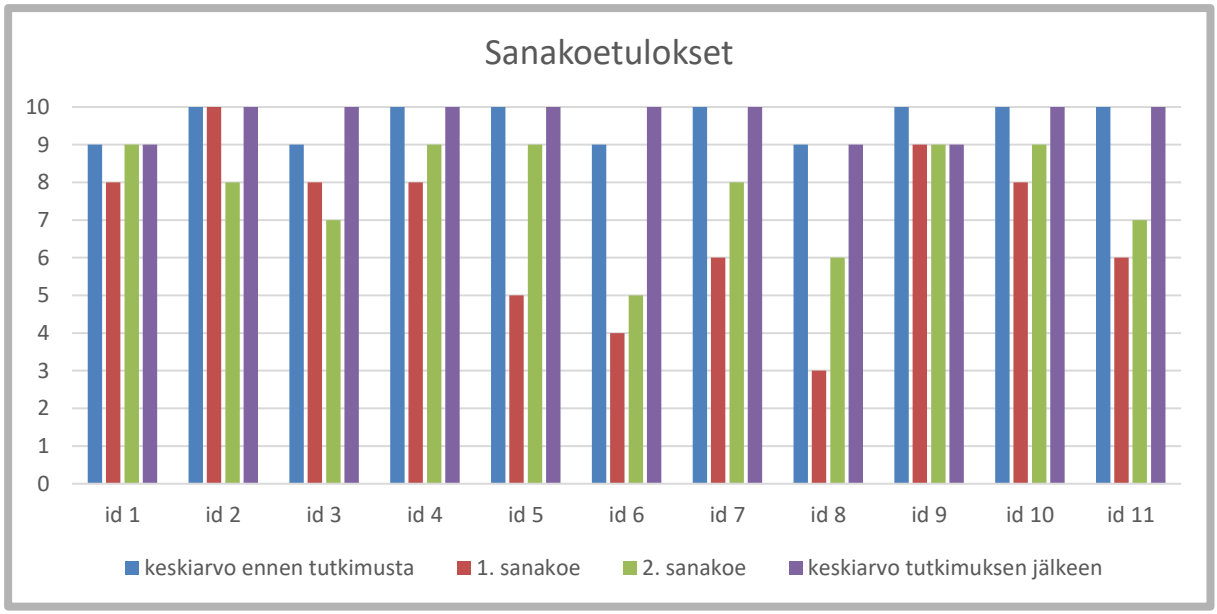
   *(a) oppikirjan kanssa **[3]** b) tietokoneen kanssa **[7]**)*

   (yksi oppilas ei osannut vastata)

Oppilaiden vastaukset erosivat toisistaan eniten multimodaalisuutta koskevissa kysymyksissä (4. & 5. kysymys). Enemmistö luokasta haluaisi jatkossakin käyttää CALL-työkalun kaltaisia ohjelmia luokkahuoneessa (6. kysymys). Kaikki oppilaat osasivat vastata viidenteen kysymykseen.

Sanakoetulokset

Tutkimuksen aikana jokaiselta oppilaalta kerättiin tietoa sanakoepisteistä, yhteensä neljä kertaa. Oppilaiden sanakoetulokset ovat nähtävissä alla olevasta taulukosta.



Luokan sanakoekeskiarvo ennen tutkimusta:

9.636~ (0.5045~ keskihajonta)

Luokan keskiarvo ensimmäisessä sanakokeessa:

6.818~, (2.1825~ keskihajonta)

Luokan keskiarvo toisessa sanakokeessa:

7.818~ (1.40129~ keskihajonta)

Luokan sanakoekeskiarvo tutkimuksen jälkeen:

9.727~ (0.4670~ *keskihajonta*)

Kahden oppilaan histogrammi kuviot olivat arvoiltaan täysin identtiset, muiden oppilaiden histogrammit taas erosivat toisistaan joillain tavoilla.

Analyysi

Oppilaiden antaman palautteen pohjalta voimme todeta, että suurin osa luokasta piti CALL-työkalun käytöstä. Kymmenen luokan yhdestätoista oppilaasta antoi ensimmäiseen ja toiseen

kysymykseen vastauksen kyllä. Ainoastaan yksi oppilas vastasi molempiin kysymyksiin ei. Kuitenkaan suurin osa luokasta ei kokenut muistavansa sanoja ohjelman avulla paremmin. Tämä saattaa olla merkki siitä, että oppilaat yksinkertaisesti haluavat oppimisympäristöönsä lisää vaihtelua tuloksista huolimatta. CALL-työkalu oli oppilaiden käytössä kuitenkin hyvin rajoitetun ajan, joten emme voi myöskään sulkea pois sitä mahdollisuutta, että oppilaiden kiinnostus ohjelmaa kohtaan vähenisi ajan myötä. Oppilaat eivät vastanneet multimodaalisuutta koskeneisiin kysymyksiin aivan yhtä yhtenäisesti. Suurin osa luokasta raportoi kiinnittäneensä eniten huomiota joko ohjelman tuottamiin kuviin tai ääniin. Puolet luokasta ilmoitti lisäksi kiinnittäneensä vähiten huomiota ohjelman tekstiin. Vaikuttaisi että ainakin tämän luokan oppilaat olisivat keskittyneet eniten juuri multimodaalisuuden tarjoamiin ominaisuuksiin. On mielenkiintoista huomata, että selvä enemmistö luokasta käyttäisi oppikirjan sijaan tietokonetta opiskellessaan. Tämä näyttäisi tukevan ajatusta siitä, että oppilailla on vahva halua hyödyntää CALL teknologiaa koulussa.

