Utrecht University

Evaluating the effectiveness of a tool to improve vocabulary skills in an academic context:

MySpeechTrainer

Linguistics RMA program

Master's thesis

Camilla Giannini

6142931

Supervisors:

dr. Helmer Strik (Radboud University)

dr. Hayo Terband

Table of contents

Abstract

Recently, many studies (Beatty, 2013; Golonka, Bowles, Frank, Richardson, and Freynik, 2014; Plonsky and Ziegler, 2016) showed that the use of CALL systems (Computer Assisted Language Learning) has a significant effect on how we learn languages. Considering that most of these applications were designed for English as a Foreign Language classrooms, it is still necessary to investigate how these systems can support students in an academic context. The present study aims at evaluating a prototype application called MySpeechTrainer (MyST). The company Novo Learning developed it and adapted it for research purposes by Radboud University. The students enrolled in a Semantics and Pragmatics course and a Phonetics course at Utrecht University used it to improve their English vocabulary in an academic context. We evaluated the participants' learning outcomes with a pre and post-tests. Moreover, we analysed the log data of MyST to understand the students' behaviour. While no significant differences were found with the analysis of the pre and post-tests, the log data provided useful information that can be used to improve the application in the future.

Keywords: Computer-assisted language learning, vocabulary acquisition, e-learning

Chapter 1

Introduction

With the advent of the internet, technology has become an integral part of everyone's daily life, as it is accessible to the public at low prices and from any location. In the past years, this progress in technology offered attractive alternatives to traditional education. It is changing the way we learn languages as new applications are becoming a component of many University curricula. The study of how computer technology is used in classrooms for language learning falls in the domain of Computer Assisted Language Learning technologies (CALL) has been a central topic for many years (Hubbard, 2009; Golonka, Bowles, Frank, Richardson, and Freynik, 2014; Plonsky and Ziegler, 2016). Many of the studies on CALL, showed that language learning technologies can positively affect the instructors' teaching strategies, by providing the tools to organize the course in a more efficient and interactive way. Moreover, the latest innovations of technology used in CALL, as ASR, Automatic Speech Recognition, can facilitate learning. They increase motivation in students and give them access to additional and personalized input and feedback at any time and everywhere.

The applications can also offer support to overcome the challenges the students face with specialized vocabulary in ESP, English for Special Purposes, classrooms. As the use of the English language is becoming compulsory in many bachelor and master courses, the main concern is that students do not have an appropriate level of English. The knowledge of vocabulary allows them to effectively communicate in academic contexts and to deal with the topics.

These are some of the advantages that CALL systems offer, and the reasons why they are more frequently an active part of many University courses. Therefore, the use of CALL is often beneficial, but it can also cause complications to the already multifaced field of language acquisition. For example, these systems are frequently used outside of the classroom, and in these cases, the students have independent access to the activities. As different studies illustrated (Hwu, 2003; Fisher, 2007), when the students use the application for self-study, they do not always perform as expected by the instructor and the designer. Many concerns remain on the use of CALL in an academic setting, where learning English as a second language is not the main purpose of the course. Previous studies on CALL (Chiu, Liou and Yeh, 2007; Golonka et al., 2014) investigated the efficiency of computer technology by testing students in EFL,

English as a Foreign Language, classrooms, but an in-depth knowledge of how e-learning can be used in an academic context is still needed.

The awareness of CALL effectiveness requires more research focused on understanding what it is the optimal use of these systems for specific language skills. An efficient method is to analyse the students' outcomes and to investigate their behaviours while using the application.

The aim of the present study is to evaluate a pilot application designed for bachelors' students in the Netherlands, with the purpose of improving their English vocabulary and speaking skills in academic contexts. The application, called MySpeechTrainer (MyST) uses Automatic Speech Recognition (ASR) software and it was developed using the content of Novo Learning Player then adapted for research purposes by Radboud University. It is believed that the results of this research will present a more in-depth knowledge of MyST application. This study will evaluate the problems, as well as the positive features of the prototype application, and it will propose solutions to enhance learning. Moreover, this research will contribute to the understanding the learners' attitudes towards the use of technology for language learning, and it will provide further information to the instructional designers and professors.

Chapter 2

Background literature

This chapter presents an overview of the theoretical studies in CALL relevant to this research. It begins with a description of CALL and how it is used for vocabulary acquisition. The chapter also provides a summary of the previous studies on CALL effectiveness for language learning. The chapter concludes with considerations on the most effective methods to evaluate students' outcomes using e-learning systems.

2.1 What is CALL?

CALL is the acronym for Computer-assisted language learning, which refers to the vast and complex field of research that investigates the use of technology for language learning. It is defined by Beatty (2013) as "any process in which a learner uses a computer and, as a result, improves his or her language" (p. 7). This definition might appear broad, but CALL research is a multifaced field where the media, the context of learning or the pedagogical theories can always differ, and they can adapt to the learners' need. Therefore, the definition provided by Beatty (2013) is interesting as it covers different types of CALL researches, highlighting the central role of language learning. Hubbard (2009) explains that the word computer not only refers to the classic use of the laptop, but it also indicates a range of technological devices, like smartphones, tablet, mp3, and so on. Despite this, the choice of the term CALL has also been criticized, as it leaves out some important components of language learning, for example, the role of teachers (Chinnery, 2008).

Computers are used in the field of language acquisition since the late 1960s, and extensive literature can be found on this topic (Hubbard, 2009; Beatty, 2013; Golonka et al., 2014; Plonsky et al., 2016) technology is continuously evolving, and research on the matter always needs to be updated. Nowadays, new tools as ASR allowed an increased interaction with users in CALL systems. This technology gave the chance to deal with the unpredictable answers of the students and to provide personalized feedback.

For these reasons, the use of CALL systems has been for long at the center of the investigations. Beatty (2013) explains that some issues discussed in the earliest researches on CALL are no longer relevant. For example, she claims that it was already demonstrated that computers should be used in classroom, and that nowadays, students and teachers are able to

use technology. Current research should focus on "*how* computers should best be used and *for what purposes"*, as "a major challenge to many studies in CALL remains a lack of empirical research." (Beatty, 2013, p. 15). It is important to know if the available applications nowadays are effective for the students' needs and if they offer a fun and interactive learning environment. In fact, Hubbard (2009) found that the most unanswered questions still lie in the topic of CALL effectiveness. In summary, the problem is not anymore if CALL should be used or not, but now it is important to understand how the students make use of the available resources.

In this context, the application tested in this study, MyST, offers an appropriate environment for second language acquisition. Students can access it from everywhere and at any time, with the possibility of creating unlimited input, and thanks to ASR they can receive immediate feedback (Neri, Cucchiarini, Strik and Boves, 2002).

2.1.1 Vocabulary learning in CALL

In language learning, vocabulary is known to be a significant factor, as it is crucial for listening and reading skills and above all being able to communicate. However, vocabulary development and acquisition have often been overlooked by traditional SLA researches (Nation, 2001; Seregély, 2008; Milla and Mayo, 2014), where the focus is mainly acquisition of grammar. The assumption of these studies is that vocabulary is incidentally acquired through context, as it happens in the L1, "the child as well as the L2 learner in whatever kind of naturalistic environment, is an extensive amount of linguistic input" (Seregély, 2008). However, as mentioned by Hubbard (2009), applications built for vocabulary training are some of the most common in CALL because they are easy to program and use.

In an academic context, the students need to have an adequate level of English and a strong knowledge of technical vocabulary, to learn and communicate in class. With rich vocabulary, students can better express their ideas in a written or oral form (Manik and Christiani, 2016). As the authors remark, the students might not even be aware of their lack of vocabulary knowledge. When the students encounter technical words in a text, they might ignore them, affecting the understanding of the topic discussed. However, when language learning is not the focus of the class, the study of vocabulary or other language skills cannot be incorporated into the syllabus, mainly for the lack of time. In addition, professors prefer to give more importance to the curriculum topics and not to the English vocabulary skills (Brooks, 2014).

CALL systems offer an opportunity for students to learn new vocabulary in a fun and interactive way. The attractive design of most of the applications motivate the students to use

the application in autonomy. In addition, most of the applications can be used everywhere with an internet connection, giving the possibility to the students to practice at their own pace. Another advantage is that self-study can develop the students' personal learning strategies that they can apply in the future in other contexts. In fact, learners are constantly exposed "to linguistic environment in which a word is encountered and facilitating full understanding of a word" (Nation, 2001).

2.1.2 Pronunciation in CALL

Another important component in language learning is pronunciation. With pronunciation we intend the way words in a sequence are spoken; and it is a general term that include different phonemic and prosodic components. In second language acquisition research, as well as in CALL, different language skills including speaking and pronunciation, reading and writing, vocabulary and grammar are often analyzed separately. In the reviews presented earlier (Chapelle, 2001; Hubbard, 2009; Beatty, 2013; Golonka et al., 2014), many studies focus on distinct aspects of language acquisition. However, other authors (ex. Kebede, 2015) showed that, as each skill is important, their combination has a positive effect on learning. For example, providing pronunciation instruction in conjunction with vocabulary-focused learning activities is beneficial for the students' fluency during communication (Ahmad, 2016). Therefore, technology for language learning should integrate different skills, rather than train the students separately for each of them. With ASR technology, MySpeechTrainer offers the possibility to integrate pronunciation with vocabulary skills practice.

2.2 Previous studies on CALL

As it was mentioned before, the research literature on CALL and language technologies, from the early 1980s until now, is abundant. The aim of these studies is to describe the CALL systems available (Hubbard, 2009) and their effectiveness for second language learning (Felix, 2005; Golonka et al., 2014; Fisher, 2004, 2007). The literature on CALL is rich, but there is still a lot to investigate, as technology is constantly changing and developing. Nowadays, many different e-learning applications are available on the market, both for individual use and for practice in class. However, many of these applications might impress the buyers for their "fancy-looking systems" (Neri et al, 2002, p.2) but they fail to assist language learning effectively. In the literature available describing CALL systems, the effectiveness of the

environments for language learning has been at the center of discussion for many years. However, in his literature review Felix (2005) points out that research has produced many ambiguous results. For the author, there is not a common definition of what is effectiveness, and he found many common problems in the research designs of the studies reviewed. Therefore, it is still essential to understand what the best research method is, and to investigate the application efficiency for language learning.

Felix (2005) defined the word *effective* as somethings "having an effect, producing a result, bringing something to pass" (pag.4). Moreover, Hubbard (2009) explained that an application can be considered effective when the "learners retain language knowledge or skills longer, make deeper associations, and/or learn more of what they need" (pag.2). Felix (2005) further reports well-designed researches which succeeded at showing the positive impact of technologies on the learning process. The positive effect of CALL was particularly found for vocabulary, reading and writing acquisition.