Ennen tutkimusta, oppilaiden sanakoekeskiarvot olivat korkeita. Kaikkien oppilaiden keskiarvot olivat joko ysin tai kympin luokkaa. Tutkimusperiodin jälkeisen syksyn aikana, oppilaiden keskiarvot pysyivät yhtä korkealla. On syytä epäillä, että tutkimukseen osallistunut luokka oli vahvasti motivoitunut opiskelemaan Saksaa. Voidaan todeta, että tutkimukseen osallistuneen luokan oppilaat kykenivät suoriutumaan sanakokeista, tutkimuksen ulkopuolella käyttämällä ”perinteisempiä” opetusmenetelmiä. Tietokoneen käyttö kielen opiskelussa ei siis todennäköisesti olisi heille välttämätön asia. Tutkimusperiodin aikana käytetystä CALL-työkalusta ei kuitenkaan vaikuttaisi olleen haittaa heidän myöhemmille sanakoetuloksilleen. Lisäksi 1. ja toisen 2. sanakokeen pisteiden pohjalta voidaan todeta, että jonkinlaista sanasto-oppimista tutkimuksen aika tapahtui. Jotkut oppilaista onnistuivat saamaan molemmista sanakokeista korkeitakin pistemääriä. Oppilaiden tuloksissa olivat selvästi havaittavissa vaihtelua.

Yhteenveto

Miten luokka siis lopulta suhtautui CALL-työkalun käyttöön? Tutkimukseen osallistuneen luokan Saksan kielen oppilaat suhtautuivat CALL-työkalun käyttöön positiivisesti. Palautelomakkeen vastausten perusteella, suurin osa oppilaista piti ohjelman käytöstä. Ohjelman käyttö ei myöskään haitannut muuta opetusta. Lisäksi luokan opettaja koki, että oppilaat hyötyivät työkalun käytöstä. Opettajan mukaan työkalu onnistui motivoimaan oppilaita ja sen käyttö auttoi myös opettajaa ymmärtämään heidän oppimisprosessiaan

paremmin. Todettakoon, että oppilaiden vuorovaikutus CALL-työkalun kanssa oli onnistunutta. Kahden tutkimuksen aikana järjestetyn sanakokeen tulosten pohjalta voidaan myös todeta, että CALL-työkalun käytön aikana jonkinlaista sanasto-oppimista tapahtui. Tilastollisten testien puutteen vuoksi, oppilaiden sanakokeista saaduista pistemääristä ei voitu tehdä sen suurempia johtopäätöksiä. Sama päti myös oppilaiden multimodaalisuus mieltymyksiin. Enemmän ja parempaa dataa oltaisiin todennäköisesti voi saada, jos tutkimuksessa oltaisiin käytetty tietokone tuutorin roolissa toimivaa CALL-ohjelmaa. Monimutkaisemman ohjelman avulla olisi voitu kerätä yksityiskohtaisempaa tietoa, esimerkiksi oppilaiden napin painalluksista ja ohjelman käyttöajoista. Todennäköisesti myös jo valmiin, Duolingon kaltaisen, oppimisympäristön hyödyntäminen olisi voinut helpottaa tutkimuksen tekoa.

Tutkimukseni tarkoituksena oli kuitenkin selvittää, kyettäisiinkö nollabudjetilla luotua, yksinkertaisempaa CALL-työkalua hyödyntää modernissa luokassa. Voi todeta, että tutkimuksessa käytetty avoimen lähdekoodin CALL-työkalu kykeni motivoimaan luokan oppilaita. Oppilaat olisivat myös jatkossa halunneet hyödyntää vastaavanlaisia ohjelmia. On hyvä nähdä, että vielä nykypäivänäkin yksinkertaisempikin CALL-työkalu kykenee aktivoimaan kielenopiskelijoita modernissa luokkahuoneessa. Opettajat hyötyisivät varmasti paljon, jos heidän olisi mahdollista hyödyntää moista teknologiaa luokissaan. Mielestäni koulujen tulisi hyödyntää avoimen lähdekoodin sovellusten tarjoamia ominaisuuksia. Tutkimuksessa käytetyn CALL-työkalun kaltaista ohjelmaa voitaisiin ainakin käyttää oppikirjojen rinnalla tuntiopetuksen monipuolistamiseen.