More recently, Golonka et al. (2014) reviewed more than 360 studies to examine the effectiveness of different technological systems for language learning. In this review, the authors focused on the studies where traditional teaching methods are compared to modern technological learning approaches. Despite the strong support that most of the studies reviewed made on the benefits of technology for teaching, the author argues that only moderate effects were found on language learning. Most of the studies show how well these systems support language learning and not if they influence the learning outcomes. Among the papers analysed by Golonka (2014), the strongest evidence for the influence of technology in language learning was found for ASR or chat (Golonka et al., 2014). In the first case, ASR can facilitate speaking more effectively than in interaction with teachers.

Moreover, Plonsky et al. (2016) described and evaluated the methodological practices adopted by the recent CALL-SLA syntheses and meta-analysis studies. The authors found that the results of the studies reviewed mostly indicate a positive effect of the use of CALL for second language learning (pag.32).

The learners' positive attitudes towards technology was a common finding in the studies reported by Hubbard (2009) and Golonka et al. (2014) who remarked that "the students prefer using technology over more traditional methods and materials" (p. 92). This type of evaluation was found not only for applications using ASR and chat, but also for other environments, like blogs or electronic dictionaries. Therefore, it is important for applications used by students to be user-friendly, to motivate them and continuously renew their interest in the learning environment, particularly if the students use them in their free time.

Even though the students seem to be positively motivated using technology, other studies focusing on students' behavior (Chapelle and Mizuno, 1989; Hwu, 2003; Fisher, 2007) showed that students don't always use the applications the way teachers and designers expect. Tracking students' behavior in CALL has become an increasingly used method in studies investigating CALL effectiveness. A study on students' learning autonomy was carried out by Chapelle and Mizuno (1989), where they found that "students are often doing something different from what instructors believe they are doing; they do not always use the optimal strategies" (p.42). In the more recent studies of Fisher (2004, 2007), the author compared student's self-reports with tracking data and found as well that the perceived use of the application was negatively correlated to the real use. In the study of 2004 the students, who had to read a document in French containing marked and unmarked hyperlinked words, were not aware of how many times they were clicking the marked words. This shows that the students are often not aware of how they use the application; therefore, the data from students' surveys should be integrated with tracking methods (Fisher, 2007, p.428). As the author further explains, before hypothesizing the positive effect of CALL on language learning, it is important to understand if students make correct use of the application.

2.3 Evaluating student outcomes in CALL

Determining the success of language learning is an important step in the evaluation of the effectiveness of e-learning tools. However, it is also central to understand how to test CALL applications correctly.

Chapelle (2001) identified two different methods, called judgmental and empirical, to evaluate CALL systems. The former is based on the researcher's observations on the topic using SLA (second language acquisition) theories, the latter refers to the interpretation of observable data, for example, the student's outcomes. Chapelle outlines three levels of analysis using these methods. The first refers to the evaluation of CALL software, which aims at understanding if it was developed in line with the students and the teacher's needs. The second type of analysis investigates the use that teachers make of CALL activities. These first two levels are conducted using SLA theories, with a judgmental method. The third type refers to the evaluation of the student's performance, which is carried out using an empirical methodology. Chapelle further claims that, for an adequate analysis of the CALL systems, the judgmental and empirical methodologies should be used in combination.

For the purpose of this research, only the empirical data on the use of MySpeechTrainer application will be analysed. The aim of this study is in fact to understand to what extend MyST is effective for language learning in an academic context. To answer this question, the use of the application needs to be analysed.

Furthermore, Chapelle of 2001 and Jamieson, Chapelle, and Preiss (2005), in their studies, present six criteria to investigate CALL systems. They aim at creating a trustworthy framework that can help the researchers in the future. The present study will focus on one of these criteria called *learning language potential*. The questions proposed by Chapelle (2001) to investigate this criterion are: "What evidence suggests that the learner has acquired the target forms that were focused on during the CALL task? What evidence indicates that learners focused on form during the CALL tasks?" (pag. 68). The research purpose is, in fact, to understand if the learners acquired the target linguistic forms found in the application. To answer these questions with empirical method, the author proposes a pre and post-test.

Considering Chapelle's (2001) framework, Hubbard (2006) also reflects on the methods that can be used for the evaluation of CALL software. However, contrarily from Chapelle (2001) and Jamieson et al. (2005), the author presents more in detail different approaches to assess the students' outcomes. This analysis is considered as the final point in the process of evaluation of CALL, and it will determine if the learning was successful or not. The students' outcomes can further help the researchers to understand how to improve the software and how to better use it in the future. The different types of empirical evaluations that the author identifies are: *observation*, *student surveys, students' journals, pre and post-testing, tracking systems*.

The first method described is observation. It is the more direct way of understanding how students use the application. For example, the teacher observes and takes notes of the students' behaviour, in a lab, or in a classroom setting.

Student surveys and students' journals are two other methods mentioned by the author. The first consists of surveys or questionnaire where the students give their opinions on the application they just tried. Despite this information being potentially useful, the author explains that it is not always a reliable method; the results can be compromised if the students know they will be graded. Furthermore, it was shown (Fisher, 2007) that what students believe they do, it's often different from how they behave. For Hubbard (2006), this method is still valuable when used with one or more of the other methods described.

The students' journals are reports made by the students where they have to reflect on their work and on the activities. For Hubbard, this type of data is mainly used for student' benefit and it is not suitable for research, as it is too subjective and not always reliable.

The other two methods described by the author are pre and post-testing and tracking systems, which will be used in the present study. We believe from the mentioned literature review, that these two methods of evaluation will provide an overview of the effectiveness of MyST application. The student's outcomes will be first analysed with a pre and post-tests. In addition, the analysis of the log data will present their behaviour after using the application. It is believed that the combination of the two methods, described in detail below, will provide evidence for the effectiveness of MyST.

2.3.1 Pre and post-tests

This type of test is used to gather information about the learners' outcomes, by comparing the scores of a pre-test, before practice, and of a post-test, after. The test should have similar structure, same number of questions, same level of difficulty. The pre-test is conducted before the start of the treatment to assess the initial level of a participant's specific skill. Further evidence of the student learning is provided if the pre-test data showed that the students did not have previous knowledge of the language skill (Chapelle, 2001). Subsequently, the post-test is conducted at the end of the treatment and it can indicate an improvement of the skills tested. This research method is used to test different language skills and not only vocabulary (Isti'anah, 2017). However, this pre and post-tests do not show enough evidence of the students' learning behaviour. For example, it does not illustrate how the students used the application or if they used it as the designer or the teacher predicted (Hubbard, 2006).

2.3.2 Tracking systems

Like many other applications for language learning, MyST is used by the students outside of the classroom, in autonomy. The aim of the activities on the application is, in fact, to provide additional support to students improving the vocabulary that they might need in class. The most effective method to understand how the students practiced in autonomy is tracking their activities. Many applications require external tracking software, while MyST platform gives the possibility to track how the students are using the application and to download the related log data.

The log data offers different information on the application usage. It saves data on the number of answers given and time they spent on each. As it was mentioned before, Hwu (2003) and Fisher (2004,2007) explained that when students can decide when and how to access to the activity, they behave differently from what the teachers and the designer of the activities expected. In his study, Fisher (2007), presents the advantages of using tracking as a research method. It provides evidence on how students really use the systems in autonomy, by tracing what they do. This method also provides a more realistic context compared to the student's journals or the questionnaires. Moreover, it offers information useful to the instructional designer to improve the systems

2.4 MySpeechTrainer

The pilot application called MySpeechTrainer (Mijn Spraak Trainer or in short, MyST), tested in this study, was developed with a collaboration between Novo Learning and the Center of Language and Speech Technology at Radboud University. The activities were designed in the Novo Learning platform, using an ICT software called Novo-CALL. The platform has an intuitive design easily accessible for the learners and for the designers. It gives the possibility to visualize the learners' progress, the report of the time spent, and the scores per activity. However, for this study, I did not have the possibility to download this information on an external file. For this reason, I had to copy the information needed from Novo Studio to Excel[1].

MyST application is available for download on different devices (iOS and Android) and web browsers (Chrome, Firefox, Edge). Moreover, it offers the possibility not only to practice the students' vocabulary skills in an academic environment, but also it allows them to train their speaking skills. This is possible with ASR technology that recognizes the speakers' pronunciation errors and provide them with an immediate feedback. Moreover, the variety of exercises allow not only to practice different language skills but also to use these skills in a specific context. The activities were entirely designed on the Novo Platform and will be thoroughly described in chapter 3 more in detail.

---

[1] Microsoft Office 365 ProPlus

2.5 Research Questions

Previous research shows that CALL systems can be effective for language learning. The applications can be used in addition to traditional methods to offer a stimulating and interactive way of learning that motivates students.

Previous studies (Felix, 2005; Golonka, 2014) focused on application in English as a foreign language (EFL) setting. This study evaluated a tool that focus on subject-specific language knowledge. The students who participated were enrolled in two courses at Utrecht University, Semantics and Pragmatics and Phonetics, which were entirely taught in English. The application was used by the students as an additional tool while attending the classes, and it offered them extra support for the specific vocabulary they needed.

This study combined a pre and post-test design with an analysis of the log data. The aim is to test whether the employment of the pilot application, MySpeechTrainer, is efficient to learn vocabulary in an academic context. Moreover, this study aims at understanding how students used the application for self-study outside of classrooms. The need to improve CALL systems, as MyST, and to understand better how they can assist and facilitate learning, raised the following research questions:

RQ1: To what extent the e-learning environment MyST improve students' vocabulary knowledge?

RQ2: How did the students use MyST in autonomy?

The first research questions will be answered using a paired t-test, by comparing the mean of the pre and post-tests results. To answer the second research question, the log file data from Novo Studio platform will be analyzed, as the pre and post-test data only evaluate the students' learning outcomes, and it does not provide evidence on how the learning process occurred.

The evaluation of the exercises, which were created to improve English vocabulary for the course of Semantics and Pragmatics and of Phonetics on MyST, will contribute to understand if this application can help students to improve their English vocabulary skills. The outcome of this research will also provide more evidence on MyST application, which will help to improve the design of the activities.

Chapter 3

Method

This chapter presents the method used to conduct this research. First, this chapter describes the participants of the study then the method used for the data collection and the structure of the activities on MyST.

3.1 Participants

The participants of this study were bachelor's students at Utrecht University enrolled in the Semantics and Pragmatics course in the first block of the second semester, and in the Phonetics course in the second block of the second semester. These courses used MySpeechTrainer as an extra, non-graded, assignment.

In the first group, the course of Semantics and Pragmatics, there were 31 students. To answer the first research question, only the data from the 11 students who completed the pre-test, post-test, and the questionnaire were analysed. The fact that most of the students did not complete the post-test, or dropped out from the course before the end, is the reason for the low number of responses. The students were ten female and one male Dutch native speakers between 19 and 24 years old. The participants were enrolled at Utrecht University in the bachelor's degree of Linguistics. Except one, all the students were finishing their second year. For the log file analysis, we examined the answers from the students who tried at least one activity on MyST, a total of 17.

The second group consisted of the students enrolled in the course of Phonetics. As it happened for the Semantics and Pragmatics group, many of them did not complete MyST course, or the tests assigned. The students who finished at least one activity were 30, among them 25 filled out the questionnaire. Moreover, the number of answers collected from the pre and post-test was 21. The questionnaire showed that 25 students were enrolled in a bachelor's course, 22 were female and 3 males. Among them, 24 were from Utrecht University and one from Sung Kyun Kwan University. There were 23 Dutch native speakers and two with different native languages, Catalan and Korean. Differently from the Semantics and Pragmatics group, in this case, most the students were enrolled in different courses at the Faculty of Humanities; moreover, there were 8 students from the Faculty of Science and other 2 from the Faculty of Social and Behavioural Science. One of the participants did not answer this question.

3.2 Material

This section describes the material used for the research. It is divided into two subsections; the first outlines how the data collection was conducted, the second describes the activities on MyST that were completed by the participants.

3.2.1 Data collection and instruments

The data was collected over the course of one semester from February until June 2019 at Utrecht University. The group of Semantics and Pragmatics was tested from February 6, 2019, until March 25, 2019; the group of Phonetics from April 21, 2019, until June 7, 2019. At the beginning and at the end of the courses, the students completed a pre-test and a post-test. The aim was to evaluate their vocabulary knowledge and to understand whether there was an improvement after the use of MyST. The tests were designed using the terms and types of exercises present in MyST.

In each test there was a total of 20 items, 15 on vocabulary and five on pronunciation. Ten words of the test were retrieved from MyST. In addition, in the post-test, only six of the words from MyST were new and the remaining four were recycled from the pre-test. As it was not possible to evaluate the pronunciation items, only the scores of the vocabulary items from MyST, a total of 7, were considered. A score of 1 was given if the answer was correct or 0 if it was incorrect. Therefore, the highest score that the students could receive, both in the pre and the post-test, was 7, and the lowest score was 0.

At the end of the block the students were asked to complete a questionnaire with demographic questions. The questionnaire was designed in Qualtrics[2], and the data was analysed with SPSS software[3].

For the second part of the analysis, the log data was collected from Novo Studio and copied in Excel for the data cleaning. In this process, the unwanted and irrelevant observations were eliminated. For example, we excluded the recordings that did not correspond to the real time spent on the activities. The fact that a student opened a slide did not necessarily mean that he or she was working on it; when using the internet browser, it is possible to minimize the Novo Studio window without interrupting the tracking time. To control for this, the

---

[2] The questionnaire was designed for the Master thesis of Anna Ovchinnikova at Radboud University.
[3] IMB SPSS Statistics 23

observations where the time was higher than 3 minutes were removed before the data analysis. For example, in one of the recordings, the average time spent on one slide was 39 minutes in another 9 hours. The process continued with the correction of typos or inconsistent capitalizations. After the data cleaning, we started the analysis, which will be presented in section 4, focusing on the time spent by the students and their scores.

3.2.2 Structure of the activities on MyST

The participants in the two groups practiced their vocabulary and pronunciation using two different courses on MyST. The structure of the two courses was the same, and the only difference consisted of the words and topics trained. One course included words on the topic of Semantics and Pragmatics and the other on Phonetics. The courses were divided into 7 weeks and consisted of different topics corresponding to the reading assigned, and to what the teacher explained in class. Every week a new activity was published and made available to the students. The topic was first introduced, and the objectives summarized, as it is illustrated in Figure 1. Before every set of exercises, a new *slide* with an instruction was provided, like the one represented in Figure 2.
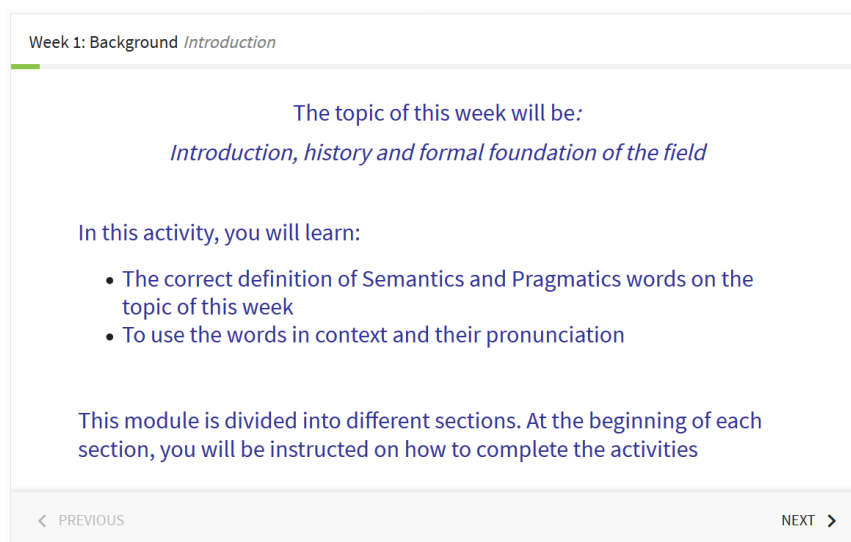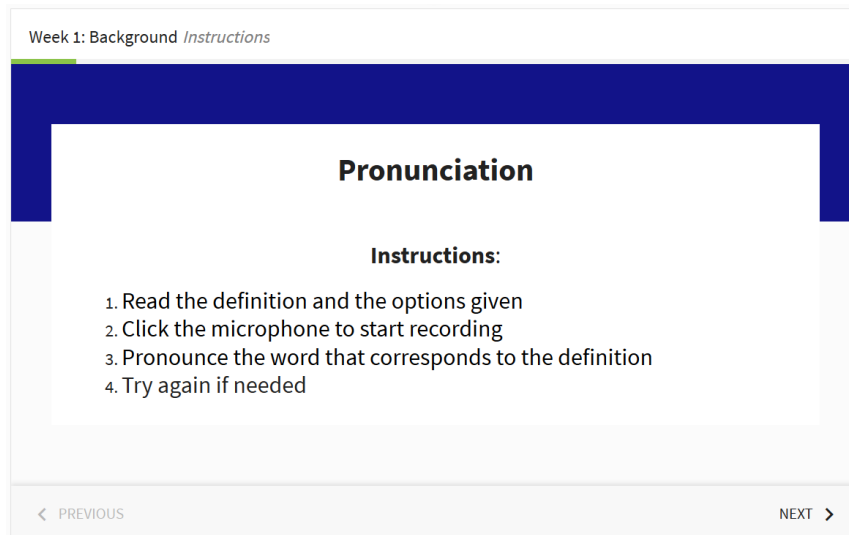

Figure 1: Sample Introduction page

Figure 2: Sample Instruction page

Each question appeared on a different *slide*. On the bottom of the screen there was the command *next* to continue or *previous* to go back; on the top, a green bar informed the students about the progress of the activity, (Figure 1 and 2). Every week the students could train their vocabulary knowledge and pronunciation with different types of activities: *pronunciation, multiple choice, gap filling, drag and drop, audio choice*.

Pronunciation: The students had to pronounce the word that appeared on the screen. They received feedback both if they answered correctly or incorrectly. They could see the IPA transcription by clicking on the word on the screen. The incorrect feedback included the transcription of their mistaken pronunciation of the target word, in grey, and the correct one, in green. They could also click on the symbols to hear the phoneme. The correct feedback only showed the IPA transcription of the target word.

Multiple choice: the correct answer to the question was chosen between different options. The students could either pronounce the correct answer or tap/click on it.

Gap filling: the gaps in a sentence needed to be filled by clicking on the correct option.

Drag and Drop: the words need to be connected to the corresponding definition.

Audio Choice: the correct answer to the question was chosen between different audios.

At the end of each activity, there was also a review and a summary, where the students could find a list with all the words they used in the exercises, figure 3 and 4.
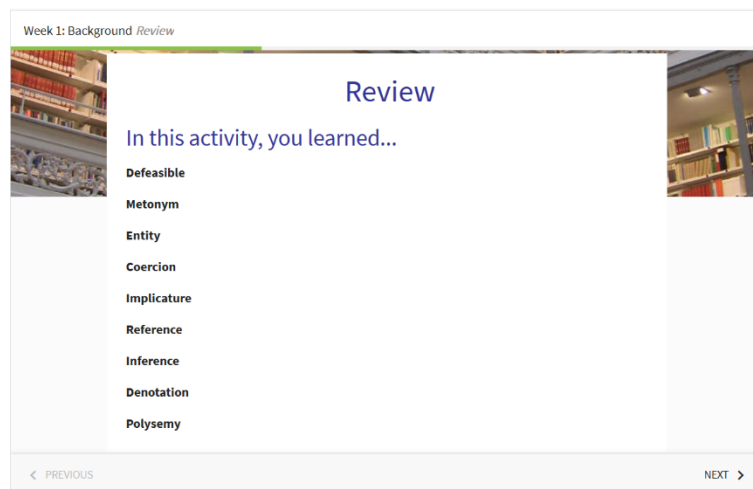
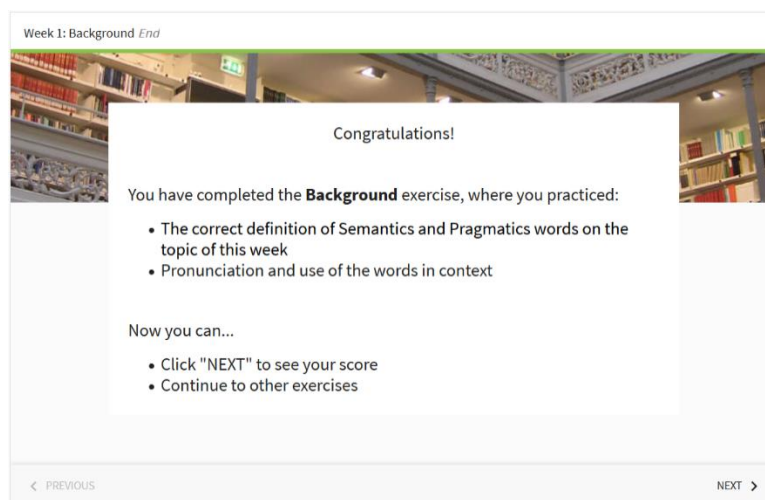

Figure 3: Sample review page 1

Figure 4: Sample review page 2

Chapter 4

Analysis and results

This section presents the analysis and the results of the pre and post-tests scores and the descriptive analysis of the log files of MyST. The results were obtained using SPSS software.

4.1 Vocabulary tests

To answer RQ1, a paired t-test between the scores of a pre and post-test was conducted. Table 1 shows the differences between the mean of the pre-test and post-test scores for both groups. The maximum score possible was 7, only considering the target words found in MyST. For the Semantics and Pragmatics group, the mean of the pre-test scores was 4.45 and the mean of the post-test was 5.36. In the Phonetic group, the scores of the pre and post-test were very similar.

Table 1: Mean scores of pre and post tests for both groups

Group	n	Pre-test	Post-test
Semantics and Pragmatics	11	4.45	5.36
Phonetics	21	5.43	5.48

A paired t-test was conducted to compare the mean scores for the Semantics and Pragmatics group. The output of the analysis is illustrated in Table 2. The difference between the scores was not significant (p=.148). The analysis of the tests demonstrated that the students' knowledge of the English terms of Semantics and Pragmatics did not improve over the semester with the use of MyST. These non-significant results could be explained by the low number of participants (N=11), or the short amount of time spent practicing with the application, as we will illustrate in section 4.2.1. Moreover, some of the students already received high scores in the pre-test, so they could not improve more their vocabulary knowledge.

Table 2: Paired Samples t-test for the Semantics and Pragmatics group

	Paired Differences					
	Mean	Std. deviation	Std. Error Mean	t	df	Sig. (2 – tailed)
Pre-test Post-test	-,909	1,921	,579	-1,569	10	,148

Figures 5 and 6 show the distribution of the pre and post test scores for the Semantics and Pragmatics group. In the pre-test the students scored mostly 4 out of 7 points, while in the post-test most of the scores were above 5.
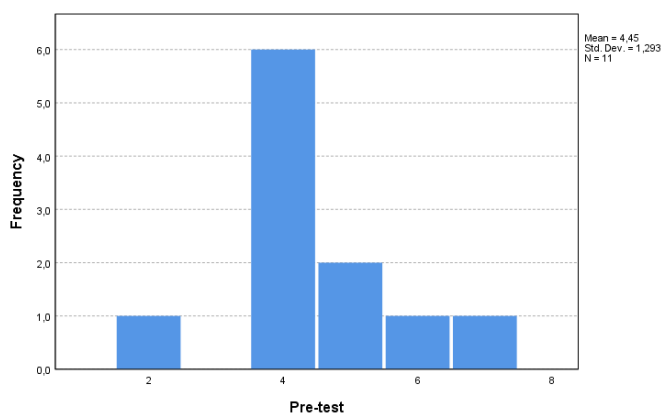

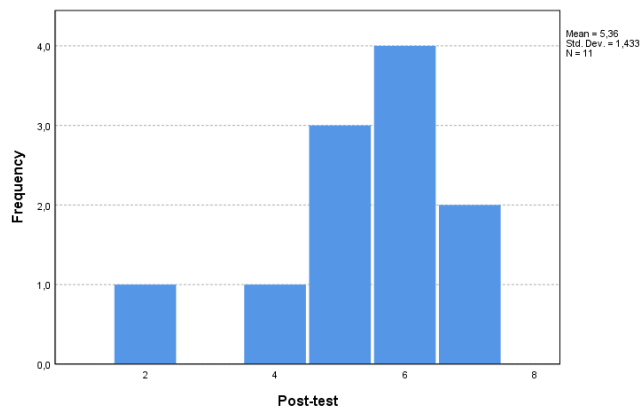Figure 5: Frequency table of pre-test scores Semantics and Pragmatics

Figure 6: Frequency table of post-test scores Semantics and Pragmatics

The graph in figure 7 represents more in detail the scores of the pre and post-test for the Semantics and Pragmatics group. One student, who scored 2 in the pre-test, received 6 in the post-test. Those who scored 4 in the pre-test either improved in the post-test or received a lower or the same score.
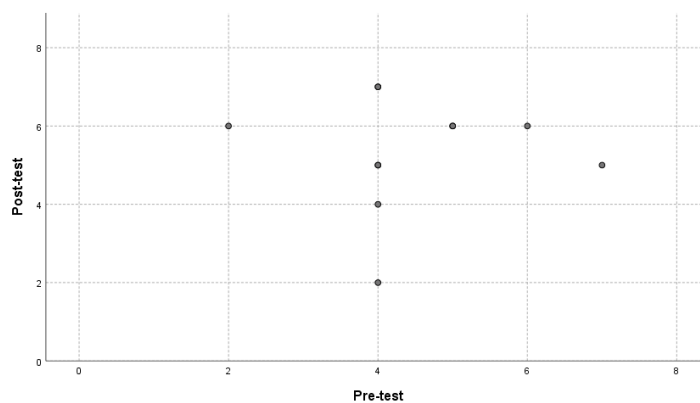

Figure 7: Scatterplot of the pre and post-tests scores for the Semantics and Pragmatics group

As we presented in table 1, the means of the scores of the pre and post-tests, for the Phonetics group, did not show any relevant difference. In fact, the mean of the pre-test scores was 5,43 and that of the post-test was 5,48. These results already suggests they did not improve

their vocabulary knowledge with MyST, and for this reason the paired t-test was not conducted for the Phonetics group.

Figures 8 and 9 illustrate the frequencies of the scores for the pre and post-test of the Phonetics group (N=21). In the pre-test, the students scored mostly 6 out of 7 points. Compared to the pre-test, in the post-test, fewer students scored 6, and more students scored 4 or 5.
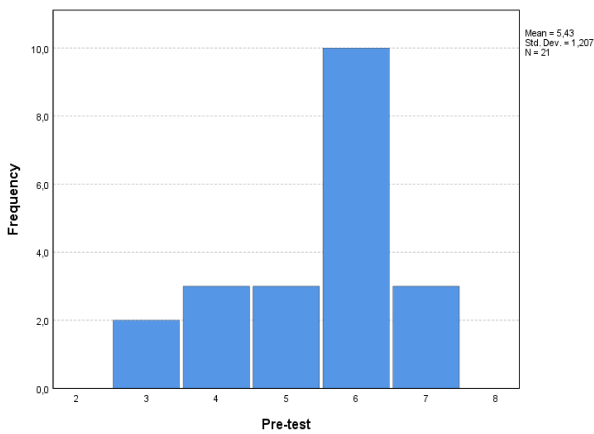

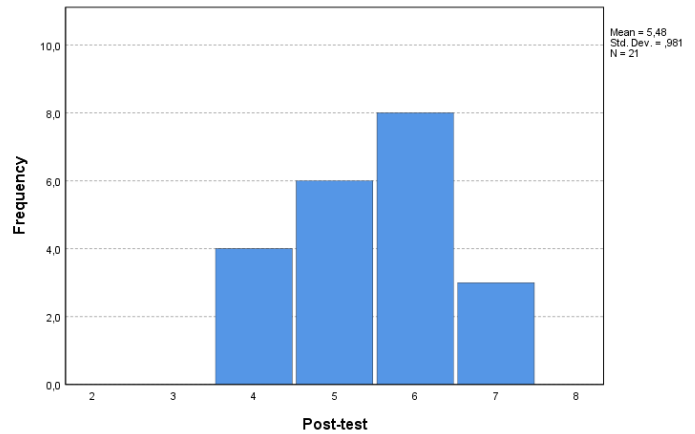Figure 8: Frequency table of pre-test scores Phonetics

Figure 9: Frequency table of post-test scores Phonetics

The scatterplot in figure 10 is a more detailed representation of the students' individual differences. Overall, the plot shows that only a few students improved in the post-test. The others either scored the same or got worse.
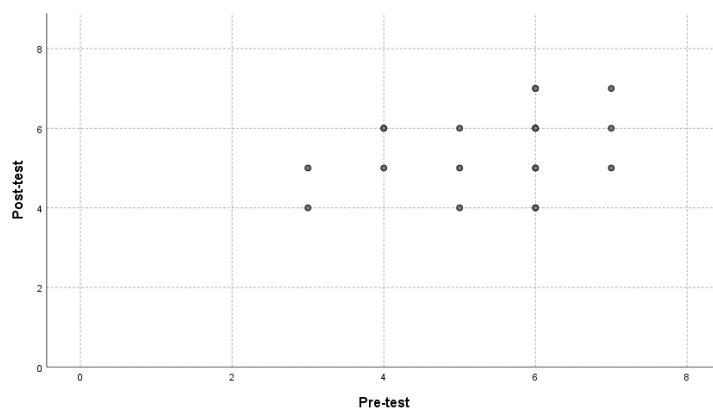

Figure 10: Scatterplot of the pre and post-test for the Phonetics group

Therefore, the results indicate that the students in the group of Semantics and Pragmatics scored better than the students who took the Phonetics course. However, in none of the groups was found a significant difference between pre and post-test scores.

4.2 Log data

This section presents the outcome of the analysis of the log data, which will be presented separately for each group. The investigation was conducted using the data extracted from Novo Learning platform. This platform provides different information about the students' behaviour when using the application, as the time spent on the exercises, the percentage of correct answers, and the attempt taken for each question.

4.2.1 Semantics and Pragmatics group

The graph below illustrates the average number of minutes spent by each of the 11 students during the 7 weeks of training with MyST. The students spent more time on MyST the first week but, overall, from week 2 to 7 the time spent remained constant. The average time in Week 1 was 8 minutes. In Week 2 and 4, the time spent decreased to less than 6 minutes, but overall it remained stable.
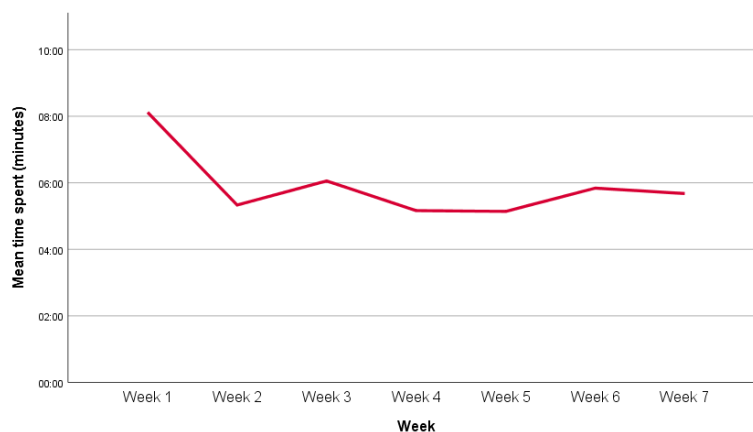

Figure 11: Average time spent on MyST by the students (N=11) per Week

The total number of participants considered for the remaining part of the log data analysis (N=17) is different from the paired t-test analysis (N=11). The aim of this study was understanding the behaviour of the students who used MyST, and we believe that including the

23

6 students who did not complete the post-test were would present a better overview of the use of MyST.

Figure 12 below illustrates the average time spent, in minutes, by all the students per week. Overall there is no noticeable difference if we include the 6 students who did not complete the pre and post-tests. The time spent is higher in week 1 than for the other weeks of the course. Differently from what we can observe in figure 11, if we consider the 17 students, the average is less than 8 minutes. From week 2 to week 7, the time spent remained under 6 minutes both in the graphs in figure 11 and 12.
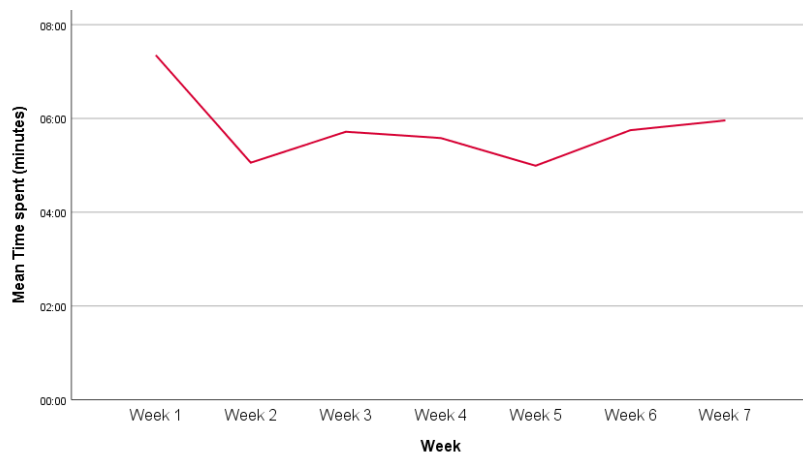
Figure 12: Average time spent on MyST by the students (N=17) per Week

Differently from the average time spent every week, the number of students who practiced with MyST decreased considerably. Figure 13 exemplify this difference.

As it will be discussed in detail later, the reason for this change can be attributed to the fact that most of the students have lost motivation to continue in using the application over the course of the 7 weeks. Moreover, the students might have stopped using MyST because they had less time or because they forgot.

Figure 13: Time spent and number of participants (N=17) per week

In section 3.2.2 we described the different kind of activities included in the course: Pronunciation, Gap Filling, Multiple choice, Drag and Drop, Audio choice. These five types were analysed separately comparing the time spent in relation to the percentage of the correct answers. Table 3 summarize the information used for the analysis. There were between 10 and 12 activities every week, the table below illustrates the number of each type of activity that the students encountered during the duration of the course. The time spent refers to the average period that the students took to finish each activity and it is expressed in seconds. The activity with the highest percentage of correct answers (86.12%) was multiple-choice, which was also the most frequent in the course. Compared to the other activities, the students scored worse in the pronunciation task (52.81%).

Table 3: Information used for the analysis of the activities: number, time spent and percentage of correct answers

Activity	Number of activities	Average time spent (seconds)	Percentage of correct answers
Pronunciation	17	00:29	52.81%
Multiple choice	33	00:18	86.12%
Gap filling	21	00:22	69.48%
Drag and drop	6	00:31	79.60%
Audio choice	5	00:28	83.00%

Figure 14 below indicates the comparison between the percentages of correct answers and the average time spent for each type of activity. Whereas for audio choice, drag and drop, gap filling the difference between the time spent and the percentage of correct answers is not large, for the multiple choice and pronunciation tasks it is. Considering multiple choice, the students spent less time and generally had better results. On the contrary, for the pronunciation tasks, the percentage of correct answers was lower, but they spent more time on the task.
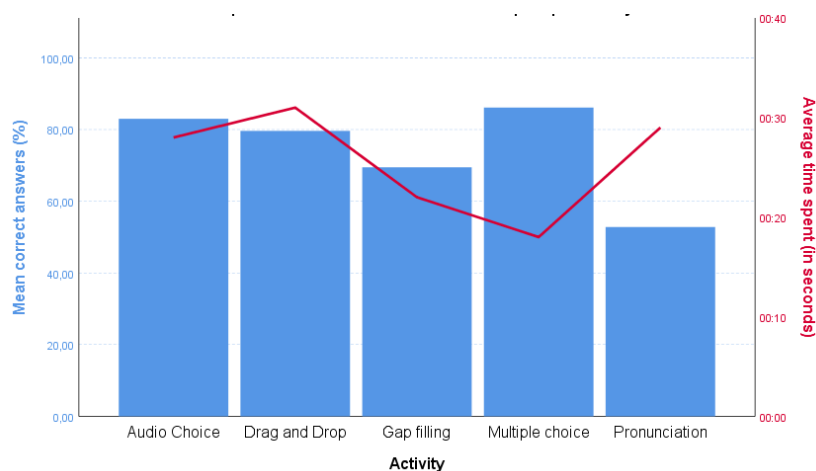

Figure 14: Comparison between percentage of correct answers and time spent per activity

In both groups, the analytics included information regarding the number of answers given for each question. This information was analysed and compared with the number of students who completed the activities. Figure 15 below illustrates this comparison per activity, and the data confirm the results presented in the chart in figure 14 above. In fact, the most obvious difference appears between the pronunciation task and other activities. For the 14 participants who completed the pronunciation task, 26 answers were registered. This happened either because the students were pronouncing the target words poorly, and they made many mistakes, or because the speech recognizer was not accurate. In fact, in Novo Studio platform, the correct and incorrect pronunciation needed to be manually inserted by the instructor when designing the activities. If there were mistakes in the options entered, the ASR system might have identified errors when there were not. Figure 15 further shows that the type of activity that required less trials was Multiple Choice.
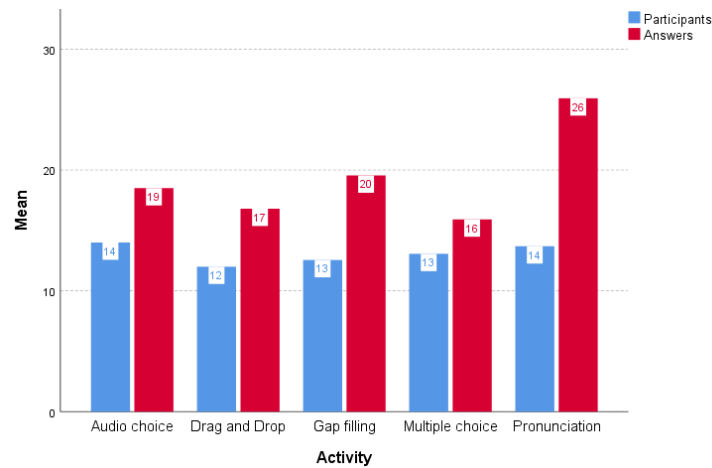
Figure 15: Comparison between answers registered and number of participants per activity

Figure 16 illustrates the comparison between the answers registered and the participants per week. Major differences between the students who completed the activities and the answers registered were found in week 1, 3 and 5. For example, in week 1, the total number of participants was 17, but the mean of the answers given for each exercise was 31. This signifies that, on average, the students had to try each activity two times or more before giving the correct answer. In week 1 there was a higher number of pronunciation type (5), which, as we saw in figure 15, was more difficult for the students. Minor differences are found in week 2, 4, 6 and 7.

Figure 15 and 16 both showed that there was a significant difference between the participants and the answers registered. The reason for this difference can be that the activities' instructions were not clear enough or that the level was too high compared to the student' knowledge. In the next chapter we will present the results of the analysis for the Phonetics group.
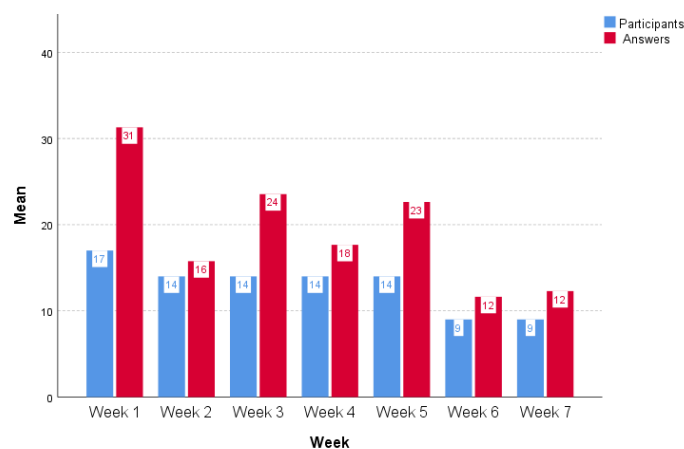

Figure 16: Comparison between answers registered and number of participants per week

4.2.2 Phonetics group

Differently to what we observed in the Semantics and Pragmatics group, for the Phonetics group the time spent per week decreased from an average of 9 minutes in Week 1 to less than 4 minutes in week 7. The graph in figure 17 shows the average number of minutes spent on MyST by the 21 students for the 7 weeks of the course.
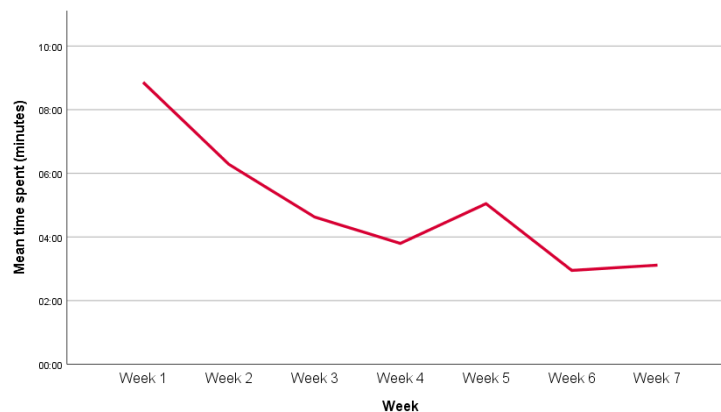
Figure 17: Average time spent on MyST by the students (N=21) per week Phonetics group

As in the Semantics and Pragmatics group, many of the students enrolled in the Phonetics course finished at least one activity (N=30), but only a few among them completed both the pre and post-test (N=21). To present a complete overview of the use of MyST, the analysis of the log data includes the information from the students who finished at least one activity on MyST.

The 9 students who did not complete the pre or the post-test were also included, and the average time spent on the application was calculated again, figure 18. If we compare the graph with the group including only the participants of the pre and post-test (N=21), we do not notice many differences. The time spent in week 1 is higher if we consider the 30 students, around 10 minutes, compared to the group of 21 students. However, for both groups, the time spent significantly decreased from week 1 to week 7.
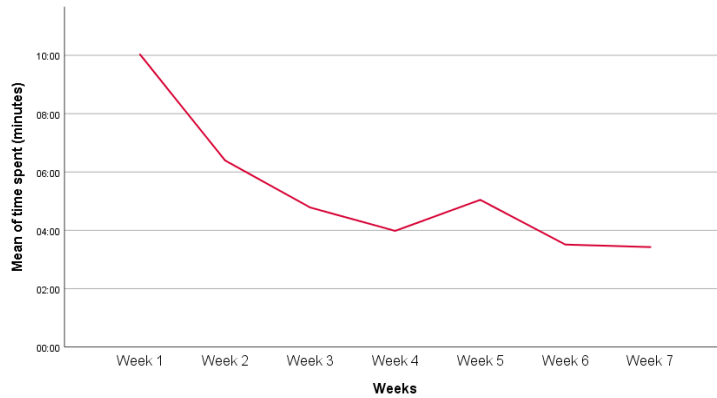
Figure 18: Average time spent on MyST by the students (N=30) per week Phonetics group

For the following analysis we considered the data of the students who finished at least one activity (N=30). It was noted that not only the time the students spent on the application drastically decreased every week, but the number of participants was also different. Figure 19 illustrate the comparison between the participants and the average of the time registered per week. In week 1 the students who participated were 30 and they spent on average 10 minutes to complete the activity. However, the number of students in week 6 and 7 decreased to 8.
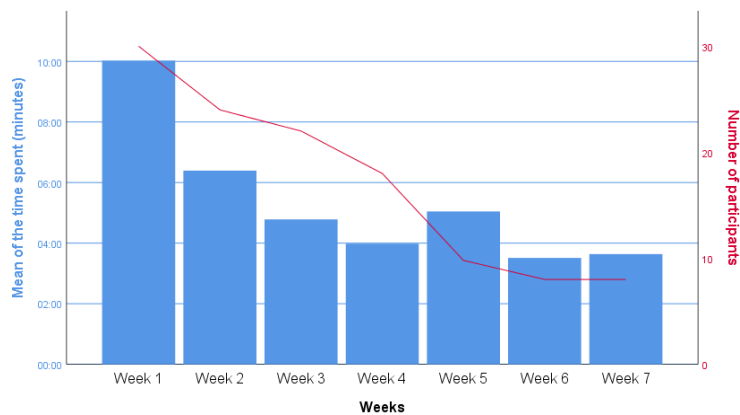
Figure 19: time spent and number of participants (N=30) per week Phonetics group

An analysis of the different types of activites was conducted also for the Phonetics group. Table 4 below illustrates the number of activites found in the Phonetics course. For each the time spent and the percentage of correct answers was calculated. The activities included Pronunciation, Multiple choice, Gap filling and Drag and drop, as described in section 3.2.2, but no Audio choice, and they were distributed differently compared to the Semantics and Pragmatics group. Most of the exercises consisted in Multiple-choice (28) and Gap filling (25), but there were only 4 Pronunciations tasks and 7 Drag and drop. However, it is important to

note that in both courses the students had to pronounce the correct answer for the Multiple-choice exercises, so they would still practice speaking skills.

Table 4: Information used for the analysis of the activities: number, time spent and percentage of correct answers

Activity	Number of activities	Average time spent (seconds)	Percentage of correct answers
Pronunciation	4	00:15	69,75%
Multiple choice	28	00:30	81,56%
Gap filling	25	00:26	73,61%
Drag and drop	7	00:50	75,57%

This data was analysed similarly to the Semantics and Pragmatics group to understand the behaviour of the students when using MyST. First, we compared the percentage of correct answers given by the students with the time they spent on each in figure 20. Overall, the students answered correctly spending a short time on the activities, except for Drag and Drop.
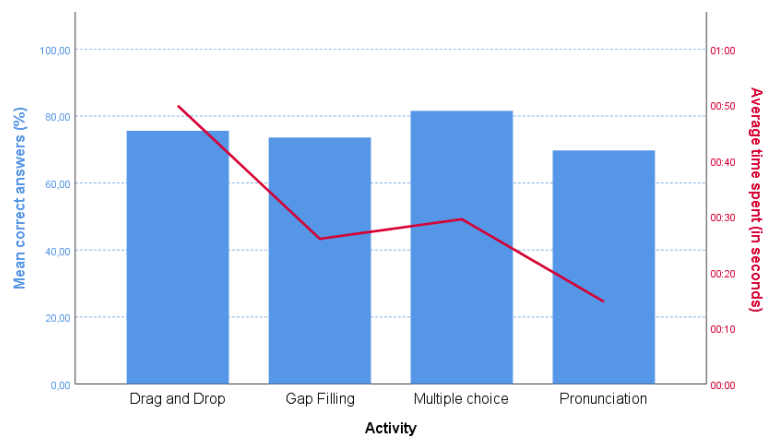
Figure 20: Comparison between percentage of correct answers and time spent per activity

The next step of the analysis was the comparison between the number of students who used the application and the number of answers registered. Overall, in every activity, the answers given were always more than the number of participants. There can be many reasons why the students repeated the same activity more than once. First, the instructions might not have been clear enough, so they might not have understood the activity at first. Moreover, if

the students made a mistake, they also had the possibility to try again. Even if there were few pronunciation tasks, compared the Semantics and Pragmatics group, there was still a great difference between the answers registered (21) and the students who completed that activity (10).
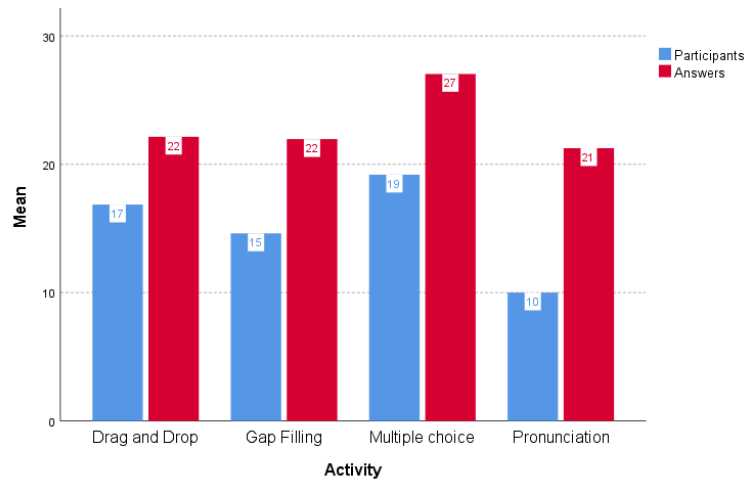
Figure 21: Comparison between answers given and number of participants per activity

This comparison was also made per week, and not only per activity. The bar graph in figure 22 shows that the students repeated the same activity more times in week 1 and 2 compared to following weeks. This data is in line with what was found for the course of Semantics and Pragmatics. If the students did not understand the tasks assigned or they were too difficult, they might have been less motivated to continue using MySpeechTrainer. This can be related to the data presented in figure 17 and 18 above.
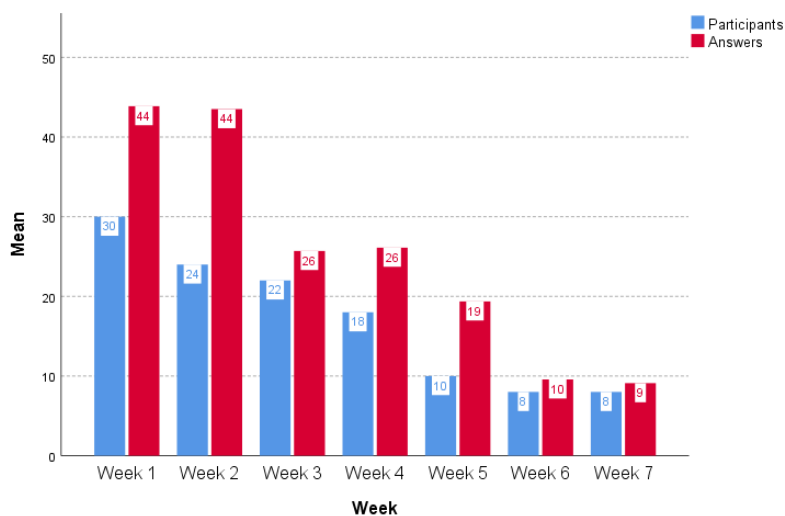
Figure 22: Comparison between answers given and number of participants per week

Chapter 5

Discussion and Conclusion

This study investigates the learning outcomes and the behaviour of the students enrolled in the courses of Semantics and Pragmatics, and Phonetics at Utrecht University using the prototype application MySpeechTrainer. The results of the pre and post-test and the log data analysis provided answers for the two research questions.

The results of the paired t-test between the mean of the pre and post-test scores did not show any significant difference. A reason for not finding any improvement after the use of MyST was the small sample size. During the first lecture, the students were asked to complete a pre-test, and after seven weeks, at the end of the course, the post-test was carried out. Almost all the students completed the pre-test, however, only 11 for the Semantics and Pragmatics and 21 for the Phonetics group finished the post-test. The reason for such large difference was that some of the students dropped out of the course, others never used their MyST account after the first day of class, and still, others were absent on the last day when the post-test was distributed. As it was discussed before, the use of MyST was in the syllabus, and it was part of the class assignments for both groups. However, correctly finishing the activities was not a condition for the final grade, and it is possible that this affected the motivation of the students to use the application. Moreover, as it was showed in section 4.1, many of the students received high scores in the pre-test, so they already had an advanced knowledge on the topic before starting the training. As we said, a small variance was found between the pre and post-test scores, which means that the use of MyST had no effect on the students who scored 6 or 7. In summary, to answer the first research question, the pre and post-test scores did not show a significant improvement in vocabulary skills using MySpeechTrainer.

They were all undergoing the same experimental condition, which was the use of MyST, and it was not feasible to include a control group. It was, in fact, not possible to find another class with the same conditions, meaning in English and on the same topics. Moreover, as the teachers included the use of MyST as one of the assignments of the classes, but without a grade, it was not possible to divide them into two groups.

The log data from Novo Studio platform was analysed to answer the second research question. The data showed that the use of the application, as well as the average time spent on each exercise, significantly decreased after the first week of class for the Phonetics group.

However, for the Semantics and Pragmatics group, the time slightly decreased from week 1 to 2 but it remained constant for the rest of the course. Overall, the time spent on MySpeechTrainer never exceeded 10 minutes. As Beck (2004) explains, even if the material proposed is pedagogically appropriate if the time spent on the activity is too low, learning might still not be efficient.

Therefore, in the case of MySpeechTrainer, the limited time spent on the application might have affected the students' learning. Additionally, this significant decrease in practice in the Phonetics group could reveal a loss of interest and motivation. In another study from Cocea and Weibelzahl (2006), the authors concluded that the students' actions, registered with the log data, reflected the level of students' motivation. Motivation is, in fact, an important factor for e-learning and study in autonomy. As the authors suggested, identifying the reasons for low motivation can help the teachers and designer to limit the drop-outs (pag.1). As it was illustrated in the analysis of the results, in many activities the students tried the same exercise many times, particularly the pronunciation tasks, before giving the correct answer. It is possible that their motivation to continue lowered every time their answer was incorrect. Furthermore, Wiebe and Kabata (2010) found that the role of the instructor has a significant effect on the student's behavior on the application. Their log-in frequencies are affected by the use of the teacher of "judiciously placed reminders and encouragement".

At the end of the seven weeks of training, the students in both groups completed a questionnaire. There were two open and optional questions at the end[4]:

- "Would you like to share something else about your experience with My Speech Trainer? If you have any suggestions or remarks, please write them down here. For example, what is one thing you liked best / least about MyST?"
- "If you have not completed any exercises in My Speech Trainer, could you share with us why?"

Most of the students declared that what they liked the most about the MyST was the variety of activities and the fact that they could practice the topic that they were studying in class. One of the students described the exercises 'engaging'. In almost all the answers, the students claimed that the ASR system did not recognize their pronunciation and. As it was illustrated in chapter 4 of this thesis, in the Semantics and Pragmatics group, the students scored low grades and spent a long time to complete the pronunciation tasks. These results can indeed

---

[4] The list of the answers of both groups can be found in Appendix C.

show a malfunction in the microphone setting, or the fact that the students did not know the correct pronunciation and they needed to practice and try more.

In summary, in the future, to ensure an efficient learning environment, some aspects need to be considered. First of all, it is important to keep the students engaged in the activities throughout the course. In the future, as some activities required more time and seemed to be more complicated for the students, we should pay particular attention to the type of feedback and the difficulty of the words used. Future studies could also examine the Pronunciation task in MySpeechTrainer, comparing the answer given by the student with the feedback received. This type of analysis could help predict what type of errors students make more often. Furthermore, future research should also focus on the methods to improve the students' time spent on the tasks, as for MySpeechTrainer it seemed to have influenced the learning outcomes.

References

Ahmad, K. (2016). Integrating Pronunciation with Vocabulary Skills from Jones, T. (Ed.). (2016) *Pronunciation in the classroom: The overlooked essential*. Tesol Press.

Beatty, K. (2013). *Teaching & researching: Computer-assisted language learning*. Routledge.

Beck, J. E. (2004). Using response times to model student disengagement. *Proceedings of the ITS2004 Workshop on Social and Emotional Intelligence in Learning Environments* (Vol. 20).

Brooks, M. (2014). The role of vocabulary in English for Specific Purposes (ESP) teaching and learning: Considerations for Asia University. *CELE Journal*, (22), 157-175.

Chapelle, C., & Mizuno, S. (1989). Student's strategies with learner-controlled CALL. *Calico Journal*, 25-47.

Chapelle, C. A. (2001). *Computer applications in second language acquisition: foundations for teaching, testing and research*. Cambridge University Press.

Chinnery, G. M. (2008). Biting the hand that feeds me: The case for e-language learning and teaching. *CALICO Journal*, *25*(3), 471-481.

Chiu, T. L., Liou, H. C., & Yeh, Y. (2007). A study of web-based oral activities enhanced by automatic speech recognition for EFL college learning. *Computer Assisted Language Learning*, *20*(3), 209-233

Cocea, M., & Weibelzahl, S. (2006). Can log files analysis estimate learners' level of motivation?. *LWA*. University of Hildesheim, Institute of Computer Science.

Felix, U. (2005). Analysing recent CALL effectiveness research – Toward a common agenda. *Computer Assisted Language Learning*, 18(1–2): 1–32.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer assisted language learning*, *27*(1), 70-105.

Hubbard, P. (2006). Evaluating CALL software. *Calling on CALL: From theory and research to new directions in foreign language teaching*, 313-338.

Hubbard, P. (2009). A general introduction to computer-assisted language learning. *Computer Assisted Language Learning: Critical Concepts in Linguistics*, *1*(1), 1-20.

Hwu, F. (2003). Learners' behaviors in computer-based input activities elicited through tracking technologies. *Computer Assisted Language Learning*, *16*(1), 5-29.

Isti'anah, A. (2017). The Effect of Blended Learning to the Students' Achievement in Grammar Class. *IJEE (Indonesian Journal of English Education)*, *4*(1), 16-30.

Jamieson, J., Chapelle, C. A., & Preiss, S. (2005). CALL evaluation by developers, a teacher, and students. *CALICO journal*, 93-138.

Kebede, D., & Getachew S. (2015). The implementation of language skills integration in teaching and learning English as a foreign language (EFL): Jimma College of teachers' education in focus. *Nawa: Journal of Language & Communication* 9.1

Manik, S., & Christiani, M. (2016). Teaching vocabulary using matching word on computer assisted language learning. *International Journal of English Language Teaching*, *4*(7), 1-26

Milla, R., & Mayo, M. P. G. (2014). Corrective feedback episodes in oral interaction: A comparison of a CLIL and an EFL classroom. *International Journal of English Studies*, *14*(1), 1-20

Nation, I.S.P. (2001*). Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer assisted language learning*, *15*(5), 441-467.

Plonsky, L., & Ziegler, N. (2016). The CALL–SLA interface: Insights from a second-order synthesis. *Language learning & Technology*, *20*(2), 17–37. Retrieved from http://llt.msu.edu/issues/june2016/plonskyziegler.pdf

Seregély, E. M. (2008). *A comparison of lexical learning in CLIL and traditional EFL classrooms* (Doctoral dissertation, uniwien). Last viewed 16/05/2019. https://celt.univie.ac.at/fileadmin/user_upload/proj_fdz_englisch/Projects_and_Research/Theses/seregely_eva.pdf

Wiebe, G., & Kabata, K. (2010). Students' and instructors' attitudes toward the use of CALL in foreign language teaching and learning. *Computer Assisted Language Learning*, *23*(3), 221-234.

Appendices

Appendix A

WORDS TRAINED IN MyST

n	Phonetics	Semantics and Pragmatics
1	velum	defeasible
2	vocal folds	metonym
3	larynx	polysemy
4	Manner and Place of articulation	implicature
5	airstream	coercion
6	pharynx	reference
7	voiced	inference
8	voiceless	entity
9	assimilation	denotation
10	syllable	connotation
11	onset	exemplar
12	coda	feature
13	amplitude	typicality
14	frequency	prototype
15	resonance	synonym

16	formant	anaphora
17	oscillogram	quantifier
18	spectrogram	type
19	spectrum	definite
20	harmonics	modifier
21	epenthesis	ambiguity
22	coarticulation	dynamic
23	assimilation	entailment
24	metathesis	composition
25	allophone	attribute
26	periodic	coercion
27	aperiodic	implicature
28	cochlea	cardinality
29	eardrum	homogeneity
30	normalization	
31	pitch	
32	prosody	
33	intonation	
34	stress	
35	rhythm	

PRE-TEST PHONETICS

Vocabulary

A. Multiple choice

1. When a sound is produced with a vibration of the vocal folds is called _____.

a. voiced

b. nasal

c. voiceless

d. sonorant

Answer: a

In MySpeechTrainer

2. Two different sounds cause a change of meaning when replaced in the same environment are _____.

a. constraint

b. underspecified

c. matching

d. contrastive

Answer: d

3. Considering a periodic waveform, the higher the number of cycles per second, the higher the _____ is.

a. Frequency

b. Resonance

c. Harmony

d. Amplitude

Answer: a

In MySpeechTrainer

4. The contraction from can not to can't involves a phenomenon that is called _____

a. assimilation

b. elision

c. attrition

d. ellipsis

<div align="right">Answer: b</div>

5. Some sounds are _____ and recognizable from a short distance and can be heard.

a. imperceptible

b. sonorant

c. audible

d. obstructed

<div align="right">Answer: c</div>

6. At the top of the trachea, containing elastic vocal folds, is found the _____.

a. larynx

b. pharynx

c. trachea

d. diaphragm

<div align="right">Answer: a</div>

<div align="right">In MySpeechTrainer</div>

7. Often in English, the only difference between a verb and a noun is the _____, as the words "object" or "record".

a. stress

b. pitch

c. intonation

d. accent

<div align="right">Answer: a</div>

<div align="right">In MySpeechTrainer</div>

8. Consonants differ from vowels as they are produced with almost always a _____, or obstacle of air.

a. movement

b. vibration

c. obstruction

d. attrition

<div align="right">Answer: c</div>

9. The speech sound is not clear and open to all sorts of interpretations, it's said to be
_____.

a. Ambiguous

b. Anaphorical

c. Obvious

d. Synonymous

Answer: a

10. A combination of different _____, the characteristics of the sounds, defines a phoneme.

a. features

b. attributes

c. stress

d. accents

Answer: a

B. Fill the gaps

11. In phonetics, _____ is known as the insertion of an extra sound in a word. Answer: Epenthesis. (In MySpeechTrainer)

12. All speakers have a different average_____, which is the adjusted by changing the tension of the vocal folds. Answer: Pitch (In MySpeechTrainer)

13. In the word café, the _____ falls on the final syllable. Answer: accent

14. Historical linguistics is a _____ study as it considers evolution of language through time. Answer: Synchronic

15. The effects of _____ involve articulatory modification of a speech sound, influenced by the preceding or following speech sound. Answer: coarticulation (In MySpeechTrainer).

Pronunciation

A. Pronounce the sentences

1. Syllables are formed by a nucleus and optional surrounding consonants.

2. I believe that the response could be affirmative.

3. <u>Intonation</u> describes how the voice rises and falls in speech.

4. A <u>feature</u> is a distinctive attribute or characteristic of something.

5. The <u>rhythm</u> of language is important in every utterance.

POST-TEST PHONETICS

Vocabulary

A. Multiple choice

1. The _____ represents the number of cycles of vibration given a period of time.

a. amplitude

b. period

c. resonance

d. frequency

<div align="right">Answer: d</div>

<div align="right">In MySpeechTrainer</div>

2. In phonology, the _____, symbolized with μ, is the unit that determines the syllable weight.

a. mora

b. onset

c. nucleus

d. stress

<div align="right">Answer: a</div>

3. An _____ is the use of an expression that depends upon an antecedent expression.

a. cataphora

b. synonym

c. anaphora

d. metaphor

<div align="right">Answer: c</div>

4. The faster the vibration of the vocal folds, the higher the _____ of the voice

a. pitch

b. signal

c. prosody

d. resonance

5. Vowels sounds are also called _____, as they are produced with non-turbulent airflow in the vocal tract.

a. articulatory

b. obstruent

c. sonorant

d. voiceless

6. There is no _____ vibration when producing voiceless sounds, but there is with voiced sounds.

a. vocal folds

b. larynx

c. velum

d. diaphragm

7. In a monosyllabic word as a *cat,* the _____ is /k/, because it precedes the nucleus of the syllable.

a. syllable

b. coda

c. onset

d. nucleus

8. The _____ of speech refers to the mapping of sounds into linguistic representations.

a. perception

b. hearing

c. discrimination

d. physiology

<div align="right">Answer: a</div>

9. /fəˈnɛtɪks/ is the _____, or representation, of a speech sounds by means of phonetic symbols.

a. transcription

b. recording

c. graphic

d. frequency

<div align="right">Answer: a</div>

10. A _____ is a language variety spoken by a group of people who live in a particular place.

a. variation

b. dialect

c. accent

d. register

<div align="right">Answer: b</div>

B. Fill the gaps

11. The same person uses a different _____ when delivering a formal speech compared to when spending time with friends. Answer: register.

12. The graph called spectrogram represents the frequencies of a signal changing in time. Answer: spectrogram (MySpeechTrainer).

13. The word _____ refers to the insertion of one or more segment in a word. Answer: epenthesis (MySpeechTrainer).

14. The word *loud* is formed by a _____, in a single syllables two vowels are combined. Answer: diphthongs.

15. In phonetics, the term prosody is an element of language that contributes to rhythmic and acoustic effects. Answer: prosody (MySpeechTrainer).

Pronunciation

A. Pronounce the sentences

1. <u>Lenition</u> is a type of alternation where the sound becomes weaker or more open.

2. The <u>cochlea</u> is part of the inner ear and is fundamental for hearing. (MySpeechTrainer)

3. The <u>rhythm</u> and beat of language are important in every utterance. (MySpeechTrainer)

4. The <u>diaphragm</u> is the largest muscle used in speech.

5. The problem of <u>normalization</u> is that listeners can perceive different signals as the same word. (MySpeechTrainer)

PRE-TEST SEMANTICS AND PRAGMATICS

Vocabulary

A. Multiple choice

1. The assumptions that hearers make about a speaker's conduct give rise to different types of _____.

a. Inference

b. Proposition

c. Prepositions

d. Facts

Answer: a

In MySpeechTrainer

2. In the sentence *my cousin said she was coming*, *she* is used as an _____ for *my cousin.*

a. Cataphora

b. Anaphora

c. Anaphorical

d. Ellipsis

Answer: b

In MySpeechTrainer

3. Different sentences allow the speaker to do different things with the same _____.

a. Proposition

b. Presuppositions

c. Preposition

d. Proposal

Answer: a

4. Semantic _____ allows a word to show up in contexts in which it would otherwise not appear.

a. Coersion

b. Inversion

c. Coercion

d. Conversion

Answer: c

In MySpeechTrainer

5. The _____ meaning of a word is perceived through visible concepts.

a. Denotational

b. Connotational

c. Denotation

d. Symbolic

Answer: a

In MySpeechTrainer

6. The more similarities that people share with each other, the more likely for them to _____ in a conversation.

a. Accomodate

b. Accommodate

c. Accomodation

d. Acomodate

Answer: b

7. In grammar, if a word or phrase is not required by the verb is called _____.

a. Argument

b. Complement

c. Predicate

d. Adjunct

8. On these considerations we must be place an equal _____.

a. Emphasis

b. Enphasis

c. Emphases

d. Emphasy

9. This text is open to all sorts of interpretations, it's _____ and contradictory.

a. Ambiguous

b. Synonymous

c. Anaphorical

d. Antonymous

In MySpeechTrainer

10. In my dictionary, "work" and "leisure" are _____.

a. Antonymous

b. Antonymy

c. Synonymy

d. Similar

Answer: b

B. Fill the gaps

11. In familiar terms, _____occurs when two or more words are joined to make one longer word. Answer: Compound.

12. the tasks to which you are assigned will depend on your_____. Answer: capabilities)

13. I have agreed to help them build their _____. Answer: Prototype (MySpeechTrainer)

14. You seem _____concerned about me. Answer: genuinely

15. This implication is easily _____. Answer: defeasible (MySpeechTrainer)

Pronunciation

A. Pronounce the sentences

1. Sometimes the voice of a verb is called also <u>diathesis</u>

2. I believe that the response could be <u>affirmative</u>

3. In a <u>polysemous</u> developments, a word can take on an opposite meaning. (MySpeechTrainer)

4. The adjective *sunny* is used <u>Attributively</u> in *a sunny day*. (MySpeechTrainer)

5. We will remain here for an <u>indefinite</u> period of time. (MySpeechTrainer)

POST TEST SEMANTICS AND PRAGMATICS

Vocabulary

A. Multiple choice

1. The characteristic of presuppositions that can be cancelled is known as _____.

a. Defeasibility

b. Defesability

c. Defisability

d. Defisable

Answer: a

In MySpeechTrainer

2. *Up to* in numerals is severely constrained as it cannot indicate a definite _____**.**

a. Dominance

b. Modifier

c. Cardinality

d. Quantifier

Answer: c

In MySpeechTrainer

3. Often different formal analyses share common _____about the goals of semantics.

a. Assumptions

b. Implications

c. Informations

d. Differences

<div align="right">Answer: a</div>

4. Presupposition failure is an extreme case of _____ use of presuppositions

a. Infelicituos

b. Infelicity

c. Inapropriate

d. Infelicitous

<div align="right">Answer: d</div>

5. Depending on the people, the word *home* has many _____, such as "warmth," "security".

a. Denotations

b. Connotations

c. Implications

d. Collocations

<div align="right">Answer: b</div>

<div align="right">In MySpeechTrainer</div>

6. The speaker's confidence that the referent is accessible to the speaker is reflected by the use of the _____.

a. Definite

b. Indefinite

c. Connotation

d. Synonym

<div align="right">Answer: a</div>

<div align="right">In MySpeechTrainer</div>

7. More concrete sources to describe a more abstract target are typically used by _____.

a. Asymmetry

b. Metaphorical

c. Metaphors

d. Concepts

<div align="center">50</div>

8. In some construction quantifying determiners, as *every, most,* may result _____

a. Anomalous

b. Anomalus

c. Anoumalous

d. Amoumalos

Answer: a

9. A mix of criteria are necessary to _____ the types of speech act possible in languages

a. Reject

b. Constrain

c. Manifest

d. Establish

Answer: d

10. Metaphors do not set up a _____ comparison between two concepts.

a. Symmetrical

b. Symmetry

c. Simmetrical

d. Symetrical

Answer: a

B. Fill the gaps

11. *Up to* can be shown to be compatible with spatially _____predicates only. Answer: homogeneous (MySpeechTrainer).

12. Within the range of each color term there is a basic focal color that speakers agree to be the best _____example of the color Answer: prototypical (MySpeechTrainer).

13. The ways in which a speaker relates references to space and time are _____systems. Answer: deictic.

14. A pronoun, as it, can refer _____ to an NP. Answer: anaphorically (MySpeechTrainer).

15. The alternatives to the formal semantics' approaches don't offer the same_____. Answer: comprehensiveness.

Pronunciation

A. Pronounce the sentences

1.An <u>Attributive</u> use of an adjective is when it modifies a noun (MySpeechTrainer)

2.An accomplishment is a situation type that is <u>dynamic</u>, durative and telic (MySpeechTrainer)

3.<u>Metonymy</u> is as an important producer of polysemy across languages (MySpeechTrainer)

4. And, or, but act like <u>conjunctions</u> by linking two main clauses or two sentences

5. Metonymy expresses simple <u>contiguous</u> relations between objects

Appendix C

STUDENTS' COMMENTS

Question 23: If you have not completed any exercises in My Speech Trainer, could you share with us why?

- Forgot, and the app often doesn't recognize my voice so it's a bit unpractical
- I often forgot to use MyST. If the app sent notifications, I would have use it more than now.
- Still need to read the materials for that week
- Because I didn't feel like I was learning anything.
- I did not have enough time next to the existing exercises of the course, and it was not clear to me that after the starting test, we had to do this every week.
- I forgot about it to be honest, i will probably use the exercises for the test!

Question 24: Would you like to share something else about your experience with My Speech Trainer? If you have any suggestions or remarks, please write them down here. For example, what is one thing you liked best / least about MyST?

- I liked the variety of exercises (not just speaking but also filling in the blank). However, for some of the speaking exercises, MyST did not find my answer correct when it was in fact correct English (I even tried an American and British accent). The stress was maybe differently implemented in the app? I would also like to get reminders/notifications when I have open exercises to do.

- It could be useful to receive some more feedback when the given answer is not the right one, now the app only tells you the answer is incorrect.

- I would recommend allowing for rhotic pronunciation, as is the norm in American English and Irish English.

- Because I do not have any trouble with speaking and understanding English, I believed the exercises were not so challenging. I am sorry to say that the exercises did not help me any further in my study program or in speaking English. However, the exercises were engaging. Suggestions for improvements are: - Do not train the users only on the relevant terms that must be learned in the course but give them full exam questions that they have to answer, corresponding to the level of the course. Give them personal feedback on the content of their answers. -Introduce a platform in which the participants of the course can anonymously ask questions and answer questions of others. -Add to each option of the app a 'spelling corrector/speech corrector'at which you give feedback on the participants' use of English.

- Sometimes the speech recognizer didn't recognize my pronunciation

- The microphone function isnt always correct. If i look at the fonetics, what i said and the answers overlap, but it is marked as wrong

- I think this application is nice and fun to improve your language But also to cover/test test material. The thing that could be improved was that the application while I had to pronounced often did not work properly. I retried often But it kept on with this difficulty. I am not sure whether this was due to my computer. It did often do work however.

- I think MyST is very useful app for learning language and contents of a university course at the same time. However, something I missed was the possibility to get the correct answer to a question when I did not know / got the answer wrong (as far as I know this was not a possibility). If I don't know the answer to a question or the correct pronunciation of a word and I've gotten the question wrong (once or perhaps multiple times), I would like to know the correct answer so I am able to learn the word.

- The "previous" button on the bottom of the screen doesn't work. The app often doesn't recognize your voice.

- getting a notification when you should make the next exercise

- I don't see how it should help with learning English, because you don't get any feedback of your pronunciations.